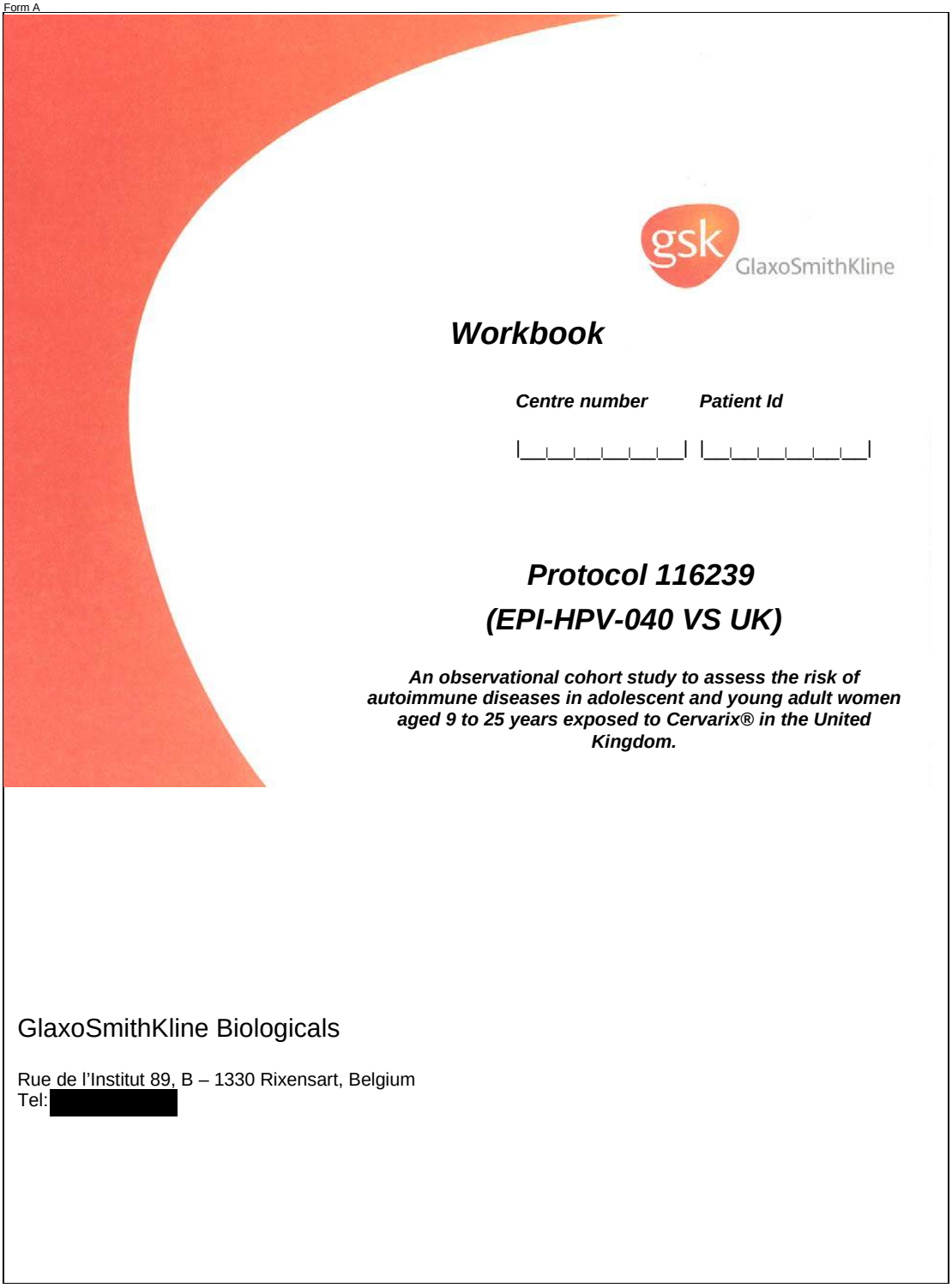
Form A



Workbook

Centre number ***Patient Id***

|__|__|__|__|__|__| |__|__|__|__|__|__|

Protocol 116239
(EPI-HPV-040 VS UK)

An observational cohort study to assess the risk of autoimmune diseases in adolescent and young adult women aged 9 to 25 years exposed to Cervarix® in the United Kingdom.

GlaxoSmithKline Biologicals

Rue de l'Institut 89, B – 1330 Rixensart, Belgium
Tel: [redacted]

CRF version template 13.2 – October 09, 2013

GENERAL INSTRUCTIONS

ABBREVIATIONS
Abbreviations for medical conditions, clinical events or drug names are to be avoided.

ERRORS/CORRECTIONS
Errors should be crossed out with a single line and the alteration made as near to the original as possible. All alterations must be printed, initialled and dated by the investigator or authorised staff.

DATES
Use the following 3-letter abbreviations to indicate months:

January	=	JAN
February	=	FEB
March	=	MAR
April	=	APR
May	=	MAY
June	=	JUN
July	=	JUL
August	=	AUG
September	=	SEP
October	=	OCT
November	=	NOV
December	=	DEC

Example: |**0 1**| |**J A N**| |**2 0 0 8**| = 1st January 2008

MISSING DATA CODES
Preferably use following codes:

ND = Not Done
NA = Not Applicable
NK = Not Known

CRF version template 13.2 – October 09, 2013

116239 (EPI-HPV-040 VS UK)

				Patient ID
		General Information		|__|__|__|__|__|__|__|__|

	CPRD	PALLAS
Cohort Number	|__|__|	
Gender*	|__|__|__|	
Birth Month*	|__|__|__|__|__|__|__|__|	
Birth Year*	|__|__|__|__|__|__|__|__|	
Practice Identifier*	|__|__|__|	
Practice Region*	|__|__|__|	
CPRD GOLD start date	|__|__| |__|__|__| |__|__|__|__| day month year	
Death Date*	|__|__| |__|__|__| |__|__|__|__| day month year	
First registration date*	|__|__| |__|__|__| |__|__|__|__| day month year	
Current registration date*	|__|__| |__|__|__| |__|__|__|__| day month year	
Registration gaps*	|__|__|__|__|__|__|__|__|	
Registration status*	|__|__|__|	
Up to standard date*	|__|__| |__|__|__| |__|__|__|__| day month year	
Acceptable quality standard met by patient*	|__|__|__|	
Transfer out date*	|__|__| |__|__|__| |__|__|__|__| day month year	
Transfer out reason*	|__|__|__|	
CPRD-HES link*	☐ No ☐ Yes	
Number of primary care resource utilization	|__|__|__|__|__|__|__|__|	
Reference Date	|__|__| |__|__|__| |__|__|__|__| day month year	
Date of end of the study period	|__|__| |__|__|__| |__|__|__|__| day month year	
Date of de-enrolment from CPRD GOLD*	|__|__| |__|__|__| |__|__|__|__| day month year	
Lost to follow-up		☐ No ☐ Yes: Please provide the date of lost to follow-up: |__|__| |__|__|__| |__|__|__|__| day month year
Comments		_____

*: data coming from CPRD GOLD

1.

CRF version template 13.2 – October 09, 2013

GlaxoSmithKline

116239 (EPI-HPV-040 VS UK)

				Patient ID
		HPV Cervarix Vaccination		|__|__|__|__|__|__|__|__|__|

Vaccination: From one year before the reference date and until the end of the follow-up

	CPRD	PALLAS
Is HPV Cervarix Vaccination administered?	☐ No ☐ Yes	☐ No ☐ Yes: please provide details below
Number of HPV Cervarix doses administered	|__|__|	|__|__|
Date of HPV Cervarix vaccination 1	|__|__| |__|__|__| |__|__|__|__| day month year	|__|__| |__|__|__| |__|__|__|__| day month year or ☐ not available
Date of HPV Cervarix vaccination 2	|__|__| |__|__|__| |__|__|__|__| day month year	|__|__| |__|__|__| |__|__|__|__| day month year or ☐ not available
Date of HPV Cervarix vaccination 3	|__|__| |__|__|__| |__|__|__|__| day month year	|__|__| |__|__|__| |__|__|__|__| day month year or ☐ not available
Date of HPV Cervarix vaccination 4	|__|__| |__|__|__| |__|__|__|__| day month year	|__|__| |__|__|__| |__|__|__|__| day month year or ☐ not available
Comments		_____

2.

GlaxoSmithKline

116239 (EPI-HPV-040 VS UK)

				Patient ID
		Other Vaccinations		|__|__|__|__|__|__|__|__|

Vaccination: From one year before the reference date and until the end of the follow-up

	CPRD	PALLAS
Is HPV unspecified or Gardasil Vaccination administered?	☐ No ☐ Yes	☐ No ☐ Yes
Date of unspecified or Gardasil HPV vaccine 1	|__|__| |__|__|__| |__|__|__|__| day month year	|__|__| |__|__|__| |__|__|__|__| day month year or ☐ not available
Date of unspecified or Gardasil HPV vaccine 2	|__|__| |__|__|__| |__|__|__|__| day month year	|__|__| |__|__|__| |__|__|__|__| day month year or ☐ not available
Is Pandemic Flu Vaccination administered?	☐ No ☐ Yes	☐ No ☐ Yes
Name of Pandemic Flu Vaccination 1	_____	_____
Date of Pandemic Flu Vaccination 1	|__|__| |__|__|__| |__|__|__|__| day month year	|__|__| |__|__|__| |__|__|__|__| day month year or ☐ not available
Name of Pandemic Flu Vaccination 2	_____	_____
Date of Pandemic Flu Vaccination 2	|__|__| |__|__|__| |__|__|__|__| day month year	|__|__| |__|__|__| |__|__|__|__| day month year or ☐ not available
Any other vaccines?	☐ No ☐ Yes	☐ No ☐ Yes
Name of Vaccine 1	_____	_____
Date of Vaccine 1	|__|__| |__|__|__| |__|__|__|__| day month year	|__|__| |__|__|__| |__|__|__|__| day month year or ☐ not available
Name of Vaccine 2	_____	_____
Date of Vaccine 2	|__|__| |__|__|__| |__|__|__|__| day month year	|__|__| |__|__|__| |__|__|__|__| day month year or ☐ not available
Name of Vaccine 3	_____	_____
Date of Vaccine 3	|__|__| |__|__|__| |__|__|__|__| day month year	|__|__| |__|__|__| |__|__|__|__| day month year or ☐ not available
Comments		_____ 3.

CRF version template 13.2 – October 09, 2013

GlaxoSmithKline

116239 (EPI-HPV-040 VS UK)

				Patient ID
		New Onset of Autoimmune Diseases		_____

	CPRD	PALLAS
Did the subject report any NOAD/s during the follow-up period?	☐ No	☐ No
	☐ Yes	☐ Yes: please report the NOAD/s in the following section/s
Comments		_____

4.

CRF version template 13.2 – October 09, 2013

116239 (EPI-HPV-040 VS UK)

				Patient ID
		New Onset of Autoimmune Diseases NOAD - 1		|__|__|__|__|__|__|__|__|

CPRD

Origin of data
- ☐ CPRD
- ☐ HES
- ☐ CPRD and HES

Name of NOAD
_____

Category of NOAD
- ☐ Multiple Sclerosis
- ☐ Transverse myelitis
- ☐ Optic neuritis
- ☐ Guillain-Barré syndrome
- ☐ Other demyelinating diseases:
- ☐ Acute disseminated encephalomyelitis
- ☐ AI peripheral neuropathies and plexopathies.
- ☐ Auto-immune uveitis
- ☐ Rheumatoid arthritis (RA)
- ☐ Juvenile rheumatoid arthritis (JRA)
- ☐ Still's disease
- ☐ Psoriatic arthritis
- ☐ Ankylosing Spondylitis
- ☐ Idiopathic thrombocytopenic purpura (ITP)
- ☐ AI haemolytic anaemia
- ☐ Type 1 diabetes mellitus
- ☐ AI thyroiditis
- ☐ Crohn's diseases
- ☐ Ulcerative colitis
- ☐ Autoimmune hepatitis

Date of diagnosis |__|__|/|__|__|__|/|__|__|__|__|
day month year

5.

CRF version template 13.2 – October 09, 2013

GlaxoSmithKline

116239 (EPI-HPV-040 VS UK)

				Patient ID
		New Onset of Autoimmune Diseases NOAD - 1		|__|__|__|__|__|__|__|__|

PALLAS/EXPERT 1/EXPERT 2/FINAL ASCERTAINMENT*

Origin of data* ☐ CPRD ☐ HES ☐ CPRD and HES

Name of NOAD* _____

Category of NOAD*

☐ **1-Neuroinflammatory/ophthalmic autoimmune diseases:**
- ☐ 1.1 - Multiple Sclerosis
- ☐ 1.2 - Transverse myelitis
- ☐ 1.3 - Optic neuritis
- ☐ 1.4 - Guillain-Barré syndrome
- ☐ 1.5 - Other demyelinating diseases:
  - ☐ 1.51 - Acute disseminated encephalomyelitis
  - ☐ 1.52 - AI peripheral neuropathies and plexopathies.
- ☐ 1.6 - Auto-immune uveitis

2- ☐ **Other autoimmune diseases:**
- ☐ 2.1 - Systemic lupus erythematous
- ☐ 2.2 - AI disease with rheumatologic condition
  - ☐ 2.21 - Rheumatoid arthritis (RA)
  - ☐ 2.22 - Juvenile rheumatoid arthritis (JRA)
  - ☐ 2.23 - Still's disease
  - ☐ 2.24 - Psoriatic arthritis
  - ☐ 2.25 - Ankylosing Spondylitis
- ☐ 2.3 **-** AI haematological conditions:
  - ☐ 2.31 - Idiopathic thrombocytopenic purpura (ITP)
  - ☐ 2.32 - AI haemolytic anaemia
- ☐ 2.4 - AI endocrine conditions:
  - ☐ 2.41 - Type 1 diabetes mellitus
  - ☐ 2.42 - AI thyroiditis
- ☐ 2.5 - Inflammatory bowel / hepatic diseases:
  - ☐ 2.51 - Crohn's diseases
  - ☐ 2.52 - Ulcerative colitis
  - ☐ 2.53 - Autoimmune hepatitis

Date of first symptom of NOAD* |__|__|__|__|__|__|__|__| ☐ Confirmed ☐ Not Confirmed
day month year

Date of diagnosis* |__|__|__|__|__|__|__|__| ☐ Confirmed ☐ Not Confirmed
day month year

Date of first abnormal lab result related to NOAD* |__|__|__|__|__|__|__|__|
day month year

Abnormal lab test and results (with the normal ranges of the test)* _____

Final classification of NOAD* ☐ Confirmed ☐ Not Confirmed

Expert Review** ☐ Need Expert Review ☐ No Need Expert Review

Comments* _____

_____

- **Note: The data panel is applicable for each review group including Final Ascertainment and must be filled whenever required.*
- *** For PALLAS Review only*

6.

CRF version template 13.2 – October 09, 2013

GlaxoSmithKline

116239 (EPI-HPV-040 VS UK)

				Patient ID
		New Onset of Autoimmune Diseases NOAD - 1		|__|__|__|__|__|__|__|__|

MEDICATION

PALLAS

Prescription of drugs

☐ Not Available
☐ No
☐ Yes → Please complete the following table.

Sl. No.	Drug Name	Chronic Use?	**Start Date**day/month/year **End Date**day/month/year or ✓ box if not available ↓
1	_____	☐	Start: |__|__|__|__|__|__|__|__|__|__| or ☐ End: |__|__|__|__|__|__|__|__|__|__| or ☐
2	_____	☐	Start: |__|__|__|__|__|__|__|__|__|__| or ☐ End: |__|__|__|__|__|__|__|__|__|__| or ☐
3	_____	☐	Start: |__|__|__|__|__|__|__|__|__|__| or ☐ End: |__|__|__|__|__|__|__|__|__|__| or ☐
4	_____	☐	Start: |__|__|__|__|__|__|__|__|__|__| or ☐ End: |__|__|__|__|__|__|__|__|__|__| or ☐
5	_____	☐	Start: |__|__|__|__|__|__|__|__|__|__| or ☐ End: |__|__|__|__|__|__|__|__|__|__| or ☐
6	_____	☐	Start: |__|__|__|__|__|__|__|__|__|__| or ☐ End: |__|__|__|__|__|__|__|__|__|__| or ☐

Comments _____

7.

GlaxoSmithKline

116239 (EPI-HPV-040 VS UK)

				Patient ID
		New Onset of Autoimmune Diseases NOAD - 2		|__|__|__|__|__|__|__|__|

CPRD

Origin of data
- ☐ CPRD
- ☐ HES
- ☐ CPRD and HES

Name of NOAD
_____

Category of NOAD
- ☐ Multiple Sclerosis
- ☐ Transverse myelitis
- ☐ Optic neuritis
- ☐ Guillain-Barré syndrome
- ☐ Other demyelinating diseases:
- ☐ Acute disseminated encephalomyelitis
- ☐ AI peripheral neuropathies and plexopathies.
- ☐ Auto-immune uveitis
- ☐ Rheumatoid arthritis (RA)
- ☐ Juvenile rheumatoid arthritis (JRA)
- ☐ Still's disease
- ☐ Psoriatic arthritis
- ☐ Ankylosing Spondylitis
- ☐ Idiopathic thrombocytopenic purpura (ITP)
- ☐ AI haemolytic anaemia
- ☐ Type 1 diabetes mellitus
- ☐ AI thyroiditis
- ☐ Crohn's diseases
- ☐ Ulcerative colitis
- ☐ Autoimmune hepatitis

Date of diagnosis |__|__||__|__|__||__|__|__|__|
day month year

8.

GlaxoSmithKline

116239 (EPI-HPV-040 VS UK)

				Patient ID
		New Onset of Autoimmune Diseases NOAD - 2		|__|__|__|__|__|__|__|__|__|

PALLAS/EXPERT 1/EXPERT 2/FINAL ASCERTAINMENT*

Origin of data* ☐ CPRD ☐ HES ☐ CPRD and HES

Name of 2nd NOAD* _____

Category of 2nd NOAD*

☐ **1-Neuroinflammatory/ophthalmic autoimmune diseases:**
- ☐ 1.1 - Multiple Sclerosis
- ☐ 1.2 - Transverse myelitis
- ☐ 1.3 - Optic neuritis
- ☐ 1.4 - Guillain-Barré syndrome
- ☐ 1.5 - Other demyelinating diseases:
  - ☐ 1.51 - Acute disseminated encephalomyelitis
  - ☐ 1.52 - AI peripheral neuropathies and plexopathies.
- ☐ 1.6 - Auto-immune uveitis

2- ☐ **Other autoimmune diseases:**
- ☐ 2.1 - Systemic lupus erythematous
- ☐ 2.2 - AI disease with rheumatologic condition
  - ☐ 2.21 - Rheumatoid arthritis (RA)
  - ☐ 2.22 - Juvenile rheumatoid arthritis (JRA)
  - ☐ 2.23 - Still's disease
  - ☐ 2.24 - Psoriatic arthritis
  - ☐ 2.25 - Ankylosing Spondylitis
- ☐ 2.3 **-** AI haematological conditions:
  - ☐ 2.31 - Idiopathic thrombocytopenic purpura (ITP)
  - ☐ 2.32 - AI haemolytic anaemia
- ☐ 2.4 - AI endocrine conditions:
  - ☐ 2.41 - Type 1 diabetes mellitus
  - ☐ 2.42 - AI thyroiditis
- ☐ 2.5 - Inflammatory bowel / hepatic diseases:
  - ☐ 2.51 - Crohn's diseases
  - ☐ 2.52 - Ulcerative colitis
  - ☐ 2.53 - Autoimmune hepatitis

Date of first symptom of 2nd NOAD* |__|__|__|__|__|__|__|__| day month year ☐ Confirmed ☐ Not Confirmed

Date of diagnosis* |__|__|__|__|__|__|__|__| day month year ☐ Confirmed ☐ Not Confirmed

Date of first abnormal lab result related to 2nd NOAD* |__|__|__|__|__|__|__|__| day month year

Abnormal lab test and results (with the normal ranges of the test)* _____

Final classification of 2nd NOAD* ☐ Confirmed ☐ Not Confirmed

Expert Review** ☐ Need Expert Review ☐ No Need Expert Review

Comments* _____

_____

- **Note: The data panel is applicable for each review group including Final Ascertainment and must be filled whenever required.*
- *** For PALLAS Review only*

9.

CRF version template 13.2 – October 09, 2013

GlaxoSmithKline

116239 (EPI-HPV-040 VS UK)

		New Onset of Autoimmune Diseases NOAD - 2		Patient ID
				|__|__|__|__|__|__|__|__|

MEDICATION

PALLAS

Prescription of drugs

☐ Not Available
☐ No
☐ Yes → Please complete the following table.

Sl. No.	Drug Name	Chronic Use?	**Start Date** day/month/year **End Date** day/month/year or ✓ box if not available ↓
1	_____	☐	Start: |__|__|__|__|__|__|__|__|__|__| or ☐ End: |__|__|__|__|__|__|__|__|__|__| or ☐
2	_____	☐	Start: |__|__|__|__|__|__|__|__|__|__| or ☐ End: |__|__|__|__|__|__|__|__|__|__| or ☐
3	_____	☐	Start: |__|__|__|__|__|__|__|__|__|__| or ☐ End: |__|__|__|__|__|__|__|__|__|__| or ☐
4	_____	☐	Start: |__|__|__|__|__|__|__|__|__|__| or ☐ End: |__|__|__|__|__|__|__|__|__|__| or ☐
5	_____	☐	Start: |__|__|__|__|__|__|__|__|__|__| or ☐ End: |__|__|__|__|__|__|__|__|__|__| or ☐
6	_____	☐	Start: |__|__|__|__|__|__|__|__|__|__| or ☐ End: |__|__|__|__|__|__|__|__|__|__| or ☐

Comments _____

10.

CRF version template 13.2 – October 09, 2013

List of abbreviations

AID	autoimmune disease
AMA	anti-mitochondrial antibody
ANA	antinuclear antibody
ANCA	antineutrophil cytoplasmic antibodies
AS	ankylosing spondylitis
CPRD	Clinical Practice Research Datalink
CT	computerized tomography
ESR	erythrocyte sedimentation rate
GSK	GlaxoSmithKline
HES	Hospital Episodes Statistics
HLA	human leukocyte antigen
IBD	inflammatory bowel disease
ITP	idiopathic thrombocytopenic purpura
JRA	juvenile rheumatoid arthritis
LKM	liver/kidney microsomal
MRI	magnetic resonance imaging
MS	multiple sclerosis
NOAD	new onset autoimmune disease
RA	rheumatoid arthritis
RDE	Remote Data Entry
RTI	Research Triangle Institute - Health Solutions
SLA	soluble liver antigen
SLE	systemic lupus erythematosus
SMA	smooth muscle antibody
TM	transverse myelitis
TSH	thyroid stimulating hormone
UC	ulcerative colitis

Annex 6 Additional information

Not applicable.

Annex 7 Report sign-off

Please refer to the modular appendices to the main study report.

MODULAR APPENDICES

List of modular appendices available for the study report and ICH-specific appendices - Study Information equivalent numbering

Modular appendices	ICH numbering
Protocol and protocol amendments	16.1.1
List of IECs or IRBs (plus name of committee chair if required by regulatory authority)	16.1.3
Representative written information for patient and sample consent forms.	16.1.3
Signatures of principal or coordinating investigator(s) or sponsor's responsible medical officer, depending on the regulatory authority's requirement	16.1.5
Important publications referenced in the report	16.1.12

Protocol and Protocol Amendments

CONFIDENTIAL

Study Protocol

Sponsor:

GlaxoSmithKline Biologicals
Rue de l'Institut 89
1330 Rixensart, Belgium

1. PASS INFORMATION

Title	An observational cohort study to assess the risk of autoimmune diseases in adolescent and young adult women aged 9 to 25 years exposed to Cervarix® in the United Kingdom
Protocol version identifier	116239 (EPI-HPV-040 VS UK)
Date of last version of the protocol:	FDA - EMA PASS Final Version 1: 09 July 2013
EU PAS Register No:	NA (Not applicable)
Active substance	J07BM02-Papillomavirus (human types 16, 18)
Medicinal product(s):	Cervarix®, Human Papillomavirus vaccine (Types 16, 18)
Product reference:	EU/1/07/419
Procedure number:	NA
Marketing Authorisation Holder	GlaxoSmithKline Biologicals Rue de l'Institut 89 1330 Rixensart, Belgium
Joint PASS	No
Research question and objectives	To assess the risk of neuroinflammatory/ophthalmic autoimmune diseases and other pre-specified autoimmune diseases within 12 months following the administration of the first dose of Cervarix
Country(-ies) of study	United Kingdom

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Authors	Coordinating author: • [redacted], Project Manager – Science Writing GSK contributors: • [redacted], Director, Lead Epidemiologist, GSK Biologicals • [redacted], Director, Observational Data Analytics, Worldwide Epidemiology, GSK • [redacted], Project Statistician, GSK Biologicals • [redacted], Consultant Statistician, Business and Decision, c/i GSK Biologicals • [redacted], Therapeutic Area Head (Cervarix), Safety Evaluation & Risk Management (SERM), VCSP • [redacted], Therapeutic Area Head (Adjuvant), SERM, VCSP • [redacted], Safety Physician, VCSP • [redacted], Epidemiology Director, North American Clinical Development • [redacted], Senior Manager, North America Regulatory Affairs • [redacted], Specialist, Global Regulatory Affairs • [redacted], Global Regulatory Lead • [redacted], Global Study Manager, [redacted], c/i GSK Biologicals • [redacted], Study Data Manager • [redacted], Study Data Manager • [redacted], Director, North American Clinical Development / Medical Affairs, Adult Vaccines Other contributors: • [redacted], Professor of Clinical Epidemiology, London School of Hygiene and Tropical Medicine • [redacted], CPRD GOLD Research Group • [redacted], CPRD GOLD Research Group

2. MARKETING AUTHORISATION HOLDER

Marketing authorisation holder(s)	GlaxoSmithKline Biologicals Rue de l'Institut 89, 1330 Rixensart, Belgium

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

TABLE OF CONTENTS

PAGE

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

LIST OF TABLES

PAGE

LIST OF FIGURES

PAGE

LIST OF ANNEXES

PAGE

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

LIST OF ABBREVIATIONS

AI	Autoimmune
CBER	Centre for Biologics Evaluation and Research (US FDA)
CI	Confidence interval
CIN	Cervical Intraepithelial Neoplasia
CPRD GOLD	Clinical Practice Research Datalink General Practitioner OnLine database
CRO	Contract Research Organisation
DH	Department of Health (UK)
FDA	Food and Drug Administration (US)
GBS	Guillain-Barré Syndrome
GP	General Practitioner
GPP	Good Pharmacoepidemiology Practices (Guidelines)
GSK	GlaxoSmithKline
HES	Hospital Episode Statistics
HIRD	HealthCare Integrated Research Database
HPV	Human papillomavirus
ICD	International Classification of Diseases
IRR	Incidence rate ratio
ISAC	Independent Scientific Advisory Committee (for Medicines and Healthcare products Regulatory Agency database research)
ISPE	International Society for Pharmacoepidemiology
ITP	Idiopathic thrombocytopenic purpura
JRA	Juvenile rheumatoid arthritis
MHRA	Medicines and Healthcare products Regulatory Agency
MMR	Measles, mumps and rubella

CONFIDENTIAL

NOAD	New onset of autoimmune disease(s)
PASS	Post Authorization Safety Study
PMS	Post-marketing surveillance
P-Y	Person-years
RA	Rheumatoid arthritis
RR	Relative Risk
SCCS	Self-control case-series
SERM	Safety Evaluation and Risk Management
SLE	Systemic Lupus Erythematous
TSS	Targeted safety study
UK	United Kingdom
US	United States (of America)
VCSP	Vaccines Clinical Safety & Pharmacovigilance

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

3. RESPONSIBLE PARTIES

GSK Biologicals has the overall responsibility for the conduct of the study.

[redacted] is the GSK Biologicals designated Head of Global Epidemiology and Lead Epidemiologist for this study.

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

4. ABSTRACT

Title	An observational cohort study to assess the risk of autoimmune diseases in adolescent and young adult women aged 9 to 25 years exposed to Cervarix® in the United Kingdom.
Rationale and background	Cervarix is GlaxoSmithKline (GSK) Biologicals' bivalent recombinant vaccine against human papillomavirus (HPV, types 16 and 18). To address a regulatory commitment made in 2009 to the US FDA, GSK initiated an observational cohort study in the USA (e-track: 113522, EPI-HPV-015) to assess the risk of new onset of autoimmune disease(s) (NOAD) within 12 months following the administration of at least one dose of Cervarix (exposed) versus a non-Cervarix vaccinated cohort (unexposed). Because of the current low Cervarix uptake in the USA which is anticipated to stay at a low level over the next few years, it will take significantly longer than the 3 years planned to complete accrual. The present protocol is submitted as an alternative epidemiological study using the Clinical Practice Research Datalink General Practitioner OnLine database (CPRD GOLD) in the UK to fulfil the post-marketing commitment. The UK has had sufficient Cervarix vaccination coverage to, in theory, enable data acquisition.
Research question and objectives	**Primary** • To assess the risk of neuroinflammatory/ophthalmic new onset of autoimmune disease(s) (NOAD) and other pre-specified NOAD within 12 months following the administration of the first dose of Cervarix **Secondary** • To describe individually the incidence of the pre-specified NOAD considering different time periods following the administration of the first dose of Cervarix: – Incidence of Guillain Barré syndrome (including Miller Fisher syndrome and other variants), and autoimmune haemolytic anaemia within two months following the administration of the first dose of Cervarix; – Incidence of idiopathic thrombocytopenic purpura (ITP) within six months following the administration of the first dose of Cervarix; – Incidence of multiple sclerosis, transverse myelitis, optic neuritis, other demyelinating diseases,

auto-immune uveitis, systemic lupus erythematous (SLE), rheumatoid arthritis (RA), juvenile rheumatoid arthritis (JRA), Still's disease, psoriatic arthritis, ankylosing spondylitis, type 1 diabetes mellitus, auto-immune thyroiditis (including Hashimoto's disease, Graves'/Basedows' disease), and inflammatory bowel / hepatic disease (Crohn's disease, ulcerative colitis and autoimmune hepatitis) within one year following the administration of the first dose of Cervarix.

Study design

- This is an observational cohort study using the CPRD GOLD data source in the UK.

  Four cohorts will be defined based on exposure to Cervarix and sex as recorded in the CPRD GOLD data source:

  – Cervarix vaccinated (exposed) female cohort

  – Unexposed historical female cohort

  – Unexposed concurrent male cohort

  – Unexposed historical male cohort

  – Study population:

    – Female population is composed of female subjects vaccinated with Cervarix between the ages of 9 to 25 years and unexposed female subjects of the same age, identified from historical data.

    – Male population is composed of 9- to 25-year-old male subjects not vaccinated with Cervarix.

  Female subjects included in the exposed cohort will have received at least one dose of GSK's vaccine Cervarix administered according to local practice.

  Female subjects in the unexposed historical cohort will be frequency matched for age and practice region identifier to the subjects included in the vaccinated (exposed) cohort.

  Comparison of the unexposed concurrent male cohort with the unexposed historical male cohort will be used as an internal control for changes over time in CPRD GOLD in reporting NOAD. The male subjects will be frequency matched for age and practice region identifier.

  A self-control case-series (SCCS) analysis for confirmed NOAD in the exposed female cohort will also be conducted, using a risk period of one year after the first Cervarix dose, a control period of one year and a

CONFIDENTIAL

six-month buffer period between risk and control periods.

Data will be extracted from CPRD GOLD and will be validated using relevant free text for the date of onset and symptoms of the identified autoimmune diseases, and full hospital discharge statistics (HES, when available) abstracted by an independent entity (according to the CPRD GOLD process).

Population

Abstract Table 1 Study groups foreseen in the study (using CPRD GOLD, UK)

Study Groups	Number of subjects	Age*
Exposed cohort	65,000	9-25 years
Non-exposed cohorts:		
Historical female cohort	65,000	9-25 years
Concurrent male cohort	65,000	9-25 years
Historical male cohort	65,000	9-25 years

* In the European Union, Cervarix is indicated for use from the age of 9 years onwards whereas in the USA, Cervarix is indicated for use in females 9 through 25 years of age.

Variables

Primary Endpoint

- Occurrence of new onset of confirmed[1] autoimmune disease during the period of one year following administration of the first dose of Cervarix (risk period) among an exposed cohort and during an equivalent time period in the unexposed cohorts for the following two co-primary composite endpoints:

  [1] **Neuroinflammatory/ophthalmic autoimmune diseases:**

  - Multiple Sclerosis
  - Transverse myelitis
  - Optic neuritis
  - Guillain-Barré syndrome, including Miller Fisher syndrome and other variants
  - Other demyelinating diseases:
    - Acute disseminated encephalomyelitis, including site specific variants: e.g. non-infectious encephalitis, encephalomyelitis, myelitis, myeloradiculomyelitis
    - AI peripheral neuropathies and plexopathies (including chronic inflammatory demyelinating

[1] Autoimmune disease diagnosis ascertainment by an expert physician panel.

polyneuropathy, multifocal motor neuropathy and polyneuropathies associated with monoclonalgammopathy).

- Auto-immune uveitis

[2] **Other autoimmune diseases:**

- Systemic lupus erythematous
- Autoimmune (AI) disease with rheumatologic conditions:
  - Rheumatoid arthritis (RA)
  - Juvenile rheumatoid arthritis (JRA)
  - Still's disease
  - Psoriatic arthritis
  - Ankylosing Spondylitis
- AI haematological conditions:
  - Idiopathic thrombocytopenic purpura (ITP)
  - AI haemolytic anaemia
- AI endocrine conditions:
  - Type 1 diabetes mellitus
  - AI thyroiditis including Hashimoto's disease, Graves'/Basedows' disease
- Inflammatory bowel / hepatic diseases:
  - Crohn's diseases
  - Ulcerative colitis
  - Autoimmune hepatitis

Secondary Endpoint

Secondary endpoint is the occurrence of new onset of individual confirmed autoimmune disease during the following specific periods:

- Occurrence of Guillain Barré syndrome (including Miller Fisher syndrome and other variants), and autoimmune haemolytic anaemia within two months following the administration of the first dose of Cervarix;
- Occurrence of idiopathic thrombocytopenic purpura (ITP) within six months following the administration of the first dose of Cervarix;
- Occurrence of multiple sclerosis, transverse myelitis, optic neuritis, other demyelinating diseases, auto-immune uveitis, systemic lupus erythematous, rheumatoid arthritis (RA), juvenile rheumatoid arthritis (JRA), Still's disease, psoriatic arthritis, ankylosing spondylitis, type 1 diabetes mellitus, auto-immune thyroiditis (including Hashimoto's

	disease, Graves'/Basedows' disease), and inflammatory bowel / hepatic disease (Crohn's disease, ulcerative colitis and autoimmune hepatitis) within one year following the administration of the first dose of Cervarix.
Data sources	The CPRD GOLD is the world's largest computerised database of linked anonymised longitudinal medical records from primary care.
Study size	Please refer to Abstract Table 1 in the Population row above.
Data Analysis	Incidence rates for NOAD will be calculated as the number of cases divided by person-time. A Poisson regression model will estimate the exposed/unexposed risk ratio and its 95% confidence interval. The Poisson model will include the number of cases in each cohort as the dependent variable, the exposure status as a binary independent variable and the log-transformed total person-year as an offset. The same statistical model will be used to compare the two cohort. The cases of NOAD in exposed subjects will be analysed using Self-control case-series (SCCS) methods.
Milestones	Provisional milestones for the study which depend on timely approval of the study in by end of Q3 2013, are provided in the cover letter to the Regulatory Authorities.

5. AMENDMENTS AND UPDATES

None

6. MILESTONES

Provisional milestones for the study which depend on timely approval of the study by end of Q3 2013, are provided in the cover letter to the Regulatory Authorities.

Milestone	Planned date
Final protocol submitted to Regulatory Authorities	31 July 2013
Start of data collection	30 September 2013
End of data collection	31 May 2014
Planned analyses completed	30 June 2014
Projected study completion	30 September 2014
Final report of study results	31 March 2015

7. RATIONALE AND BACKGROUND

Cervarix is a GlaxoSmithKline (GSK) Biologicals' bivalent recombinant vaccine against human papillomavirus (HPV, types 16 and 18). It is currently licensed in more than 120 countries worldwide, including the European Union (EU) via the Centralised Procedure. Cervarix was granted approval by the European Medicines Agency (EMA) in September 2007 and the US Centre for Biologics Evaluation and Research (CBER) in October 2009. In the US, Cervarix is indicated for the prevention of cervical cancer, cervical intraepithelial neoplasia (CIN) grade 2 or worse and adenocarcinoma *in situ*, and cervical intraepithelial neoplasia (CIN) grade 1, caused by oncogenic human papillomavirus (HPV) types 16 and 18, in females 9 through 25 years of age. In the EU, Cervarix is indicated for use from the age of 9 years for the prevention of premalignant cervical lesions and cervical cancer causally related to certain oncogenic Human Papillomavirus (HPV) types.

Pre-licensure clinical studies provide key vaccine safety data, but their power to detect rare outcomes such as new onset of auto-immune disease(s) (NOAD) is limited due to their sample size, since incidence rates of different NOAD vary roughly from 1 to 20/100,000 per year [Cooper, 2003]. A pooled analysis of NOAD data from 68,000 subjects exposed to AS04-adjuvanted HPV-16/18, herpes simplex virus and hepatitis B vaccines in the GSK development programs did not suggest any excess risk associated with the AS04-adjuvanted vaccines compared to control vaccines [Verstraeten, 2008]. A pooled safety analysis of data from 57,580 adolescent and adult females aged 9 years and above, of whom 33,339 received at least one dose of HPV vaccine, showed the vaccine to be generally well tolerated in women of all ages [GSK confidential document. Prophylactic HPV-16/18 L1 VLP Vaccine Formulated with AS04. Investigator Brochure Edition 11 March 2012; Descamps, 2009].

Analysis of the end-of-study PATRICIA efficacy trial showed the vaccine to be generally well tolerated, which included the recorded incidence of NOAD in a broad range of women, including those of different nationalities and ethnicities [Lehtinen, 2012]. The percentage of subjects experiencing a NOAD as assessed by GSK or the investigators was low and comparable between the two groups (99 (1.1%) in the Cervarix group and 95 (1.0%) in the Hepatitis A (Havrix) group), and no imbalances between groups were observed for any event classified by the MedDRA Preferred Term.

Moreover, a recent publication by the UK Medicines and Healthcare products Regulatory Agency (MHRA) reviewed the safety profile of Cervarix use in the UK from September 2008 to July 2012 [Medicines and Healthcare products Regulatory Agency (MHRA), 2012]. No new safety concerns were identified and the number and nature of ADR reports received was as expected after administration of at least 6 million doses of the vaccine in the UK.

Siegrist et al. [Siegrist, 2007] suggested the use of population-based data which allows identification of issues of potential concern, monitoring of the impact of large-scale interventions and rapid action if any vaccine safety issues occur, that could compromise vaccine programs. The Gardasil vaccine from Merck was approved by the US FDA in 2006 and a post-licensure commitment to the regulatory authorities was established to conduct a safety surveillance study to estimate NOAD. The post-licensure study has been published in 2011 showing the Gardasil vaccine was not associated with any "autoimmune safety signal" in a large Californian database study (Kaiser Permanente Southern California & Kaiser Permanente Northern California managed care organisation databases), and "no pre-specified autoimmune condition examined demonstrated any cluster of disease onset in relation to vaccination timing, dose sequence or age" [Chao, 2011]. In this study involving 189,629 women who received one of more doses of Gardasil between August 2006 and March 2008, the incidence of potential cases of autoimmune diseases after vaccination was investigated within 3 pre-specified categories of NOAD: rheumatologic/autoimmune disorders [2a], endocrine conditions [b], and neurological/ophthalmic conditions [c]. Overall, 1014 potential onset of new cases of the pre-specified autoimmune diseases were electronically identified: 719 were eligible for case review; 31-40% were confirmed as onset of new cases. No cluster of NOAD in relation to vaccination timing, dose sequence or subject age was found. No estimated incidence rate ratios (IRR) were elevated, except for Hashimoto's disease (IRR = 1.29,

---

[2] Autoimmune conditions of interest were pre-specified and composed of three groups:

a) Rheumatologic/autoimmune disorders: immunethrombocytopenia (ITP), autoimmune haemolytic anaemia, systemic lupus erythematous (SLE), rheumatoid arthritis (RA) and juvenile rheumatoid arthritis (JRA)

b) Autoimmune endocrine conditions: type 1 diabetes, Hashimoto's disease and Graves' disease/Basedows' disease

c) Autoimmune neurological/ophthalmic conditions: multiple sclerosis (MS), acute disseminated encephalomyelitis, other demyelinating diseases of the central nervous system, vaccine-associated demyelination, Guillain-Barre´ syndrome, neuromyelitis optica, optic neuritis and uveitis

CONFIDENTIAL

95% confidence interval: 1.08-1.56) but the authors reported that "there was no consistent elevation in incidence for autoimmune thyroid conditions in the vaccinated cohort [IRR = 0.72 (0.50–1.01) for Graves' disease] and several confirmed new-onset autoimmune thyroid condition cases were likely pre-existing cases at the time of vaccination" [Chao, 2011].

GSK also committed to develop a post-licensure study of NOAD as outlined in the 2009 approval letter for Cervarix in the US:

"To conduct an observational study in a US managed care organization to evaluate the incidence of new onset autoimmune disease among at least 50,000 Cervarix recipients. The final protocol will be submitted by March 2010. Projected completion of patient accrual, subject to vaccine uptake, will be completed by March 2013. Projected study completion, subject to vaccine uptake, will be completed by September 2014. The final study report is projected to be submitted by March 2015 (6 months after study completion)".

In order to address this regulatory commitment, GSK initiated an observational cohort study to assess the risk of NOAD within 12 months following the administration of at least one dose of Cervarix (exposed) versus a non-Cervarix vaccinated cohort (unexposed). This study (e-track: 113522, EPI-HPV-015) planned to include 140,000 females, aged 9 to 25 years, enrolled in US health plans. Based on the low incidence of NOAD in this age-group, composite endpoints were defined and agreed with the FDA. Its primary objective is to evaluate whether there is an increased incidence of neuroinflammatory NOAD or other NOAD within 12 months following the administration of at least one dose of Cervarix. Data are retrieved from a large insurance administrative claims database (HealthCore Integrated Research Database (HIRD), HealthCore Inc.,WellPoint, New York, US) which includes data from 14 health plans distributed throughout the US, representing claims information from the largest commercially insured population in the US.

In the US, the commercial distribution of Cervarix began in November 2009. However, the uptake of Cervarix in the US is currently lower than initially expected (51,000 doses were distributed in 2009; 234,710 doses were distributed in 2010; 153,730 doses were distributed for 2011; 134,720 doses were distributed in 2012; and between 100,000 to 150,000 doses are projected for 2013). During the time period of 16 October 2009 to 30 April 2012, 851 females in the Cervarix® exposed and 851 females in the unexposed cohorts were accrued, who received a total of 1,516 cumulative doses of Cervarix®. This is 1.2% of the target number of 70,000 females in the Cervarix® exposed and unexposed cohorts, respectively, and 1.1% of the target of 135,000 cumulative Cervarix® doses. At this rate of accrual, it will take significantly longer than the 3 years planned to complete accrual in this study.

The present protocol is submitted as an alternative epidemiological study using the Clinical Practice Research Datalink General Practitioner OnLine database (CPRD GOLD) in the UK to fulfil the post-marketing commitment.

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

7.1. Autoimmune disease(s) incidence rates for UK and USA

The background incidence rates of auto-immune diseases in the UK and USA for male and female subjects were derived from published literature and have been tabulated by GSK Biologicals in Table 1, showing no difference in magnitude between the two countries for the age range from 9 to 25 years and for events for which data are available in both countries.

CONFIDENTIAL

Table 1 Background incidence rates of NOAD in the UK and US*

Disease	Age (years)	Incidence rate UK (per 100,000/year)		Age (years)	Incidence rate US (per 100,000/year)		References
		Males	Females		Males	Females	
Diabetes mellitus type 1	0-14	15.4-26.8	15.3-25.9	0-14	14.1-19.1	15.1-16.4	UK: Diamond Project Group, 2006**;Harron, 2011; Imkampe, 2011 US: Diamond Project Group, 2006**
	10-19	35.0	26.0	10-17	6.7-33.1	6-28.2	UK: Gonzalez, 2009; US: Kostraba, 1992; Lipton, 1995; MacDonald, 1989; Allen, 1986; NCKP$$
	15-34	20.0	10.1	18-25	10.0-15.2	7.9-19.2	UK: Imkampe, 2011; US: Fishbein, 1982; Allen, 1986; NCKP$$
	NA	NA	NA	9-26	NA	18.0	US: Chao, 2011
Inflammatory bowel disease***	10-19	4.2-10.7	3.4-11.0	10-19	2.1-8.7	3.5-9.4	UK: Steed, 2010; Henderson, 2012 US: Herrinton, 2008; Abramson, 2010
	20-29	15.5	21.9	20-29	8.3-15.7	8.1-13.4	UK: Steed, 2010; Henderson, 2012; US: Herrinton, 2008
Multiple sclerosis	15-19	0	0.9	10-17	0	0-2.3	UK: Alonso, 2007$; US: Mayr, 2003; NCKP$$
	20-24	1.7	5.9	18-25	5.3	7.5-8.6	UK: Alonso, 2007$; US: Mayr, 2003; NCKP$$
	NA	NA	NA	9-26	NA	2.5	US: Chao, 2011
Immune or idiopathic thrombocytopenic purpura (ITP)	0-18	4.7	3.7	9-26	NA	5.9	UK: Yong, 2010; US: Chao, 2011
	6-17	2.1-2.6	2.7-3.4	10-17	NA	1.5-15	UK: Schoonen, 2009; US: Simpson, 1989; NCKP$$
	18-29	0.6-1.6	3.6-4.9	18-25	NA	3.3-15	UK: Abrahamson, 2009; US: Simpson, 1989, NCKP$$
Guillain-Barré syndrome	NA	NA	NA	10-17	NA	0.8-1.8	UK: Hughes, 2006 US: Shui, 2012; Beghi, 1996; Koobatian, 1991; Riggs, 1989; NCKP$$
	NA	NA	NA	5-17	1.1-1.9	0.8-1.2	
	15-24	0.6	1.1	18-25	1.4-2.2	0.4-2	
Systemic lupus erythematous	10-19	0.1	2.3	10-17	0-0.3	1.5-3.4	UK: Nightingale, 2006; US: Hochberg, 1985; McCarty, 1995; Naleway, 2005
	20-29	0.0	4.7	18-25	1.3-1.7	5.6-19.2	
	NA	NA	NA	9-26	NA	10.3	US: Chao, 2011
Rheumatoid arthritis	NA	NA	NA	9-34	3.6	7.0-13.8	US: Myasoedova, 2010; Chao, 2011
Autoimmune thyroiditis	NA [a]	NA	NA	10-17	NA	19.5	US: NCKP$$
	NA	NA	NA	18-25	NA	37.8	
Optic neuritis	NA	NA	NA	9-26	NA	3.9	US: Chao, 2011
Uveitis	NA	NA	NA	9-26	NA	11.9	US: Chao, 2011

NA = not available; ** Study conducted in several parts of the UK and US. UK: the ranges represent data from Scotland, Leicestershire, Northern Ireland, Oxford, Plymouth and Yorkshire. US: The ranges represent data from Allegheny and Jefferson; *** Inflammatory bowel disease includes Crohn's disease, ulcerative colitis and non-Crohn's colitis (ulcerative colitis and unclassified inflammatory bowel disease combined); $ Used data from the beginning of the GPRD, likely to be incomplete;
$$ Northern California Kaiser Permanente. Only female data available from the previous GSK review. See EPI-HPV-015 (e-track: 113522)

[a] For all ages – Incidence in male = 22/100,000/year and in female = 99/100,000/year in Scotland, UK [Leese, 2008]

8. RESEARCH QUESTIONS AND OBJECTIVES

8.1. Primary objective

- To assess the risk of neuroinflammatory/ophthalmic new onset of autoimmune disease(s) (NOAD) and other pre-specified NOAD within 12 months following the administration of the first dose of Cervarix:

  [1] **Neuroinflammatory/ophthalmic autoimmune diseases:**

  - Multiple Sclerosis
  - Transverse myelitis
  - Optic neuritis
  - Guillain-Barré syndrome, including Miller Fisher syndrome and other variants
  - Other demyelinating diseases:
    - Acute disseminated encephalomyelitis, including site specific variants: e.g. non-infectious encephalitis, encephalomyelitis, myelitis, myeloradiculomyelitis
    - AI peripheral neuropathies and plexopathies (including chronic inflammatory demyelinating polyneuropathy, multifocal motor neuropathy and polyneuropathies associated with monoclonalgammopathy).
  - Auto-immune uveitis

  [2] **Other autoimmune diseases:**

  - Systemic lupus erythematous
  - Autoimmune (AI) disease with rheumatologic conditions:
    - Rheumatoid arthritis (RA)
    - Juvenile rheumatoid arthritis (JRA)
    - Still's disease
    - Psoriatic arthritis
    - Ankylosing Spondylitis
  - AI haematological conditions:
    - Idiopathic thrombocytopenic purpura (ITP)
    - AI haemolytic anaemia
  - AI endocrine conditions:
    - Type 1 diabetes mellitus
    - AI thyroiditis including Hashimoto's disease, Graves' /Basedows' disease
  - Inflammatory bowel / hepatic diseases:
    - Crohn's diseases
    - Ulcerative colitis
    - Autoimmune hepatitis

8.2. Secondary objectives

- To describe individually the incidence of the pre-specified NOAD considering different time periods following the administration of the first dose of Cervarix:
  - Incidence of Guillain Barré syndrome (including Miller Fisher syndrome and other variants), and autoimmune haemolytic anaemia within two months following the administration of the first dose of Cervarix;
  - Incidence of idiopathic thrombocytopenic purpura (ITP) within six months following the administration of the first dose of Cervarix;
  - Incidence of multiple sclerosis, transverse myelitis, optic neuritis, other demyelinating diseases [3], auto-immune uveitis, systemic lupus erythematous (SLE), rheumatoid arthritis (RA), juvenile rheumatoid arthritis (JRA), Still's disease, psoriatic arthritis, ankylosing spondylitis, type 1 diabetes mellitus, auto-immune thyroiditis (including Hashimoto's disease, Graves'/Basedows' disease), and inflammatory bowel / hepatic disease (Crohn's disease, ulcerative colitis and autoimmune hepatitis) within one year following the administration of the first dose of Cervarix.

Refer to Section 9.3 for the definition of the primary and secondary endpoints and the pre-specified list of NOAD.

8.3. Exploratory objective

- To evaluate if temporal clustering of the individual NOADs comprising the primary endpoint and the secondary endpoints (i.e., those NOADs on the pre-defined list) occurs following the administration of at least one dose of Cervarix® within the 12-month follow-up period.

9. RESEARCH METHODS

9.1. Study Design

9.1.1. Overview

- This is an observational cohort study using the CPRD GOLD data source in the UK.
- Four cohorts will be defined based on exposure to Cervarix and sex as recorded in the CPRD GOLD data source:

[3] Other demyelinating diseases:
- Acute disseminated encephalomyelitis, including site specific variants: e.g. non-infectious encephalitis, encephalomyelitis, myelitis, myeloradiculomyelitis
- AI peripheral neuropathies and plexopathies (including chronic inflammatory demyelinating polyneuropathy, multifocal motor neuropathy and polyneuropathies associated with monoclonalgammopathy).

CONFIDENTIAL

1. Cervarix vaccinated (exposed) female cohort
2. Unexposed historical female cohort
3. Unexposed concurrent male cohort
4. Unexposed historical male cohort

- Study population:
  - Female population is composed of female subjects vaccinated with Cervarix between the ages of 9 to 25 years and unexposed female subjects identified from historical data.
  - Male population is composed of 9- to 25-year-old male subjects not vaccinated with Cervarix.

  Female subjects included in the exposed cohort will have received at least one dose of GSK's vaccine Cervarix administered according to local practice.

  Female subjects in the unexposed historical cohort will be frequency matched for age and practice region identifier to the subjects included in the vaccinated (exposed) cohort.

  Study design:

  Comparison of the unexposed concurrent male cohort with the unexposed historical male cohort will be used as an internal control for changes over time in CPRD GOLD in reporting NOAD. The male subjects will be frequency matched for age and practice region identifier as described in Section 9.2.

  A self-control case-series (SCCS) analysis for confirmed NOAD in the exposed female cohort will also be conducted, using a risk period of one year after the first Cervarix dose, a control period of one year and a six-month buffer period between risk and control periods.

9.1.2. Rationale for study design

NOAD represent a heterogeneous group of diseases with different clinical conditions and disease progression. Some NOAD present with a chronic disease pattern of relapse over time e.g. multiple sclerosis or systemic lupus erythematous, or an acute disease pattern (e.g. Guillain Barré Syndrome).

In addition to the comparison of the exposed vs. non-exposed cohort (cohort design), the confirmed NOAD in the exposed cohort will also be analysed using a self-control case-series (SCCS) analysis.

In the cohort design, specified NOAD will be collected over a period of one year following the administration of at least one dose of Cervarix in an exposed cohort and over a comparable period in the unexposed cohorts as shown in Figure 1. Four cohorts will be constituted as shown in Figure 1. The unexposed male cohorts will be enrolled in order to assess a possible change over time in the incidence rate of NOAD in CPRD GOLD independent of Cervarix introduction. The cohorts will be frequency

CONFIDENTIAL

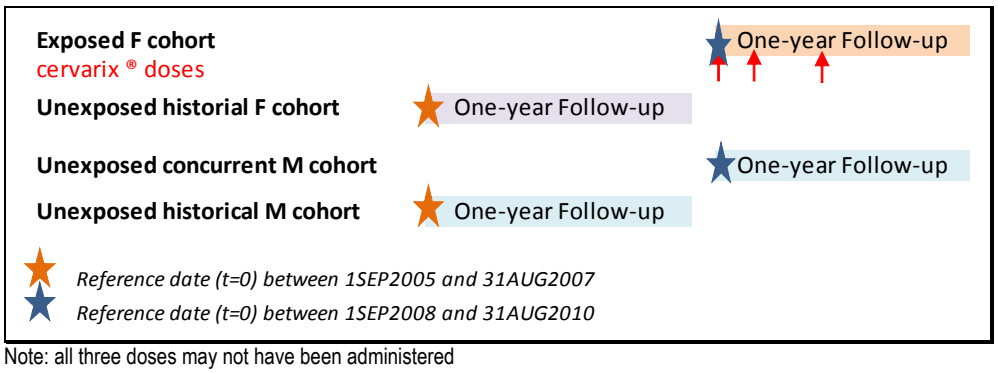
matched for the age (age class of one year) and practice region identifier at reference date (age at first dose of Cervarix).

An unexposed concurrent female cohort has been not chosen as the unvaccinated group for the following reasons:

- There is a under-reporting of HPV vaccination in CPRD GOLD as evidenced by the difference in vaccine coverage in CPRD GOLD population and the UK population (see Table 2), an exposed concurrent cohort would include both true and false unexposed subjects;
- Moreover, because of the large vaccine coverage in the UK, non-vaccinated women could have different health care behaviour compared to the vaccinated women, resulting in a difference in the probability to detect auto-immune disease.

Therefore, the reference date (time = 0) for the vaccinated (exposed) cohort will be the date of the first dose of Cervarix recorded in CPRD GOLD. The reference date for the unexposed (unvaccinated) cohorts will be a date randomly selected among the reference dates of the exposed subjects and minus 3 years for the historical cohorts.

Figure 1 Cohort design



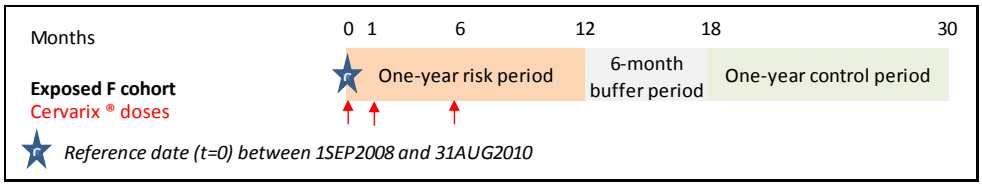
Note: all three doses may not have been administered

For the SCCS analysis, the exposed female cohort will be followed during a total period of 30 months from the first dose. The risk period will be defined as one year after the first dose, including the six months after the last dose when the full three-dose vaccination course is administered. A control period of the same duration will be defined, excluding the periods after subsequent doses. To control for a possible late effect of the vaccine, the control and the risks period will be separated by a six-month buffer period (Figure 2). For analysis of individual diseases, specific risk period could be defined.

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Figure 2 Self-control case-series analysis



Note: all three doses may not have been administered

The SCCS uses each subject as its own control, preventing virtually all potential confounding factors which do not vary with time (e.g. socio-economic status, gender, location) [Whitaker, 2006]. Additionally, fewer cases are usually required, as compared to a case-control design. The control period is chosen subsequent to the at-risk period. Age bias is not anticipated because the follow-up period is short compared to the age effect on incidence of autoimmune diseases (see Table 1).

9.1.3. HPV vaccine coverage in UK and in CPRD GOLD

The submitted study is in the UK, which is a country with a high Cervarix vaccine coverage. The recommended schedule in the UK for HPV vaccination of 12-13 year old girls involves three doses of Cervarix given over at least a 6-month period [DH, 2008; HPA, 2010].

In the UK public HPV immunization program (12-13 year olds), HPV vaccination coverage in the UK for 2010/11 was 89.0%, 87.6% and 83.8% for the first, second and third dose respectively [Health Protection Agency, 2012].

Vaccination records in the CPRD GOLD cover only part of the Cervarix-vaccinated population in the UK. For instance, 70.7% of the 01-SEP-1996 – 31-AUG-1997 birth cohort were reported as HPV vaccinated in CPRD GOLD versus 85.9% reported by the UK Health Protection Agency. Corresponding numbers for the 01-SEP-1994 – 31-AUG-1995 birth cohort are 70.6% versus. 81.9% (Table 2).

Table 2 HPV vaccination coverage in CPRD GOLD versus HPA-DH UK data

	CPRD†			UK‡
Birth cohort	N vaccinated	N total	%	%
01 SEP 1996 and 31 AUG 1997	16028	22685	70.7%	85.9%
01 SEP 1994 and 31 AUG 1995	16343	23143	70.6%	81.9%

† Data extracted from CPRD GOLD version ffgprd_smart_2012Q3
‡ From Health Protection Agency Department of Health [Health Protection Agency, 2012]

A first exploration of the CPRD GOLD database identified 148,731 subjects vaccinated with Cervarix between 2007 and 2010. The number of vaccinated women in the relevant age-range included in the CPRD GOLD is provided in Table 3. The number of HPV-vaccinated females in the CPRD GOLD appears to be sufficient for studying adverse events which have a low incidence after vaccination with Cervarix.

CONFIDENTIAL

Table 3 Number of HPV-vaccinated females in the relevant age range included in the CPRD GOLD

Age range [years]	Number of subjects
[9-10]	16
[11-15]	82326
[16-20]	65080
[21-25]	866

Created by GSK Biologicals', 2012
Date extracted from CPRD GOLD version ffgprd_smart_2012Q3

9.2. Setting

Setting and study population

9.2.1. The UK HPV National Immunization Programme

The submitted study is in the UK, which is a country with a high Cervarix vaccine coverage. The recommended schedule in the UK for routine HPV vaccination of all girls at 12 to 13 years of age is [DH, 2008; HPA, 2010]:

- Vaccination at this age starts with a first dose of 0.5ml of Cervarix HPV vaccine
- The second dose of 0.5ml follows one to two months after the first dose
- A third dose of 0.5ml follows at least six months after the first dose

The UK has had sufficient Cervarix vaccination coverage to, in theory, enable data acquisition. A public immunisation programme targeting girls between 12-13 years of age including a catch-up programme for young women up to 18 years was undertaken during the academic year 2008/09. A phased catch-up programme for females born 1 September 1991 to 31 August 1995 during the 2008/09 academic year was completed by the end of the 2009/10 academic year. The programme was delivered largely through secondary schools [Crawford, 2009; Sheridan, 2009; Sheridan, 2010]. In the UK public HPV immunization program (12-13 year olds), HPV vaccination coverage in the UK for 2010/11 was 89.0%, 87.6% and 83.8% for the first, second and third dose respectively [Health Protection Agency, 2012]. The recommended age range for the UK programme matches the age range required by the FDA (9-25 years of age) for the post-licensure safety study.

The study population will be composed of female and male subjects, 9 to 25 years of age, registered in the CPRD GOLD.

The exposed cohort will be composed of female subjects vaccinated with at least one dose of Cervarix, with or without other recommended vaccines.

The unexposed historical female cohort will consist of frequency age-matched and practice region-matched female subjects from the period before the introduction of Cervarix. The unexposed concurrent male cohort will consist of frequency age-matched and practice region-matched by one-year classes (15, 16, 17, etc.) male subjects from the

period after the introduction of Cervarix. The unexposed historical male cohort will consist of frequency age-matched and practice region-matched male subjects from the period before the introduction of Cervarix.

9.2.2. Cohort identification and creation

The exposed eligible cohort have been identified based on the stepwise approach defined in Annex 5. Among eligible exposed subjects, 65,000 subjects will be randomly selected using the RANUNI function of SAS. The RANUNI function returns a number that is generated from the uniform distribution on the interval (0, 1). The corresponding subject number will be computed as *random_subj_number=ranuni(seed)*. The subjects will be ordered according the *random_subj_number* and the first 65,000 subjects will be included in the exposed cohort.

The unexposed eligible cohorts have been identified based on the stepwise approach define in Annex 5.

All the unexposed subjects who matched exposed subject for age (birth cohort) and region (frequency matching) will be identified. This represents 234 combinations of birth cohort-region (18 birth cohorts and 13 regions in CPRD GOLD). In each combination, the subjects will be randomly selected based on the distribution in the exposed cohort.

A reference date is randomly attributed to all potential unexposed subjects. For the concurrent male cohort, the reference date is a random date between 01-SEP-2008 and 31-AUG-2010. For the historical cohorts, the reference date is a random date between 01-SEP-2005 and 31-AUG-2007 (reference dates for exposed cohort – 3 years).

The random reference dates in each 'birth cohort-region" combination will be attributed randomly using the RANUNI function of SAS.

The age at reference date will be calculated for the unexposed cohort. After applying the exclusion criteria, the eligible subjects (65,000 subjects in each cohort) will be randomly selected in each of the combinations 'birth cohort-region' taking into account the distribution in the exposed cohort. The random selection will also use the RANUNI function of SAS.

9.2.3. Number of subjects

For the cohort design, the target sample size is 65,000 subjects for each cohort. Refer to Section 9.5 for a detailed description of the estimation of the sample size.

9.2.4. Inclusion criteria

Note: Other vaccines are allowed in this study regardless of the time of administration and the time interval between subsequent doses.

CONFIDENTIAL

9.2.4.1. Inclusion criteria for the exposed female cohort

Exposed females must satisfy ALL the following criteria at study entry:

- Female aged from 9 to 25 years at the reference date (01 September 2008 through 31 August 2010)
- Recorded in the CPRD GOLD for at least 12 months before the reference date
- The first dose of Cervarix received between 01 September 2008 through 31 August 2010, Full date (day/month/year) of Cervarix vaccination(s) available
- Subject defined as acceptable in CPRD GOLD

9.2.4.2. Inclusion criteria for the unexposed historical female cohort

Unexposed females must satisfy ALL the following criteria at study entry:

- Female aged 9 to 25 years at the reference date (01 September 2005 through 31 August 2007)
- Recorded in the CPRD GOLD for at least 12 months before the reference date
- Subject defined as acceptable in CPRD GOLD

9.2.4.3. Inclusion criteria for the unexposed concurrent male cohort

Unexposed concurrent males must satisfy ALL the following criteria at study entry:

- Male aged 9 to 25 years at the reference date (01 September 2008 through 31 August 2010)
- Recorded in the CPRD GOLD for at least 12 months before the reference date
- Subject defined as acceptable in CPRD GOLD

9.2.4.4. Inclusion criteria for the unexposed historical male cohort

Unexposed historical males must satisfy ALL the following criteria at study entry:

- Male aged 9 to 25 years at the reference date (01 September 2005 through 31 August 2007)
- Recorded in the CPRD GOLD for at least 12 months before the reference date
- Subject defined as acceptable in CPRD GOLD

CONFIDENTIAL

9.2.5. Exclusion criteria

9.2.5.1. Exclusion criteria for all cohorts

- Subjects with a diagnostic code of any auto-immune disease during the year prior to the reference date.
- Subjects who received at least one dose of unspecified HPV vaccine or Gardasil at any time before the reference date.
- Subjects who have been included in the other cohort.

9.2.5.2. Exclusion criteria for the non-exposed cohorts

- Subjects who received any dose of Cervarix at any time before the reference date.

9.3. Variables

9.3.1. Primary endpoint

- Occurrence of new onset of confirmed[4] autoimmune disease during the period of one year following administration of the first dose of Cervarix (risk period) among an exposed cohort and during an equivalent time period in the unexposed cohorts for the following two co-primary composite endpoints:

  [1] **Neuroinflammatory/ophthalmic autoimmune diseases:**

  – Multiple Sclerosis
  – Transverse myelitis
  – Optic neuritis
  – Guillain-Barré syndrome, including Miller Fisher syndrome and other variants
  – Other demyelinating diseases:
    – Acute disseminated encephalomyelitis, including site specific variants: e.g. non-infectious encephalitis, encephalomyelitis, myelitis, myeloradiculomyelitis
    – AI peripheral neuropathies and plexopathies (including chronic inflammatory demyelinating polyneuropathy, multifocal motor neuropathy and polyneuropathies associated with monoclonalgammopathy).
  – Auto-immune uveitis

  [2] **Other autoimmune diseases:**

  – Systemic lupus erythematous
  – Autoimmune (AI) disease with rheumatologic conditions:
    – Rheumatoid arthritis (RA)

[4] Auto-immune disease diagnosis ascertainment by an expert physician panel (Section 9.4.2).

CONFIDENTIAL

  - Juvenile rheumatoid arthritis (JRA)
  - Still's disease
  - Psoriatic arthritis
  - Ankylosing Spondylitis
- AI haematological conditions:
  - Idiopathic thrombocytopenic purpura (ITP)
  - AI haemolytic anaemia
- AI endocrine conditions:
  - Type 1 diabetes mellitus
  - AI thyroiditis including Hashimoto's disease, Graves' /Basedows' disease
- Inflammatory bowel / hepatic diseases:
  - Crohn's diseases
  - Ulcerative colitis
  - Autoimmune hepatitis

9.3.2. Secondary endpoints

Secondary endpoint is the occurrence of new onset of individual confirmed autoimmune disease during the following specific periods:

- Occurrence of Guillain Barré syndrome (including Miller Fisher syndrome and other variants), and autoimmune haemolytic anaemia within two months following the administration of the first dose of Cervarix;
- Occurrence of idiopathic thrombocytopenic purpura (ITP) within six months following the administration of the first dose of Cervarix;
- Occurrence of multiple sclerosis, transverse myelitis, optic neuritis, other demyelinating diseases (see the two sub-bullets for these diseases in Section 9.3.1), auto-immune uveitis, systemic lupus erythematous (SLE), rheumatoid arthritis (RA), juvenile rheumatoid arthritis (JRA), Still's disease, psoriatic arthritis, ankylosing spondylitis, type 1 diabetes mellitus, auto-immune thyroiditis (including Hashimoto's disease, Graves'/Basedows' disease), and inflammatory bowel / hepatic disease (Crohn's disease, ulcerative colitis and autoimmune hepatitis) within one year following the administration of the first dose of Cervarix.

9.3.3. Data to be collected

9.3.3.1. Subjects characteristics

The following data will be extracted for the analysis population:

- Demographic characteristics: birth month and birth year, sex, region, practice region identifier, date of death (if applicable) and acceptable patient flag

CONFIDENTIAL

- CPRD GOLD information: CPRD GOLD start date, first registration date, current registration date, registration gaps, registration status, transfer-out date, transfer-out reason
- HES information: Linkage to HES data
- Vaccines:

  Administration of any other vaccine from one year before the reference date and until the end of follow-up will be collected: date of vaccination, medcodes and immunization type will be extracted from the immunisation file. Cross-tabulation of medcodes and vaccine names/class is detailed in Annex 5.
- Health care resource utilization: number of primary care resource utilization during the year before the reference date.

9.3.3.2. Clinical outcomes

Occurrence of auto-immune diseases defined as study endpoints will be identified using defined algorithms (see Annex 5).

For each case, the following data will be extracted:

- Medcode(s)
- Date of event

The associated "free text" (event text ID) if any will be identified. For all the identified cases of auto-immune diseases, the associated free text will be reviewed by a GSK-identified reviewer for confirmation and determination of the date of first symptoms of NOAD.

9.3.3.3. Other derived variables

The following variables will be derived from the CPRD GOLD data:

- Subject's date of birth will be defined as the 15th of the birth month and birth year. If the birth month is missing, the birth date will be defined as the 30th June of the birth year
- Incomplete dates (except for vaccination date which is the reference date) will be substituted as follows for calculation of age and/or time to event; if the day is missing the date will be defined as the 15th of the month, if both the day and the month are missing, the date will be defined as 30th June of the year
- Age at a specific event will be computed as the difference between the date of the event and the date of birth

9.4. Data Sources

9.4.1. The UK Clinical Practice Research Datalink General Practitioner OnLine database (CPRD GOLD)

The CPRD GOLD is the world's largest computerised database of linked anonymised longitudinal medical records from primary care. The data are drawn from the computer systems used by general practitioners (GPs) to maintain the clinical records within their practices. As of March 2011, CPRD GOLD contains records from over 12 million patients contributing 64 million person-years of prospectively recorded high-quality primary healthcare data [Williams, 2012].

The CPRD GOLD is operated on a non-profit basis by the UK Medicines and Healthcare products Regulatory Agency (MHRA), containing coded longitudinal medical records from general practices and more recently from hospital-based care (e.g., Hospital Episode Statistics, HES). The current linkage between CPRD GOLD primary care data and HES data is around 50% as of Q1 2013. The CPRD GOLD database is licensed in-house by GSK. Data quality is monitored continuously by the MHRA and practices that fail to maintain the required standards are removed from the database.

The latest update provided by the CPRD GOLD team in Q1 of 2013 (first release of 2013) contains data for 10,960,947 research standard patients, drawn from 660 practices throughout the UK. A total of 4,727,669 patients from 548 practices are currently active in the database. The CPRD GOLD population closely matches the age and gender distribution of the UK population as a whole. Mean follow-up is 6.9 years (median 5.0 years). Recorded data include demographic information, prescription details, clinical events, preventive care provided, specialist referrals, hospital admissions and their major outcomes. Data are retrieved by means of the READ classification system; READ codes are a coded thesaurus of clinical terms, which are the basic means by which clinicians record patient findings and procedures in health and social care IT systems across primary and secondary care (e.g. GP surgeries and pathology reporting of results). The Medcodes are the abbreviated terms which mean CPRD GOLD medical codes. Medcodes consisting of READ codes are used to enter medical diagnosis in the CPRD GOLD database.

9.4.2. Data source for case ascertainment

First, the CPRD GOLD is based on data from GPs, while most auto-immune diseases would probably be diagnosed in specialist settings. Consequently, the number of auto-immune diseases, the quality of the information, and the diagnostic certainty might be lower compared to other databases that include hospital data only. In particular the specific information related to the onset of clinical symptoms, and radiological and biological data associated with the etiologic diagnosis of auto-immune diseases may not all be available in the CPRD GOLD database and associated resources. Besides, not all general practices participating to CPRD GOLD have consented to the linkage between CPRD GOLD primary care data and HES data (current linkage around 50% as of Q3 2012). Specific algorithms for each outcome of interest have been developed (Annex 5)

CONFIDENTIAL

and available "free text" that is related to auto-immune disease diagnosis will be requested from CPRD GOLD, when needed.

Moreover, in a recent study by [Chao & Jacobsen, 2012], the authors recommended that expert case review of medical records is used in autoimmune safety studies, and case identification can be expanded by use of laboratory test results and other relevant measures in addition to specific ICD-10 diagnosis codes.

The ascertainment of the etiologic diagnosis and date of disease onset for all identified auto-immune diseases will be performed by a contract research organisation (CRO) to ensure the correct classification of each case. The CRO will review all subject data retrieved from CPRD GOLD including medcodes (including clinical, laboratory, and treatment files), the relevant "free text" and HES (including specific ICD-10 diagnostic codes), when available. They will assess whether the aetiology of the auto-immune disease is confirmed or not and whether the date of disease onset falls within the observation period of the study, which is one year after the reference date.

In the event that the aetiology or the date of onset could not be confirmed, a second review step will be conducted with an expert physician panel to reach an agreement about the case ascertainment, and, if they cannot reach an agreement, then the different clinical opinion of the experts will be listed. The experts will be blinded with regards to HPV vaccine exposure.

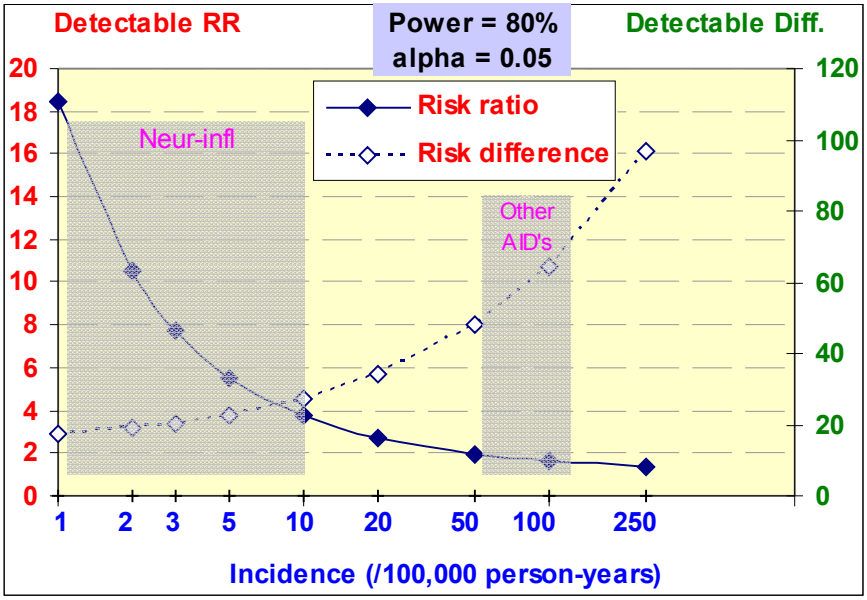
9.5. Study size

9.5.1. Sample size for cohort design

The target sample size is 50,000 subjects in each cohort. The relative risk (RR) that would be detected with 80% power and alpha = 0.05 is given in Figure 3 versus the incidence rate in the (Cervarix) unexposed cohort. The detectable difference in incidence rate (= additional cases per 100,000 person-years) is also depicted.

CONFIDENTIAL

Figure 3 **Detectable relative risk and difference versus the incidence rate in the (Cervarix) unexposed cohort**



(Method: Comparison of two independent proportions using a likelihood ratio test, PASS 2005)

Cohorts of 50,000 subjects each should allow detection, with 80% power, of a RR between 3.7 and 18.7 for the neuro-inflammatory NOAD (incidence rate between 10 and 1/100,000 person-years) and between 1.6 and 2.0 for other NOAD (incidence rate between 100 and 50/100,000 person-years).

Because of risk of lost to follow-up, the sample size is increased by approximately 30% in each cohort to approx. 65,000 subjects.

9.5.2. Sample size for self-control case-series

The power of the SCCS analysis depends on the number of cases and the ratio between the duration of the risk and the control periods. With a risk and a control period of 12 months each, the number of cases needed versus the detectable risk ratio (incidence rate ratio) is summarized in Table 4. Figure 4 shows the total number of vaccinated subjects needed to follow-up for 30 months after vaccination versus the number of cases and the background incidence.

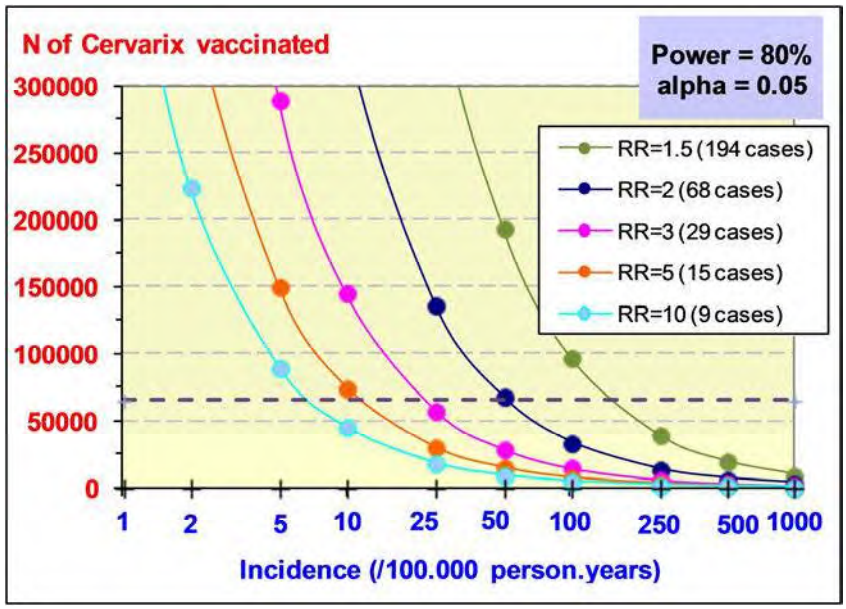
CONFIDENTIAL

Table 4 Sample size for a SCCS analysis - Number of cases in vaccinated subjects versus the incidence rate ratio [a]

Incidence rate ratio	Total number of cases
1.5	194
2	68
3	29
5	15

Method: sample for case-series analysis based on the signed root likelihood ratio [Musonda, 2006]
[a] 80% power using a two-sided test and alpha = 0.05

Figure 4 Population size for a SCCS analysis versus the incidence rate ratio and the background incidence in the general population



Dotted line: target sample size of the exposed cohort

9.6. Data Management

9.6.1. Remote Data Entry instructions

Remote Data Entry (RDE), using a validated computer application will be used by the GSK identified reviewer to enter the information obtained from the free text review and final case ascertainment classification.

In all cases, subject initials will not be collected nor transmitted to GSK. Subject data necessary for analysis and reporting will be entered/transmitted into a validated database or data system. Clinical data management will be performed in accordance with applicable GSK standards and data cleaning procedures.

No monitoring will be done. The GSK identified reviewer remains accountable for the data entry.

9.6.2. Final study database

The final study database will consist of data extracted from CPRD GOLD and additional data from free text review. The study database will be locked and stored by GSK Biologicals' data management according to GSK Biologicals Standard Procedures.

9.7. Data Analysis

Examples of statistical tables and figure templates are given in Annex 6.

9.7.1. Hypotheses

9.7.1.1. Hypotheses for the cohort analysis

Null hypothesis (H0): the incidence of neuroinflammatory/ophthalmic autoimmune diseases (other autoimmune diseases) in the exposed female cohort is equal to the incidence in the historical non-exposed female cohort (historical cohort).

Alternative hypothesis (H1): the incidence of neuroinflammatory/ophthalmic autoimmune diseases (other autoimmune diseases) in the exposed female cohort is different from the incidence in the historical unexposed female cohort (historical cohort).

These hypotheses will be tested separately for each of the two co-primary endpoints. No alpha adjustment will be done.

The same hypotheses will be tested between the two male cohorts.

9.7.1.2. Hypotheses for the self-control case-series analysis

Null hypothesis (H0): the incidence rate of neuroinflammatory/ophthalmic autoimmune diseases (other autoimmune diseases) in the exposed female cohort is the same during the risk period and the control period.

Alternative hypothesis (H1): the incidence rate of neuroinflammatory/ophthalmic autoimmune diseases (other autoimmune diseases) in the exposed female cohort is different during the risk period and during the control period.

These hypotheses will be tested separately for each of the two co-primary endpoints. No alpha adjustment will be done. These hypotheses will be tested for individual diseases using the specific risk periods provided that at least 10 cases be recorded (with 10 cases for SCCS analysis, we will have 80% power to detect incidence rate of 6 to 8 depending on the risk period versus control period ratio).

CONFIDENTIAL

9.7.2. Analysis Population

9.7.2.1. Population for the cohort design

The study population for the cohort design will comprise all exposed and unexposed subjects that satisfy the inclusion criteria.

9.7.2.2. Population for the SCCS analyses

Only the cases of NOAD recorded in the exposed cohort during either the risk or the control periods will be included in the SCCS analysis.

9.7.3. Subject disposition

Subject disposition will be summarized by cohort and overall by computing:

- Number of screened subjects.
- Number (%) of non-eligible subjects for each of the following reasons of non-eligibility:
  - Diagnostic code of NOAD during the year prior to the reference date (for all cohorts);
  - Subject not actively registered with the practice during the study period (for all cohorts);
  - Subject not flagged as acceptable in CPRD GOLD (for all cohorts);
  - Subject not recorded for at least 12 months within CPRD GOLD at reference date;
  - At least one dose of unspecified HPV vaccine or Gardasil at any time before the reference date(for all cohorts);
  - At least one dose of Cervarix vaccine at any time before the reference date (for unexposed cohorts);
  - The first dose of Cervarix received before 01 September 2008 or after 31 August 2010 (for the exposed female cohort).
- Number of eligible subjects in each cohort.
- After frequency matching for age and for practice-region, number of included subjects in each cohort.

A detailed, comprehensive list of reasons for elimination from exposed and unexposed cohort analyses will be established at the time of data cleaning.

9.7.4. Demographic and baseline characteristics

Demographic and baseline characteristics of all included subjects (age at reference date, region (GP practice), other vaccination during the previous year) will be summarized per cohort and overall, using descriptive statistics.

Frequency tables will be generated for categorical variables.

Mean, standard error, median and range will be provided for continuous variables.

The two female cohorts and the two male cohorts will be compared for their demographic and baseline characteristics using Fisher's exact test or Student t-test.

9.7.5. Analysis of co-primary endpoints

9.7.5.1. Cohort analysis

The primary analysis will compare the incidence rates of the primary outcomes of interest between the Cervarix exposed female cohort and the historical unexposed female cohort. Results will be presented as the incidence rate ratio and the incidence difference. Exposed person-time will be defined as the period between the reference date and the earliest of the following events:

- End of study period (defined as 12 months after the reference date);
- Date of de-enrolment from CPRD GOLD;
- Date of unspecified HPV vaccine or Gardasil or Cervarix for unexposed cohort;
- Date of first diagnosis of the outcome of interest.

Incidence rates for auto-immune diseases will be calculated by dividing the number of cases by person-time. A Poisson regression model will estimate the risk ratio and its 95% confidence interval. The Poisson model will include the number of cases in each cohort as the dependent variable, the exposure status as a binary independent variable, the age-group ([9-18],[18-25]) as a covariate, and the log-transformed total person-year as an offset.

Sensitivity analyses

Sensitivity analyses will be performed:

- Incidence rates for NOAD in the exposed cohort will be calculated after each dose by dividing the number of cases by the total person-time. Exposed person-time will be defined as the period between the date of the dose administration and the earliest of the following events:
  - End of risk period (defined as 6 months after each dose);
  - Date of the next Cervarix dose;
  - Date of de-enrolment from CPRD GOLD;

CONFIDENTIAL

  - Date of unspecified HPV vaccine or Gardasil;
  - Date of first diagnosis of the outcome of interest.

- Separate analysis will be performed for subjects younger/older than 18 years.
- Analysis including possible confounding factors (if available): other vaccination, age, region, healthcare resource utilization, categorization will be determined based on the available data.
- In case of more than 5% of NOAD with unknown/incomplete date, a sensitivity analysis will be done including these events as occurring during the risk period. This imputation will be done for the four cohorts.
- The same comparison will be done between the two male cohorts. In case of significant difference between these two male cohorts, the primary analysis will be adjusted for time effect other than Cervarix (see Section 9.7.5.1 for the detailed statistical models).

For the SCCS analysis, the incidence rates in the exposed cohort during the risk period will be compared with the incidence rates in the exposed cohort during the control period using a conditional Poisson regression model [Whitaker, 2006].

Analysis of individual disease could also be performed depending on the number of cases (at least 10 cases in both exposed and non-exposed cohorts for each defined risk period).

9.7.6. Secondary endpoints

The new cases of individual autoimmune disease during the specific period (see Section 9.3.2) will be analysed by descriptive statistics per cohort. Incidence rate during the specific period will be computed per cohort for each individual disease as the total number of new cases divided by the total person-year as for the primary endpoint.

In case of more than 10 cases in the exposed female and unexposed female cohorts, a Poisson regression model will estimate the risk ratio and its 95% confidence interval. The Poisson model will include the number of cases in each cohort as the dependent variable, the exposure status as a binary independent variable and the log-transformed total person-year as an offset. Same analysis will be performed for the two male cohorts.

A SCCS analysis of individual disease will be carried out if there are at least 10 cases of disease in the exposed female cohort and during the total of the risk and the control period. The risk period of each disease is the disease specific period : 2 months, 6 months, or 1 year depending on the disease (see Section 9.3.2), the control period is, 22 months, 18 months, or 12 months, respectively. The risk and the control periods are separated by a 6-month buffer period.

These analyses will not be considered as confirmatory analysis.

CONFIDENTIAL

9.7.7. Exploratory analysis

We will employ the SaTScan software package to investigate temporal clustering of adverse events. Using this package, we will analyze data to test whether the events are randomly temporal distributed. This analysis will be done for each of the four cohorts.

9.7.8. Statistical calculations

All the statistical calculations will be done in SAS 9.2 or higher.

All the statistical tests will be two-sided at alpha level of 0.05.

9.7.8.1. Handling of missing data

Missing data will not be substituted.

9.7.8.2. Descriptive statistics

Age at reference date will be summarized by descriptive statistics per cohort and overall: n of subjects, mean, SD, median, minimum and maximum and compared among the four cohorts using a one-way ANOVA. In case of an overall significant difference, pair-wise comparison using t-test will be carried out.

Exposure to other vaccines (Yes/No) during the year prior to the reference date and during the one year after will be summarized in frequency tables (n, %) per cohort and overall. The four cohorts will be compared using Chi-square test. In case of an overall difference, pair-wise comparison will be carried out using Chi-square test.

All auto-immune diseases will be summarized by descriptive statistics per cohort. The statistics will be computed:

- Number of cases
- Proportion computed as the number of cases divided by the total number of subjects

All the cases of auto-immune disease will be listed with data about exposure status and demographic characteristics.

9.7.9. Statistical models

9.7.9.1. Poisson regression

Poisson regression will be computed using the SAS GENMOD procedure. The dependent variable is the number of events (*Y*). The main model (Model 1) will include the exposure status (exposed (*X=1*) vs. non-exposed (*X=0*)) as a binary independent variable, the age-group ([9-18],[18-25]) as a covariate, and the log-transformed total person-time (*PY*) of each exposed and unexposed cohort as an offset.

CONFIDENTIAL

Female cohort model: $\ln(Y) = \beta_0 + \beta_1 X + \beta_2 Z + \ln(PY)°$

The coefficients β_1 and β_2 are the coefficients associated to exposure effect and age-group, respectively. The risk ratio (exposed/unexposed) will be derived as the exponential of the coefficient associated with the exposure status and its 95% Wald confidence interval.

The SAS code is:

```
PROC  GENMOD data=<filename> ;
 MODEL Y= X  Z / offset=Ln_PY dist=poisson link=log ;
RUN;
```

The same model will be run for comparing the two male cohorts (concurrent vs. historical)

Male cohort model: $\ln(Y) = \beta_0' + \beta_1' X + \beta_2' Z + \ln(PY)°$

The coefficient β_1' is the time effect (concurrent vs. historical) in males.

The exposure effect in females adjusted for temporal effect in males model will be:

$$\ln(Y) = \beta_0 + \beta_{11} X_1 + \beta_{12} X_2 + \beta_{13} X_3 + \beta_2 Z + \ln(PY)°$$

Where:

X_1 is '1' for the exposed female cohort and '0' for the other cohorts;
X_2 is '1' for the concurrent male cohort and '0' for other cohorts;
X_3 is '1' for the historical male cohort and '0' for the other cohort.

The Cervarix effect in females adjusted for the temporal effect in males will be computed as :

$$\beta_{1*} = \beta_{11} - (\beta_{12} - \beta_{13})$$

The SAS code is:

```
PROC  GENMOD data=<filename> ;
 Class X (ref=' 1');
 MODEL Y= X  Z / offset=Ln_PY dist=poisson link=log ;
 Contrast "vaccine adjusted effect in Females" X 1 -1 +1/ / estimate=exp;;
RUN;
```

Note: X=1 for the non-exposed female cohort, 2 for the exposed female cohort, 3 for concurrent male cohort, and 4 for the historical male cohort.

Sensitivity analysis: A Poisson regression model including, in addition to the exposure and the age-group, other covariates

- Region (class variable: 13 regions defined in CPRD)
- Vaccination during the year prior to reference data (2 classes: yes, no)

CONFIDENTIAL

- Use of healthcare resources during the previous year (categories will be quartiles computed from all cohorts)

Covariates occurring in less than 5% of the subjects (percentage will be computed over both exposed and non-exposed cohorts) will not be included in the model. If the number of subjects or the number of events is too low in some categories (for example regions), categories could be grouped.

9.7.9.2. Self-control case-series

Background

The self-control case-series method (SCCS) was developed to investigate associations between vaccination and acute potential adverse events [Farrington, 1996]. The SCCS is based only on cases, and provides consistent estimates of the relative incidence. It controls implicitly for all fixed confounders, that is to say, confounders that do not vary with time over the observation period, e.g. genetics, location, socio-economic status, gender, individual frailty, severity of underlying disease.

The effect estimate is calculated as the ratio of the rate (or hazard) of events in a given post-exposure period (risk period), to the rate of events in the absence of the exposure (control period).

Risk and control periods

For the two co-primary endpoint analysis, the period at risk will be 12 months after the first dose of Cervarix. The control period will also be a period of 12 months after the risk period but separated by a buffer period of 6 months. The new events occurring during the buffer period will not be included in the model (Figure 5). A buffer period is defined to control the risk of late effect of vaccine (risk period for many auto-immune diseases is not clearly defined). We have judged reasonably that the vaccine effect if any will be absent after 18 months.

A sensitivity analysis of the co-primary endpoints will be carried out taking into account individual risk and control periods. For each subject the risk period is the period from the first dose of Cervarix until 6 months after the last dose, and the control period is the period from 6 months after the end of the risk period until month 30. This sensitivity analysis is an SCCS adjusted for the individual variability in the vaccination schedule.

For individual diseases (secondary endpoint), specific risk periods will be defined (Figure 5). SCCS analysis of individual diseases will be carried out only if there are at least 10 cases in the risk and the control periods.

Censoring will not be applied for this analysis, since date of death from the population register will not be available at the time.

CONFIDENTIAL

Figure 5 Risk and control periods for the various endpoints

Statistical Calculations

The statistical calculation will be done using the specific SAS macro developed by Whitaker et al. [Whitaker, 2006] and available online from http://statistics.open.ac.uk/sccs

Mathematical model

Because of relatively short duration of the total period of observation (30 months) compared to a possible age effect, no age effect will not be included in the model.

Each individual ***i*** is observed during a time **[a_i, b_i]**. This interval is the observation period for the individual ***i.*** The observation period for individual ***i*** is then partitioned into k=0,1 periods. Risk period, ***k = 1,*** correspond to an increased risk relative to control period which is coded ***k = 0.***

Conditioning on the exposure history over the entire observation period, we assume that events of interest for individual ***i*** arises as a non-homogeneous Poisson process with rate λ_{ik}. If n_{ik} is the number of events arising for individual ***i*** and risk period ***k***, then

$$n_{ik} \approx \text{Poisson}\left(\lambda_{ik} e_{ik}\right)$$

where e_{ik} is the time spent by subject ***i*** in period ***k***.

Conditioning on the total number of events $n_i = \varphi_i + \beta_k$ arising in **[a_i, b_i]**, the log-likelihood contribution of individual ***i*** is

CONFIDENTIAL

$$l_i = \sum_k n_k \log\left(\frac{\lambda_{ik} e_{ik}}{\sum_s \lambda_{is} e_{is}}\right)$$

With a log-linear model for the Poisson rate of the form

$$\log(\lambda_{ik}) = \varphi_i + \beta_k$$

Where φ_i is an individual effect, and β_k is the exposure effect associated with risk period. The parameter β_k is the log relative incidence.

The log-likelihood estimate of β_k is

$$l(\beta) = \sum_i n_{ik} \log\left(\frac{\exp(\beta_k) e_{ik}}{\sum_r \exp(\beta s) e_{ir}}\right)$$

9.7.9.3. Scan Statistics

Scan statistics will be used to detect temporal clusters of cases. This is done by gradually scanning a window across time noting the number of observed and expected observations inside the window. The scanning window will be an interval of time. Two time window sizes will be used: 2 months and 4 months. The window with the maximum likelihood is the most likely cluster, that is, the cluster least likely to be due to chance. A p-value is assigned to this cluster. Since two time windows will be used, p-values will be compared to a Bonferroni adjusted alpha ($\alpha/2 = 0.025$). Scan statistics use a different probability model depending on the nature of the data. Under the null hypothesis, the observed events occur randomly following a uniform distribution according to a discrete Poisson model during the total observation period.

This scan statistics analysis will be done in each of the four cohorts to detect possible clustering of the two co-primary composite endpoints and of the individual diseases.

9.7.10. Conduct of analyses

9.7.10.1. Sequence of analyses

The study feasibility assessment is intended to confirm that the data are of sufficient quality to confirm the diagnoses, and that the target sample sizes can be reached. A brief report of the study feasibility assessment will tabulate the numbers of confirmed diagnoses and the number of vaccinees in the periods considered for the exposed and historical cohort.

The final analyses will be performed according to a two-step schedule:

1. Between-cohort analysis will be done when the primary and secondary endpoints occurring during the one-year follow-up period will be available for all subjects and the corresponding database will be frozen.
2. The SCCS analysis will be carried out when the primary and secondary endpoints occurring during the 30-month follow-up period will be available for all exposed female subjects and the corresponding database will be frozen.

9.7.10.2. Statistical considerations for interim analyses

There is no interim analysis.

9.7.10.3. Changes from planned analyses

Not applicable.

9.8. Quality control

Validation of clinical outcomes is described in Section 9.3.3 and Section 9.4.2.

Clinical data management will be performed in accordance with applicable GSK standards and data cleaning procedures.

The final study dataset will be archived and stored on a secured, access limited, computer platform SAS Drug Development (SDD) according to GSK Biological Standard Procedures. Specific statistical programs will be written in SAS 9.2 (or higher) and validated according to the GSK standard procedures. The validation of the quality control (QC) of the statistical analysis will be documented. All statistical programs, output files and QC documentation will be saved as read-only files on SDD.

The final study protocol and possible amendments, the final statistical report and the QC document, and the final study report(s) will be archived on a Document management system based on the Documentum platform: Computer Aided Regulatory Submission (CARS).

9.9. Limitations of the research methods

Limitations and recommendations for the research methods

As mentioned previously, Cervarix vaccination administration was done in the UK in the schools and administration of Cervarix is not reported for all subjects in CPRD. The vaccination record in the CPRD GOLD may not be complete; if for a certain subject no vaccination code for Cervarix is registered, vaccination status is uncertain. Due to the potential for under-reporting of HPV vaccination, an unexposed historical cohort (before introduction of Cervarix in the UK) was chosen instead of an unexposed concurrent cohort.

CONFIDENTIAL

Although, diagnosis and/or coding of auto-immune diseases could change over time. Moreover, a link with HES was recently implemented in CPRD GOLD since 2009. The HES data coverage period is actually longer than 2009: currently April 1997 to March 2012. As such, historical controls will have the potential to have linked data in the same way as the more recent exposed group. In order to assess the possible impact of changes in auto-immune diseases diagnosis/coding in CPRD GOLD and HES over time, two male cohorts will also be enrolled: a cohort concurrent to the exposed female cohort and a historical cohort selected at the same time period as the unexposed historical female cohort. However, it is recognized that incidence of auto-immune diseases might vary according to sex as well as according to the age group. These two cohorts will be useful to study potential change over time in CPRD GOLD due to recent access to HES data that might increase the number of auto-immune diseases coded in the database.

A multitude of medcodes are used by the GP to enter diagnosis of autoimmune disease in CPRD GOLD. There is no ICD-10 code mapping. As far as possible , an exhaustive list of medcodes has been defined for each disease and specific algorithms have been developed. In addition, associated “free text” will be reviewed. However, risk of false positive cases (lack of specificity) could not be totally excluded. Lack of specificity can bias the risk ratio estimate to the null hypothesis. There can also be a risk of false negative (lack of sensitivity) however for a rare event, a lack of sensitivity does not bias the risk ratio estimation. Determination of the time to onset of first symptoms is also a limitation. Time between diagnosis and first symptoms is largely variable and depends on the disease and this is why not only medcodes will be used, but also HES and the free text when available.

The definition of the risk period for autoimmune disease is challenging. Too short a risk period would underestimate the actual risk whereas too long a risk period would dilute the actual risk. A period of one year after the first dose of Cervarix will be used for the co-primary composite endpoints. Analysis of the incidence rate of individual diseases will be used for a specific-disease risk period.

These limitations were addressed during a feasibility assessment: with the planned number of subjects for each cohort (65,000), the expected number of autoimmune disease cases to be reviewed should be around 280 cases, which is achievable with the developed methodology for case reviews.

10. PROTECTION OF HUMAN SUBJECTS

10.1. Regulatory and ethical considerations, including the informed consent process

The study will be conducted in accordance with all applicable regulatory requirements.

The study will be conducted in accordance with the Guidelines for Good Pharmacoepidemiology Practices (GPP) [ISPE, 2007], all applicable subject privacy requirements and the guiding principles of the Declaration of Helsinki.

Conduct of the study includes, but is not limited to, the following: CPRD GOLD's internal Independent Scientific Advisory Committee (ISAC) favourable approval of study protocol and any subsequent amendments. This approval has been obtained on August 30, 2012 (ISAC reference 12_086R, see Annex 4).

No patient informed consent will be obtained. The patient information in the CPRD GOLD database is fully coded and GSK Biologicals personnel will not be able to make a link between the data and specific individuals.

The CPRD GOLD has an ethical approval from a Multi-centre Research Ethics Committee (MREC) for purely observational research (i.e. studies that do not include patient involvement [Clinical Practice Research Datalink (CPRD GOLD) Website, 2012]).

10.2. Data privacy

The CPRD GOLD database is a fully coded, MHRA-approved database with an international reputation in the field of drug safety signal evaluation [Williams, 2012].

GSK has a licence to use this database from CPRD GOLD, in order to perform analyses. GSK has access to an online extract from CPRD GOLD which is continuously updated. Data will be not identifiable by GSK as the key-codes are maintained by CPRD GOLD and not available online and never shared with external parties. When GSK requests "free text" to CPRD GOLD, CPRD GOLD has internal processes to secure the maintenance of confidentiality concerning subject identifiers. Identifiers will never be transferred to GSK.

11. MANAGEMENT AND REPORTING OF ADVERSE EVENTS / ADVERSE REACTIONS

This study intends to collect data only on auto-immune diseases recorded in the CPRD GOLD. Where required, the results of this study will be communicated to regulators when the final study report becomes available.

CONFIDENTIAL

12. PLANS FOR DISSEMINATING AND COMMUNICATING STUDY RESULTS

To comply with GPP or other applicable guidelines administrative obligations relating to data collection, archiving data, audits, confidentiality and publications must be fulfilled.

12.1. Posting of information on public registers

Study information from this protocol will be posted on public registers (e.g. GSK Clinical Study Register, clinicaltrials.gov) before the start of analysis as applicable.

12.2. Ownership and publication

12.2.1. Ownership

The source data are the property of the UK Secretary of State. GSK has received the authorisation to use this data for study purposes. All information provided by GSK and data generated as a result of the analysis are property of GSK.

12.2.2. Posting to the clinical trials registers and publication

The results summary will be posted to the GSK Clinical Study Register and other public registers as applicable, in accordance with regulatory and policy mandated timelines. In addition, a manuscript will be submitted to a peer reviewed journal for publication within the policy defined timelines. The manuscript will be co-authored by the CPRD GOLD Research Group, an external expert from the London School of Hygiene and Tropical Medicine (LSHTM), and coordinated by GSK. When manuscript publication in a peer-reviewed journal is not feasible, further study information will be posted to the GSK Clinical Study Register (e.g. write-up).

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

13. REFERENCES

Abrahamson PE, Hall SA, Feudjo-Tepie M, Mitrani-Gold FS, Logie J. The incidence of idiopathic thrombocytopenic purpura among adults: a population-based study and literature review. *Eur J Haematol* 2009;83(2):83-89.

Abramson O, Durant M, Mow W, Finley A, Kodali P, Wong A, et al. Incidence, prevalence, and time trends of pediatric inflammatory bowel disease in Northern California, 1996 to 2006. *J Pediatr* 2010;157(2):233-239.e1.

Allen C, Palta M, D'Alessio DJ. Incidence and differences in urban-rural seasonal variation of type 1 (insulin-dependent) diabetes in Wisconsin. *Diabetologia* 1986;29(9):629-633.

Alonso A, Jick SS, Olek MJ, Hernan MA. Incidence of multiple sclerosis in the United Kingdom : findings from a population-based cohort. *J Neurol* 2007;254(12):1736-41.

Beghi E, Bogliun G. The Guillain-Barre syndrome (GBS). Implementation of a register of the disease on a nationwide basis. Italian GBS Study Group. *Ital J Neurol Sci* 1996;17(5):355-61.

Burnham JM, Shults J, Weinstein R, Lewis JD, Leonard MB. Childhood onset arthritis is associated with an increased risk of fracture: a population based study using the General Practice Database. *Ann Rheum Dis.* 2006;65(8):1074-1079.

Cervarix (Human Papillomavirus vaccine [Types 16, 18] (Recombinant, adjuvanted, adsorbed) EU Product Information) Product Information. February, 2011.

Chao C, Klein NP, Velicer CM, Sy LS, Slezak JM, Takhar H, et al. Surveillance of autoimmune conditions following routine use of quadrivalent human papillomavirus vaccine. *Journal of Internal Medicine.* 2011;271(2):193-203.

Chao C & Jacobsen S. Evaluation of autoimmune safety signal in observational vaccine safety studies. *Human Vaccines & Immunotherapeutics* 2012;8(9):1302-1304.
URL:
http://www.landesbioscience.com/journals/vaccines/2012HV0049.pdf?nocache=2093950154

Clinical Practice Research Datalink (CPRD GOLD) Website. ISAC (Independent Scientific Advisory Committee) – Submission of CPRD GOLD Protocols to ISAC. http://www.cprd.com/ISAC. Accessed: 09 May 2012.

Cooper GS, Stroehla BC. The epidemiology of autoimmune diseases. *Autoimmun Rev.* 2003;2:119-125.

Crawford C, White J. Health Protection Agency survey of Primary Care Trust teenage vaccination programmes; Immunisation, Hepatitis and Blood Safety Department HPA Centre for Infections: December 2009.

CONFIDENTIAL

Department of Health (DH). *Immunisation against infectious disease.* The Green Book, DH publication; 2008 Gateway reference: 7523. The 'Green Book' chapter on Human Papillomavirus (HPV); dh_087787.

Descamps D, Hardt K, Spiessens B, Izurieta P, Verstraeten T, Breuer T, Dubin G. Safety of human papillomavirus (HPV)-16/18 AS04-adjuvanted vaccine for cervical cancer prevention: A pooled analysis of 11 clinical trials. *Hum Vaccine.* 2009;5:332-340.

Diamond Project Group. Incidence and trends of childhood Type 1 diabetes worldwide 1990-1999. *Diabet Med* 2006;23(8):857-66.

Farrington CP, Nash J, Miller E. Case series analysis of adverse reactions to vaccines: a comparative evaluation. *Am J Epidemiol.* 1996;143:1165-1173.

Fishbein HA, LaPorte RE, Orchard TJ, Drash AL, Kuller LH, Wagener DK. The Pittsburgh insulin-dependent diabetes mellitus registry: seasonal incidence. *Diabetologia* 1982;23(2):83-85.

Gonzalez EL, Johansson S, Wallander MA, Rodriguez LA. Trends in the prevalence and incidence of diabetes in the UK: 1996-2005. *J Epidemiol Community Health* 2009;63(4):332-336.

GSK confidential document. Prophylactic HPV-16/18 L1 VLP Vaccine Formulated with AS04. Investigator Brochure Edition 11. March 2012.

Harron KL, McKinney PA, Feltbower RG, Bodansky HJ, Norman PD, Campbell FM, et al. Incidence rate trends in childhood type 1 diabetes in Yorkshire, UK 1978-2007: effects of deprivation and age at diagnosis in the South Asian and non-South Asian populations. *Diabet Med* 2011;28(12):1508-13.

Health Protection Agency Department of Health (HPA). *Routine programme for year 8 girls (12-13 years old) and catch-up campaign for year 13 girls (17-18 years old).* Sheridan A and White J, London, UK; 2010. Annual HPV vaccine uptake in England: 2008/09.

Health Protection Agency (HPA; Immunisation section). Annual HPV vaccine coverage in England in 2010/11. 22 March 2012.

Henderson P, Hansen R, Cameron FL, Gerasimidis K, Rogers P, Bisset WM, et al. Rising incidence of pediatric inflammatory bowel disease in Scotland. *Inflamm Bowel Dis* 2012; 18:999-1005.

Herrinton LJ, Liu L, Lewis JD, Griffin PM, Allison J. Incidence and prevalence of inflammatory bowel disease in a Northern California managed care organization, 1996-2002. *Am J Gastroenterol* 2008;103(8):1998-2006.

Hochberg MC. The incidence of systemic lupus erythematosus in Baltimore, Maryland, 1970-1977. *Arthritis Rheum* 1985;28(1):80-86.

CONFIDENTIAL

Hughes RA, Charlton J, Latinovic R, Gulliford MC. No association between immunization and Guillain-Barré syndrom in the United Kingdom, 1992 to 2000. *Arch Intern Med.* 2006;166(12):1301-1304.

Imkampe A-K & Gulliford MC. Trends in Type 1 diabetes incidence in the UK in 0- to 14-year-olds and in 15- to 34-year-olds, 1991-2008. *Diabetic Medicine* 2011;28(7):811-4

International Society for Pharmacoepidemiology (ISPE). *Guidelines for Good Pharmacoepidemiology Practices.* GPP, http://www.pharmacoepi.org/resources/guidelines_08027.cfm; 2007. Accessed 23 October 2011.

Koobatian TJ, Birkhead GS, Schramm MM, Vogt RL. The use of hospital discharge data for public health surveillance of Guillain-Barre syndrome. *Ann Neurol* 1991;30(4):618-21.

Kostraba JN, Gay EC, Cai Y, Cruickshanks KJ, Rewers MJ, Klingensmith GJ, et al. Incidence of insulin-dependent diabetes mellitus in Colorado. *Epidemiology* 1992;3(3):232-238.

Leese GP, Flynn RV, Jung RT, MacDonald TM, Murphy MJ, & Morris AD. Increasing prevalence and incidence of thyroid disease in Tayside, Scotland: The Thyroid Epidemiology, Audit and Research Study (TEARS). *Clinical Endocrinology* 2008;68:311-316.

Lehtinen M, Paavonen J, Wheeler CM, Jaisamrarn U, Garland SM, Castellsague X et al. Overall efficacy of HPV-16/18 AS04-adjuvanted vaccine against grade 3 or greater cervical intraepithelial neoplasia: 4-year end-of-study analysis of the randomised, double-blind PATRICIA trial. *The Lancet.* 2012;13:89-99.

Lipton RB, Fivecoate JA. High risk of IDDM in African-American and Hispanic children in Chicago, 1985-1990. *Diabetes Care* 1995;18(4):476-482.

MacDonald MJ, Johnson SD, Gottschall JL, Hunter JB, Winter KL, Grieshop RJ, et al. Epidemiology of insulin dependent diabetes before age 20 in Wisconsin, with particular reference to seasonality. *Diabetes Res* 1989;12(4):151-60.

Mayr WT, Pittock SJ, McClelland RL, Jorgensen NW, Noseworthy JH, Rodriguez M. Incidence and prevalence of multiple sclerosis in Olmsted County, Minnesota, 1985-2000. *Neurology* 2003;61(10):1373-1377.

McCarty DJ, Manzi S, Medsger TA, Jr., Ramsey-Goldman R, LaPorte RE, Kwoh CK. Incidence of systemic lupus erythematosus. Race and gender differences. *Arthritis Rheum* 1995;38(9):1260-70.

Medicines and Healthcare products Regulatory Agency (MHRA). Drug Safety Update 2012;6(4):H2. URL: http://www.mhra.gov.uk/Safetyinformation/Safetywarningsalertsandrecalls/UKsafetyPublicAssessmentReports/CON221607 (click on .pdf document at bottom of page; last accessed on 06-FEB-2013).

Musonda P, Farrington CP, Whitaker HJ. Sample sizes for self-controlled case series studies. *Stat Med.* 2006;25:2618-2631.

Myasoedova E, Crowson CS, Kremers HM, Therneau TM, Gabriel SE. Is the incidence of rheumatoid arthritis rising?: results from Olmsted County, Minnesota, 1955-2007. *Arthritis Rheum* 2010;62(6):1576-1582.

Naleway AL, Davis ME, Greenlee RT, Wilson DA, McCarty DJ. Epidemiology of systemic lupus erythematosus in rural Wisconsin. *Lupus* 2005;14(10):862-866.

Nightingale AL, Farmer RD, de Vries CS. Incidence of clinically diagnosed systemic lupus erythematosus 1992-1998 using the UK General Practice Research Database. Pharmacoepidemiol Drug Saf 2006;15(9):656-61.

Riggs JE, Gutmann L, Whited JD. Guillain-Barre syndrome. Another immune-mediated disease with a predilection for young women? *West Virginia Med J* 1989;85(9):382-383.

Schoonen WM, Kucera G, Coalson J, Li L, Rutstein M, Mowat F, Fryzek J, Kaye JA. Epidemiology of immune thrombocytopenic purpura in the General Practice Research Database. *Br J Haematol.* 2009;145:235-244.

Sheridan A and White J. Annual HHPV vaccine uptake in England: 2008/09: Routine programme for year 8 girls (12-13 years old) and catch-up campaign for year 13 girls (17-18 years old); Health Protection Agency Department of Health, 2009.

Sheridan A and White J. Annual HPV vaccine coverage in England in 2009/2010; Health Protection Agency Department of Health, 2010.

Shui IM, Rett MD, Weintraub E, Marcy M, Amato AA, Sheikh SI et al. Guillain-Barré Syndrome incidence in a large United States cohort (2000-2009). *Neuroepidemiology* 2012;39:109-115.

Siegrist CA, Lewis EM, Eskola J, Evans SJ, Black SB. Human papilloma virus immunization in adolescents and young adults: a cohort study to illustrate what events might be mistaken for adverse reactions. *Pediatr Infect Dis J.* 2007;26:979-984.

Simpson KN, Coughlin CM, Eron J, Bussel JB. Idiopathic thrombocytopenia purpura: Treatment patterns and an analysis of cost associated with immunoglobulin and anti-D therapy. *Seminars in Hematology* 1989;35:58-64.

Steed H, Walsh S, Reynolds N. Crohn's disease incidence in NHS Tayside. *Scott Med J* 2010;55(3):22-25.

Verstraeten T, Descamps D, David MP, Zahaf T, Hardt K, Izurieta P, Dubin G, Breuer T. Analysis of adverse events of potential autoimmune aetiology in a large integrated safety database of AS04 adjuvanted vaccines. *Vaccine.* 2008;26:6630-6638.

Whitaker HJ, Farrington CP, Spiessens B and Musonda P. Tutorial in biostatistics: the self-controlled case series method. *Stat Med.* 2006;25:1768-1797.

Williams T, van Staa TP, Puri S, Eaton S. Recent advances in the utility and use of the General Practice Research Database as an example of a UK Primary Care Data resource. *Therapeutic Advances in Drug Safety* 2012;3(2): 89-99.

Yong M, Schoonen WM, Li L, Kanas G, Coalson J, Mowat F, et al. Epidemiology of paediatric immune thrombocytopenia in the General Practice Research Database. *Br J Haematol* 2010;149(6):855-864.

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Annex 1 List of stand-alone documents

No.	Document Reference No	Date	Title
1	116239 (EPI-HPV-040 VS UK)	09-JUL-2013	List of stand-alone documents
2	116239 (EPI-HPV-040 VS UK)	11-FEB-2013	ENCePP Checklist for study
3	116239 (EPI-HPV-040 VS UK)	11-FEB-2013	Glossary of terms
4	116239 (EPI-HPV-040 VS UK)	11-FEB-2013	ISAC evaluation of protocols for research involving CPRD GOLD
5	116239 (EPI-HPV-040 VS UK)	11-FEB-2013	Algorithms
6	116239 (EPI-HPV-040 VS UK)	11-FEB-2013	Example of table and figure templates
7	116239 (EPI-HPV-040 VS UK)	11-FEB-2013	Trademarks
8	116239 (EPI-HPV-040 VS UK)	09-JUL-2013	Protocol sponsor signatory approval

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Annex 2 ENCePP Checklist for study protocols

ANNEX 2. ENCePP Checklist for study protocols

Section 1: Milestones	Yes	No	N/A	Page Number(s)
1.1 Does the protocol specify timelines for				
1.1.1 Start of data collection[1]	X	☐	☐	24-25
1.1.2 End of data collection[2]	X	☐	☐	25-25
1.1.3 Study progress report(s)	☐	☐	X	
1.1.4 Interim progress report(s)	☐	☐	X	
1.1.5 Registration in the EU PAS register	X	☐	☐	See comments
1.1.6 Final report of study results.	X	☐	☐	See comments

Comments:

For 1.1.3 and 1.1.4 – no progress reports are planned for this study.

For 1.1.5 and 1.1.6 – see regulatory submission cover letter. The EU PAS register number will be generated at the time of the final version of the protocol.

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

[2] Date from which the analytical dataset is completely available.

Section 2: Research question	**Yes**	**No**	**N/A**	**Page Number(s)**
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)	X	☐	☐	14-16
2.1.2 The objective(s) of the study?	X	☐	☐	18-19
2.1.3 The target population? (i.e. population or subgroup to whom the study results are intended to be generalised)	X	☐	☐	23
2.1.4 Which formal hypothesis(-es) is (are) to be tested?	X	☐	☐	32-33
2.1.5 If applicable, that there is no a priori hypothesis?	☐	X	☐	

Comments:

Section 3: Study design	**Yes**	**No**	**N/A**	**Page Number(s)**
3.1 Is the study design described? (e.g. cohort, case-control, randomised controlled trial, new or alternative design)	X	☐	☐	20-21
3.2 Does the protocol specify the primary and secondary (if applicable) endpoint(s) to be investigated?	X	☐	☐	26-27
3.3 Does the protocol describe the measure(s) of effect? (e.g. relative risk, odds ratio, deaths per 1000 person-years, absolute risk, excess risk, incidence rate ratio, hazard ratio, number needed to harm (NNH) per year)	X	☐	☐	35-40

Comments:

The statistical analysis plan is detailed in Section 9.7.7 of the protocol (pages 35-40).

Section 4: Source and study populations	**Yes**	**No**	**N/A**	**Page Number(s)**
4.1 Is the source population described?	X	☐	☐	28-29
4.2 Is the planned study population defined in terms of:				
4.2.1 Study time period?	X	☐	☐	24-25
4.2.2 Age and sex?	X	☐	☐	24-25
4.2.3 Country of origin?	X	☐	☐	22-23
4.2.4 Disease/indication?	X	☐	☐	16
4.2.5 Co-morbidity?	☐	☐	X	
4.2.6 Seasonality?	☐	☐	X	
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	☐	☐	24-25 & 30-31

Comments:

Section 5: Exposure definition and measurement	**Yes**	**No**	**N/A**	**Page Number(s)**
5.1 Does the protocol describe how exposure is defined and measured? (e.g. operational details for defining and categorising exposure)	X	☐	☐	26-28
5.2 Does the protocol discuss the validity of exposure measurement? (e.g. precision, accuracy, prospective ascertainment, exposure information recorded before the outcome occurred, use of validation sub-study)	X	☐	☐	40-41
5.3 Is exposure classified according to time windows? (e.g. current user, former user, non-use)	X	☐	☐	21
5.4 Is exposure classified based on biological mechanism of action and taking into account the pharmacokinetics and pharmacodynamics of the drug?	☐	☐	X	
5.5 Does the protocol specify whether a dose-dependent or duration-dependent response is measured?	☐	☐	X	

Comments:

Section 6: Endpoint definition and measurement	**Yes**	**No**	**N/A**	**Page Number(s)**
6.1 Does the protocol describe how the endpoints are defined and measured?	X	☐	☐	34-35
6.2 Does the protocol discuss the validity of endpoint measurement? (e.g. precision, accuracy, sensitivity, specificity, positive predictive value, prospective or retrospective ascertainment, use of validation sub-study)	X	☐	☐	37 & 40/41

Comments:

Section 7: Confounders and effect modifiers	**Yes**	**No**	**N/A**	**Page Number(s)**
7.1 Does the protocol address known confounders? (e.g. collection of data on known confounders, methods of controlling for known confounders)	X	☐	☐	37
7.2 Does the protocol address known effect modifiers? (e.g. collection of data on known effect modifiers, anticipated direction of effect)	X	☐	☐	37

Comments:

Section 8: Data sources	**Yes**	**No**	**N/A**	**Page Number(s)**
8.1 Does the protocol describe the data source(s) used in the study for the ascertainment of:				
8.1.1 Exposure? (e.g. pharmacy dispensing, general practice prescribing, claims data, self-report, face-to-face interview, etc.)	X	☐	☐	28-30
8.1.2 Endpoints? (e.g. clinical records, laboratory markers or values, claims data, self-report, patient interview including scales and questionnaires, vital statistics, etc.)	X	☐	☐	26-27
8.1.3 Covariates?	X	☐	☐	27-29
8.2 Does the protocol describe the information available from the data source(s) on:				
8.2.1 Exposure? (e.g. date of dispensing, drug quantity, dose, number of days of supply prescription, daily dosage, prescriber)	X	☐	☐	28-29
8.2.2 Endpoints? (e.g. date of occurrence, multiple event, severity measures related to event)	X	☐	☐	26-27
8.2.3 Covariates? (e.g. age, sex, clinical and drug use history, co-morbidity, co-medications, life style, etc.)	X	☐	☐	27-29
8.3 Is a coding system described for:				
8.3.1 Diseases? (e.g. International Classification of Diseases (ICD)-10)	X	☐	☐	28-29
8.3.2 Endpoints? (e.g. Medical Dictionary for Regulatory Activities (MedDRA) for adverse events)	X	☐	☐	26-27
8.3.3 Exposure? (e.g. WHO Drug Dictionary, Anatomical Therapeutic Chemical (ATC)Classification System)	X	☐	☐	27-29
8.4 Is the linkage method between data sources described? (e.g. based on a unique identifier or other)	X	☐	☐	29-30

Comments:

Section 9: Study size and power	**Yes**	**No**	**N/A**	**Page Number(s)**
9.1 Is sample size and/or statistical power calculated?	X	☐	☐	30-31

Comments:

The study is using CPRD GOLD + HES data when available.

Section 10: Analysis plan	**Yes**	**No**	**N/A**	**Page Number(s)**
10.1 Does the plan include measurement of excess risks?	☐	☐	X	
10.2 Is the choice of statistical techniques described?	X	☐	☐	36-41
10.3 Are descriptive analyses included?	X	☐	☐	36
10.4 Are stratified analyses included?	☐	☐	X	
10.5 Does the plan describe methods for adjusting for confounding?	X	☐	☐	36-41
10.6 Does the plan describe methods addressing effect modification?	X	☐	☐	36-41

Comments:

Section 11: Data management and quality control	**Yes**	**No**	**N/A**	**Page Number(s)**
11.1 Is information provided on the management of missing data?	X	☐	☐	36
11.2 Does the protocol provide information on data storage? (e.g. software and IT environment, database maintenance and anti-fraud protection, archiving)	X	☐	☐	42
11.3 Are methods of quality assurance described?	X	☐	☐	40
11.4 Does the protocol describe possible quality issues related to the data source(s)?	X	☐	☐	40-41
11.5 Is there a system in place for independent review of study results?	☐	X	☐	

Comments:

Section 12: Limitations	**Yes**	**No**	**N/A**	**Page Number(s)**
12.1 Does the protocol discuss:				
12.1.1 Selection biases?	X	☐	☐	40-41
12.1.2 Information biases? (e.g. anticipated direction and magnitude of such biases, validation sub-study, use of validation and external data, analytical methods)	X	☐	☐	40-41
12.2 Does the protocol discuss study feasibility? (e.g. sample size, anticipated exposure, duration of follow-up in a cohort study, patient recruitment)	X	☐	☐	See separate document
12.3 Does the protocol address other limitations?	X	☐	☐	40-41

Comments:

A feasibility assessment report has been submitted together with the protocol.

Section 13: Ethical issues	**Yes**	**No**	**N/A**	**Page Number(s)**
13.1 Have requirements of Ethics Committee/Institutional Review Board approval been described?	X	☐	☐	42 + Protocol Annex 4
13.2 Has any outcome of an ethical review procedure been addressed?	☐	X	☐	
13.3 Have data protection requirements been described?	X	☐	☐	42

Comments:

Section 14: Amendments and deviations	**Yes**	**No**	**N/A**	**Page Number(s)**
14.1 Does the protocol include a section to document future amendments and deviations?	☐	X	☐	

Comments:

Section 15: Plans for communication of study results	**Yes**	**No**	**N/A**	**Page Number(s)**
15.1 Are plans described for communicating study results (e.g. to regulatory authorities)?	X	☐	☐	43
15.2 Are plans described for disseminating study results externally, including publication?	X	☐	☐	43

Comments:

Name of the main author of the protocol: [redacted]

Date: 11/FEB/2013

Signature: [redacted]

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Annex 3 GLOSSARY OF TERMS

Coded:	Information is associated with a subject number i.e. a code number. Coded information can only be linked back to the individual via a key code i.e. a listing of the research participant and their code. Within the pharmaceutical industry coding data is the usual mechanism used for protecting an individual's research data. The key code is kept secure, usually by the investigator, and GSK researchers cannot identify the research individual other than in exceptional and controlled circumstances.
Cohort study:	A form of epidemiology study where subjects in a study population are classified according to their exposure status and followed over time (prospective / retrospective) to ascertain the outcome(s) (disease).
Eligible:	Qualified for enrolment into the study based upon strict adherence to inclusion/exclusion criteria.
eTrack:	GSK's tracking tool for clinical/epidemiological trials.
Medcode	The Medcodes are the abbreviated terms which mean CPRD GOLD medical codes. Medcodes consisting of READ codes are used to enter medical diagnosis in the CPRD GOLD database.
Non-interventional (observational) Human Subject Research:	Studies where medicinal products, should they be administered, are prescribed in normal (routine) medical practice. No medical care or medical/scientific procedures as required in a research protocol are administered to participants except as part of routine medical care.
Post-Authorization Safety Study (PASS)	A pharmacoepidemiological study or a clinical trial carried out in accordance with the terms of the marketing authorisation, conducted with the aim of identifying or quantifying a safety hazard relating to an authorised medicinal product. This includes all GSK sponsored non-interventional studies and clinical trials conducted anywhere in the world that are in accordance with the terms of the European marketing authorisation and where the investigation of safety is the specific stated objective.

CONFIDENTIAL

Self-control case-series (SCCS):	Statistical method for assessing the association between a transient exposure and an adverse event. The method was developed to study adverse reactions to vaccines. The method uses only cases; no controls are required as the cases act as their own controls. Each case's given observation time is divided into control and risk periods. Risk periods are defined during or after the exposure. The method estimates a relative incidence rate, that is, the incidence in risk periods relative to the incidence in control periods. Time-varying confounding factors such as age can be allowed for by dividing up the observation period further into age categories. An advantage of the method is that confounding factors that do not vary with time, such as genetics, location, socio-economic status are controlled for implicitly.
Study population:	Sample of population of interest.
Subject:	Term used throughout the protocol to denote an individual who has been contacted in order to participate or participates in the clinical/epidemiological study, or a person about whom some medical information have been recorded in a database.
Subject number:	A unique number identifying a subject, assigned to each subject consenting to participate in the study.
Targeted Safety Study (TSS)	Studies specifically planned or conducted to examine an actual or hypothetical safety concern in a product marketed anywhere in the world. This includes any GSK sponsored pharmaco-epidemiology study or clinical trial conducted anywhere in the world with the aim of identifying or quantifying a safety hazard. Although all clinical trials collect safety information as a matter of routine, only those initiated to examine a specific safety concern are considered a targeted safety study.

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Annex 4 ISAC Evaluation Of Protocols For Research Involving CPRD GOLD

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

ISAC EVALUATION OF PROTOCOLS FOR RESEARCH INVOLVING GPRD DATA

FEED-BACK TO APPLICANTS

CONFIDENTIAL			*by e-mail*
PROTOCOL NO:	12_086R		
PROTOCOL TITLE:	An observational cohort study to assess the risk of autoimmune diseases in adolescent and young adult women aged 9 to 25 years exposed to Cervarix in the United Kingdom [eTrack: 116239, EPI-HPV-040 VS UK]		
APPLICANT:	Dr [redacted], Director, Observational Data Analytics, Worldwide Epidemiology, GSK, [redacted]		
APPROVED ☒	**APPROVED SUBJECT TO MINOR AMENDMENT (resubmission not required)** ☐	**REVISION/ RESUBMISSION REQUESTED** ☐	**REJECTED** ☐

INSTRUCTIONS:

Please include your response/s to the Reviewer's feedback below only if you are required to Revise/ Resubmit your protocol.

Protocols with an outcome of 'Approved' or 'Approved subject to minor amendments' do not require resubmission to the ISAC

REVIEWER COMMENTS:

Protocol 12_086R is approved.

DATE OF ISAC FEEDBACK:	30 August 2012
DATE OF APPLICANT FEEDBACK:	

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Annex 5 Algorithms

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Annex 5 Version 4.0 (11 FEB 2013)

eTrack study number and Abbreviated Title	116239 (EPI-HPV-040 VS UK)
Title	An observational cohort study to assess the risk of autoimmune diseases in adolescent and young adult women aged 9 to 25 years exposed to Cervarix® in the United Kingdom

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

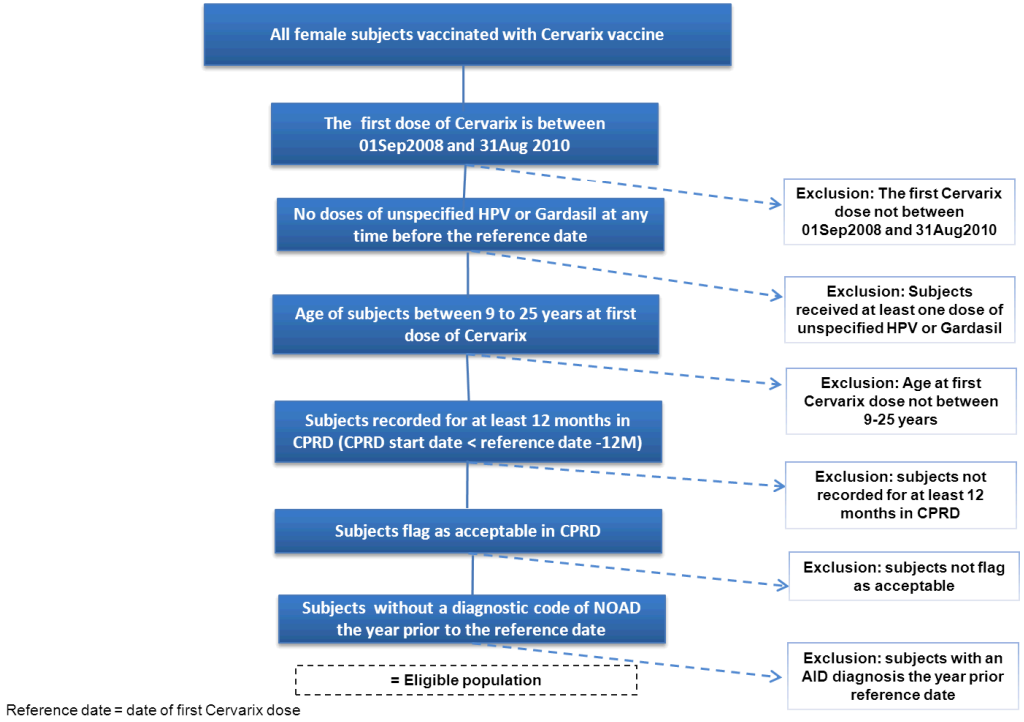
TABLE OF CONTENTS

PAGE

CONFIDENTIAL

Section 1: Identification of the study populations

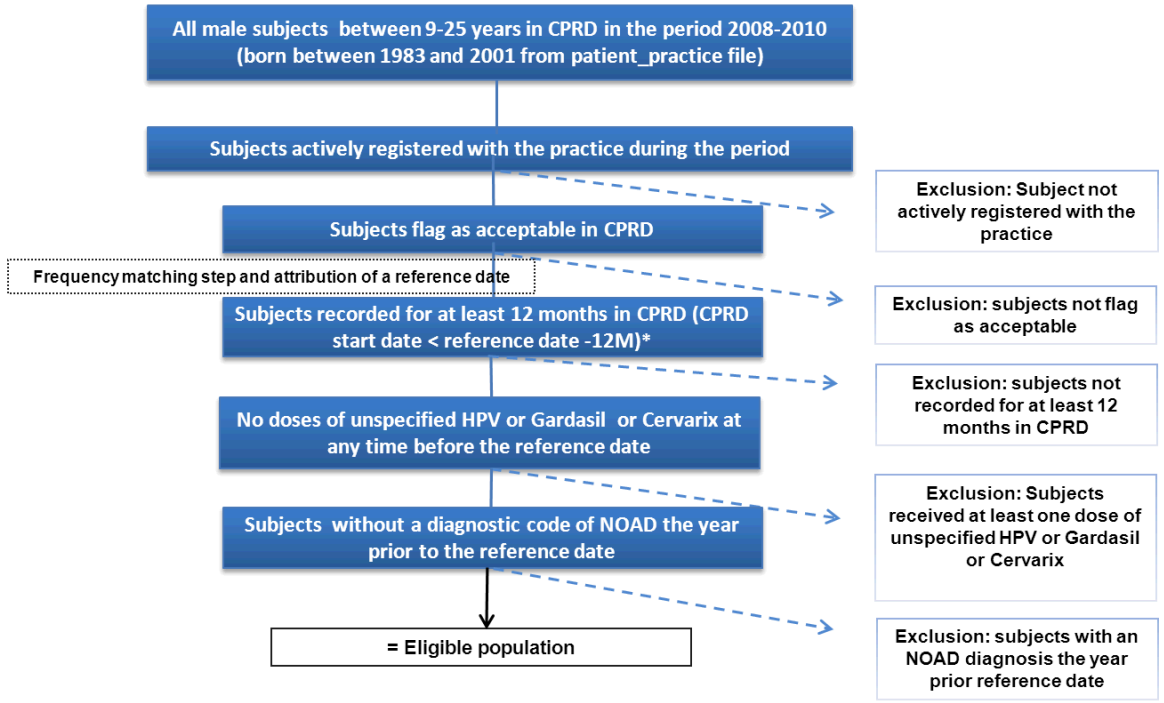
Exposed cohort (Cervarix vaccinated females cohort)



Reference date = date of first Cervarix dose

CONFIDENTIAL

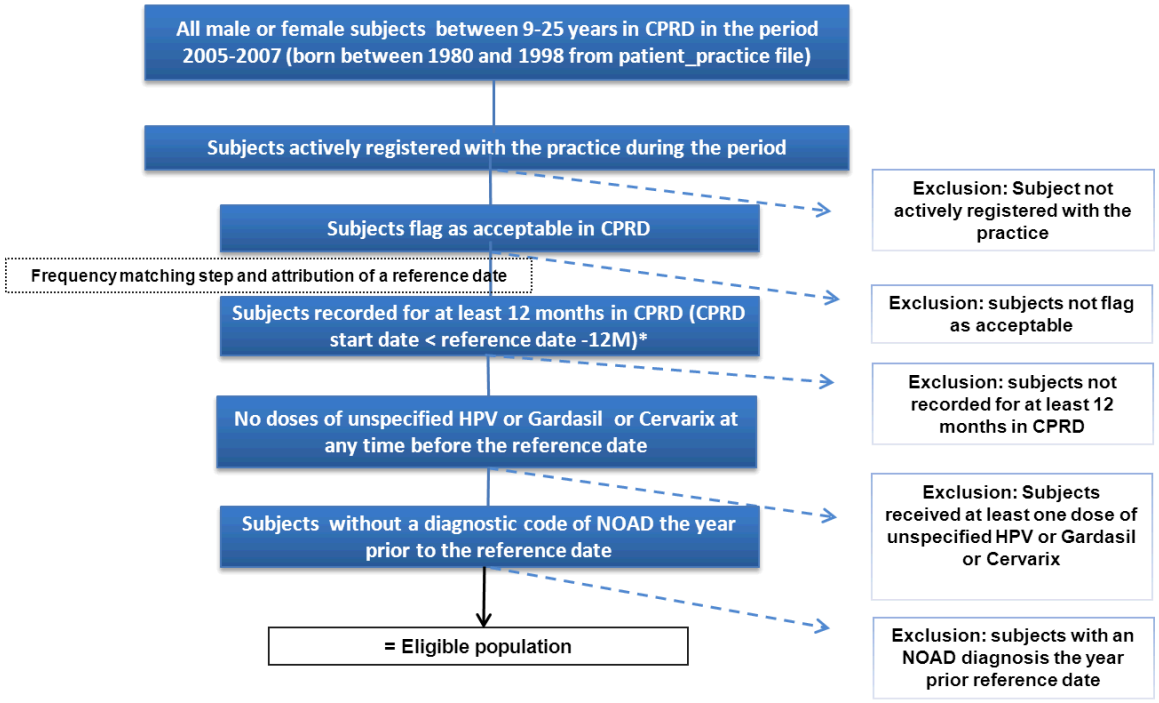
Unexposed Concurrent Male Cohort



*Reference date = random date between 01SEP2008 and 31AUG2010

CONFIDENTIAL

Unexposed Historical Cohorts



*Reference date = random date between 01SEP2005 and 31AUG2007

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Section 2: Key Variables (variables name)

I Directly extracted from CPRD-GOLD:

1. Patient Identifier
2. Gender
3. Birth Month
4. Birth Year
5. Practice Identifier
6. Region
7. Date of death
8. First registration date
9. Current registration date
10. Registration gaps
11. Registration status
12. Transfer Out Date
13. Transfer Out Reason
14. Up To Standard Date
15. Acceptable Patient Flag
16. Matching CPRD-HES

I Derived from CPRD-GOLD:

A. Subject's characteristics

17. Date of Cervarix vaccination
18. Date of unspecified HPV or Gardasil vaccine
19. Date of any other vaccine
20. Date of birth
21. CPRD start date

CONFIDENTIAL

22. Health care resource utilization
23. Date of autoimmune disease diagnosis

B. Autoimmune disease - Endpoints

24. Multiple Sclerosis
25. Transverse Myelitis
26. Optic neuritis
27. Guillain Barre Syndrome
28. Other demyelinating diseases
29. Autoimmune Uveitis
30. Systemic Lupus Erythematosus
31. Rheumatoid Arthritis
32. Juvenile Rheumatoid Arthritis
33. Still's disease
34. Psoriatic Arthritis
35. Ankylosing Spondylitis
36. Idiopathic Thrombocytopenic Purpura
37. Autoimmune Haemolytic Anaemia
38. Type 1 diabetes mellitus
39. Autoimmune thyroiditis
40. Crohn's disease
41. Ulcerative colitis
42. Autoimmune Hepatitis

Notes: Algorithms to be used for data extraction are described from pages 10 to 22. Annexes to this document (starting page 23) detail the Medical codes or vaccine classification or product codes (available upon request).

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Variables directly extracted from CPRD-GOLD

	Column name	Field name	Description	CPRD-GOLD file	Codelist	Type of variable	Comments
1	***Patient Identifier***	***patid***	unique identifier given to a patient	Patient	-	Num8.	-
2	***Patient Gender***	***gender***	Patient's gender	Patient	Lookup SEX	Num3.	-
3	***Birth Month***	***mob***	Patient's month of birth (for those aged under 16)	Patient	-	Num8.	Limited information
4	***Birth Year***	***yob***	Patient's year of birth	Patient	-	Num8.	Value+1800
5	***Practice Identifier***	***pracid***	unique identifier given to a specific practice	Patient_practice	-	Num3.	Link to patient file by patid variable
6	***Practice Region***	***region***	Practice region: Value to indicate where in the UK the practice is based	Patient_practice	Lookup PRG	Char22	Link to patient file by patid variable
7	***Death Date***	***deathdate***	Date of death of patient – derived using an algorithm	Patient	dd/mm/yyyy[1]	Num8.	-
8	***First Registration Date***	***frd***	First registration date: Date the patient first registered with the practice. If patient only has 'temporary' records, the date is the first encounter with the practice; if patient has 'permanent' records it is the date of the first 'permanent' record (excluding preceding temporary records)	patient	dd/mm/yyyy[1]	Num8.	-
9	***Current Registration Date***	***crd***	Date the patient's current period of registration with the practice began (date of the first 'permanent' record after the latest transferred out period). If there are no 'transferred out periods', the date is equal to 'frd'	patient	dd/mm/yyyy[1]	Num8.	-

CONFIDENTIAL

	Column name	Field name	Description	CPRD-GOLD file	Codelist	Type of variable	Comments
10	***Registration Gaps***	***reggap***	Number of days missing in the patients registration details	patient	PAT_GAP[2]	Num8.	-
11	***Registration Status***	***regstat***	Registration status: Status of registration detailing gaps and temporary patients	patient	PAT_STAT[3]	Num3.	-
12	***Transfer Out Date***	***tod***	Date the patient transferred out of the practice	Patient	dd/mm/yyyy[1]	Num8.	-
13	***Transfer Out Reason***	***toreason***	Reason the patient transferred out of the practice. Includes 'Death' as an option	Patient	Lookup TRA	Num3.	-
14	***Up To Standard Date***	***uts***	Date at which the practice data is deemed to be of research quality. Derived using an algorithm that primarily looks at practice death recording and gaps in the data	practice	dd/mm/yyyy[1]	Num8.	-
15	***Acceptable Patient Flag***	***accept***	Flag to indicate whether the patient has met certain quality standards: 1 = acceptable, 0 = unacceptable	Practice	Boolean	Num3.	-
16	***Matching CPRD-HES***	***HES_e***	Flag (0,1) indicating whether patient is eligible for linkage to HES data	HES_eligibility		Num3.	The flag should be equal to 1

[1] **dd/mm/yyyy**: Valid dates are in the format DD/MM/YYYY. Missing dates are NULL, and invalid dates are set to 01/01/2500

[2] **PAT_GAP:** Number of days between patient's transferred out date and re-registration date for the patient's 'transferred out periods', regardless of whether the transfer was internal or not.

[3] **PAT_STAT:** Transferred out period is the time between a patient transferring out and re-registering at the same practice. If the patient has transferred out for a period of more than 1 day, and the transfer is not internal, this value is incremented. 0 means continuous registration, 1 means one 'transferred out period', 2 means two periods, etc. If the patient only has 'temporary' records then this value is set to 99.

Variables Derived from CPRD-GOLD

	Column Name	Algorithms	Comments
17	***Date of Cervarix vaccination***	Search for the subject in *Immunisation file* where: – Immstype equals 67 (HPVCER) and status equals 1 – Retrieve the eventdates For the subjects with at least one recorded dose in *Immunisation file*: 1) Search for additional Cervarix vaccination in *Therapy file* where: – Cervarix prodcode = 36952 The vaccinations from *Therapy* file will be considered as additional Cervarix vaccination if the eventdate is not equals to eventdate (+/- 14 days) from *Immunisation file*. 2) Search for additional Cervarix vaccination in *Clinical file* (medcode=93489 93621 95554): - if the eventdate is equal to eventdate from *Immunisation or Therapy file* then vaccination is similar than the one from *Immunisation or Therapy file*. - if the eventdate is not equal to eventdate but in an interval of +/- 14 days from *Immunisation or Therapy file* then the vaccination is not an additional unspecified HPV vaccination – ntPV. the date is different, the dose is considered as an unspecified additional doses - if the eventdate is not equal to eventdate +/- 14 days from *Immunisation or Therapy file* then the vaccination is an additional unspecified HPV vaccination If the 1st dose of Cervarix is between 01Sep2008 and 31Dec2010, the subject will be included in the exposed cohort. The date of 1st dose of Cervarix vaccination is the reference date for exposed cohort. Eventdate of all of recorded Cervarix doses will be retrieved.	

	Column Name	Algorithms	Comments
18	***Date of unspecified HPV or Gardasil vaccine***	This variable will retrieve a date of unspecified HPV OR Gardasil vaccination. Search for the subject in *Immunisation file* where: – Medcode in (93489, 93621, 95554) (HPV 1st, 2d, 3rd dose) AND immstype equals 58 or not specified – Retrieve the eventdates Search for additional unspecified HPV or Gardasil vaccine in *Therapy file* where: – prodcode = 32424 37955 /Gardasil prodcode =32147 The vaccinations from *Therapy file* will be considered as additional unspecified HPV or Gardasil vaccination if the eventdate is not equals to eventdate (+/- 14 days) from *Immunisation file*. Search for additional HPV vaccinations in *Clinical file* (medcode=93489 93621 95554): - if the eventdate is equal to eventdate from *Immunisation or Therapy file* then vaccination is similar than the one from *Immunisation or Therapy file*. - if the eventdate is not equal to eventdate but in an interval of +/- 14 days from *Immunisation or Therapy file* then the vaccination is not an additional unspecified HPV vaccination – ntPV. the date is different, the dose is considered as an unspecified additional doses. - if the eventdate is not equal to eventdate +/- 14 days from *Immunisation or Therapy file* then the vaccination is an additional unspecified HPV vaccination	
19	***Date of any other vaccine***	This variable checks if a vaccine (other than HPV) was administered during the year before the reference date. Search for the subject in *Immunisation file and Therapy file* if a medcode (for vaccine) exist, retrieve eventdate, immstype and medcode and status=1 for *Immunisation file*. Search in *Therapy file* if a prodcode for vaccine exist, retrieve eventdate, drugsubstance, productname. Eventdate of vaccination should be between the reference date -365 and end of follow-up. Cross tabulation of medcode/prodcode and names of vaccines in Annex 20.	Annex 20
20	***Date of birth***	Date of birth will be derived from month of birth (mob) and year of birth (yob) in *Patient file*. If month of birth and year of birth are present, the date of birth will be read as “15mmyyyy”. If month	

	Column Name	Algorithms	Comments
		of year is not present, it will be read as “30JUNyyyy”.	
21	***CPRD Start Date***	From *Patient and Practice file*: If crd < Up to Standard Date then CPRD Start Date= Up to Standard Date (uts) If crd > Up To Standard Date then CPRD Start Date=Current registration Date (crd)	
22	***Health care resource utilization***	The number of GP/primary care consultations during the year before the reference date will be retrieved from *Consultation file*.	
23	***Date of autoimmune disease diagnosis***	The autoimmune diagnosis will be identified by applying the algorithm 24 from 42. If the same recorded medcode has more than one event date, then the first event will be used as the first date of autoimmune diagnosis.	
24-42	***Autoimmune disease name***	Each autoimmune disease will be retrieved from algorithms (see below algorithm 24-42)	

CONFIDENTIAL

Algorithm for each autoimmune disease

	Column Name	Algorithms	Comments
24	**Multiple Sclerosis**	Multiple sclerosis (MS) cases will be identify where: – in *Clinical and Referral file*: a medcode for MS is listed (see annex_1) AND/OR – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for MS is listed (see annex_1) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to MS in the study period will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to MS is listed (see annex_21) – in *Test file*: MRI scan (enttype=300) and cerebrospinal fluid examination (enttype=410)	Annex 1
25	**Transverse Myelitis**	Transverse Myelitis (TM) cases will be identified where: – in *Clinical and Referral file*: a medcode for TM is listed (see annex_2) – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for TM is listed (see annex_2) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to TM in the study period will be retrieved and the case will be sent for expert review.	Annex 2

	Column Name	Algorithms	Comments
26	**Optic neuritis**	Optic neuritis cases will be identified where: – in *Clinical and Referral file*: a medcode for Optic neuritis is listed (see annex_3) – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for optic neuritis is listed (annex_3) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to Optic neuritis in the study period will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to optic neuritis is listed (see annex_21)	Annex 3
27	**Guillain Barré Syndrome**	Guillain Barré Syndrome (GBS) cases will be identified where: – in *Clinical and Referral file*: a medcode for GBS is listed (see annex_4) – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for GBS is listed (see annex_4) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to GBS in the study period will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to MS is listed (see annex_21) – in *Test file*: Nerve conduction studies (enttype=343) and cerebrospinal fluid examination (enttype=410)	Annex 4

	Column Name	Algorithms	Comments
28	**Other demyelinating diseases**	Other demyelinating disease cases were identified where: – in *Clinical and Referral file*: a medcode for Other demyelinating disease is listed (see annex_5) – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for Other demyelinating disease is listed (see annex_5) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to Other demyelinating disease (medical, product codes) in the study period will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to Other demyelinating disease is listed (see annex_21)	Annex 5
29	**Autoimmune Uveitis**	Autoimmune Uveitis cases will be identified where: – in *Clinical and Referral file*: a medcode for Autoimmune Uveitis is listed (see annex_6) – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for Autoimmune Uveitis is listed (see annex_6) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to Autoimmune Uveitis (medical, product codes) in the study period will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to Autoimmune Uveitis is listed (see annex_21)	Annex 6

	Column Name	Algorithms	Comments
30	**Systemic Lupus Erythematosus**	Systemic Lupus Erythematosus (SLE) cases will be identified where: – in *Clinical and Referral file*: a medcode for SLE is listed (see annex_7) – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for SLE is listed (see annex_7) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to SLE (medical, product codes) in the study period will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to SLE is listed (see annex_21) – in *Test file*: Clotting test (enttype=215); Erythrocyte sedimentation (enttype=273); Anti nuclear antibodies (enttype=279); C Reactive protein (enttype=280); Anti Smooth muscles (enttype=328) and Complement tests (enttype=422)	Annex 7
31	**Rheumatoid Arthritis**	Rheumatoid Arthritis (RA) cases will be identified where: – in *Clinical and Referral file*:: a medcode for RA is listed (see annex_8) – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for RA is listed (see annex_7) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to RA in the study period will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to RA is listed (see annex_21)	Annex 8

	Column Name	Algorithms	Comments
		– in *Test file*: Erythrocyte sedimentation (enttype=273); Anti nuclear antibodies (enttype=279); C Reactive protein (enttype=280); Rheumatoid factor (enttype=292); Anti Smooth muscles (enttype=328); DNA blinding autoantibodies (enttype=330); LE cells (enttype=364); Serum fibrinogen level (enttype=397); HLA tissue typing (enttype=405) and Complement tests (enttype=422)	
32	**Juvenile Rheumatoid Arthritis**	Juvenile Rheumatoid Arthritis (JRA) cases will be identified where: – in *Clinical and Referral file*:: a medcode for RA is listed (see annex_9) – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for JRA is listed (see annex_8) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to JRA in the study period will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to JRA is listed (see annex_21)	Annex 9
33	**Still's disease**	Still's disease cases will be identified where: – in *Clinical and Referral file*: a medcode for Still's disease is listed (see annex_10) – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for Still's disease is listed (see annex_10) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to Still's disease in the study period will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to Still's disease is listed (see annex_21)	Annex 10

	Column Name	Algorithms	Comments
34	**Psoriatic Arthritis**	Psoriatic Arthritis (PA) cases will be identified where: – in *Clinical and Referral file*: a medcode for PA is listed (see annex_11) – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for PA is listed (see annex_11) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to PA in the study period we will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to PA is listed (see annex_21)	Annex 11
35	**Ankylosing Spondylitis**	Ankylosing Spondylitis (AS) cases will be identified where: – in *Clinical and Referral file*: a medcode for AS is listed (see annex_12) – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for AS is listed (see annex_12) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to AS in the study period will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to AS is listed (see annex_21)	Annex 12

	Column Name	Algorithms	Comments
36	**Idiopathic Thrombocytopenic Purpura**	Idiopathic Thrombocytopenic Purpura (ITP) cases will be identified where: – in *Clinical and Referral file*: a medcode for ITP is listed (see annex_13) – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for ITP is listed (see annex_13) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to ITP in the study period will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to AS is listed (see annex_21) – in *Test file*: a platelets test (enttype=189)	Annex 13
37	**Autoimmune Haemolytic Anaemia**	Autoimmune (AI) Haemolytic Anaemia cases will be identified where: – in *Clinical and Referral file*: a medcode for AI Haemolytic Anaemia is listed (see annex_14) – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for AI Haemolytic Anaemia is listed (see annex_14) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to AI Haemolytic Anaemia in the study period will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to AI Haemolytic Anaemia is listed (see annex_21)	Annex 14

	Column Name	Algorithms	Comments
38	**Type 1 diabetes mellitus**	Diabetes type 1 mellitus cases will be identified where: – in *Clinical and Referral file*: a medcode for Diabetes is listed (see annex15) – in *HES*: an ICD10 diagnosis code for Diabetes is listed (see annex_15) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to Diabetes in the study period will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to Diabetes is listed (see annex_21) – in *Test file*: diagnostic tests for diabetes Diagnostic tests: Test* / Entity type / Abnormal range Glucose tolerance test / 222 / >140 mg/dl Fasting glucose / 274 / >100 mg/dl HbA1c – diabetic control / 275 / >36 mmol/mol or >5.7% *prediabetes and potential abnormal results are flagged as abnormal	Annex 15
39	**Autoimmune Thyroiditis**	AI Thyroiditis cases will be identified where: – in *Clinical and Referral file*: a medcode for AI Thyroiditis is listed (see annex_16) – in *HES (HES_diagnosis_epi file)*: an ICD10 diagnosis code for AI Thyroiditis is listed (see annex_16) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to AI Thyroiditis in the study period will be retrieved and the case will be sent for	Annex 16

CONFIDENTIAL

	Column Name	Algorithms	Comments
		expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to AI Thyroiditis is listed (see annex_21) – in *Test file*: diagnostic tests for AI thyroiditis: enttype=203 (Thyroid stimulating hormone), 236 (Thyroid function), 329 (Thyroid autoantibodies), 400 (Serum parathyroid hormone) – in *Clinical/Test file*: where enttype=110 (=Thyroid disease)	
40-41	**Inflammatory Bowel Diseases**	IBD, Ulcerative Colitis or Crohn diseases cases will be identified where: – in *Clinical and Referral file*: a medcode for IBD/Crohn or UC is listed (see annexes_17, 18,19) – in *HES*: an ICD10 diagnosis code is listed (see annexes_17, 18,19) Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to IBD/UC/Crohn disease in the study period will be retrieved and the case will be sent for expert review. Additional information will be retrieved in order to complete the patient profile, the eventdate should be 1 year before the reference date or during the follow-up period: – in *Therapy file*: a prodcode related to IBD/UC/Crohn disease is listed (see annex_21) – in *Test file*: Red blood cell count (enttype=194); Colonoscopy (enttype=296); Sigmoidoscopy (enttype=297); Abdominal xray (enttype=341) If a medcode of IBD (annex_17) is the unique medcode for a subject then autoimmune disease diagnosis is IBD, but if for the same subject medcode for IBD or Crohn/ UC is found simultaneously, the priority of autoimmune disease diagnosis date will be done to Crohn or UC disease.	Annex 17, Annex 18, Annex 19

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

	Column Name	Algorithms	Comments
42	**Autoimmune hepatitis**	AI hepatitis cases will be identified where: – in *Clinical and Referral file*: medcode= 18652 – in *HES (HES_diagnosis_epi file)*: ICD10 diagnosis code =K75.4 Eventdate should be between reference date and (reference date+365 days) for unexposed cohort and between reference date and (reference date+30months) for exposed cohort. Freetext related to autoimmune hepatitis in the study period will be retrieved and the case will be sent for expert review.	

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Section 3: Annexes

List of Annexes

[1] Neuroinflammatory/ophthalmic autoimmune disease

Annex 1: Medical codes and ICD-10 codes for Multiple Sclerosis
Annex 2: Medical codes and ICD-10 codes for Transverse Myelitis
Annex 3: Medical codes and ICD-10 codes for Optic Neuritis
Annex 4: Medical codes and ICD-10 codes for Guillain Barre Syndrome
Annex 5: Medical codes and ICD-10 codes for Other Demyelinating Diseases
Annex 6: Medical codes and ICD-10 codes for Autoimmune Uveitis

[2] Other autoimmune diseases

Annex 7: Medical codes and ICD-10 codes for Systemic Lupus Erythematosus
Annex 8: Medical codes and ICD-10 codes for Rheumatoid Arthritis
Annex 9: Medical codes and ICD-10 codes for Juvenile Rheumatoid Arthritis
Annex 10: Medical codes and ICD-10 codes for Still's disease
Annex 11: Medical codes and ICD-10 codes for Psoriatic Arthritis
Annex 12: Medical codes and ICD-10 codes for Ankylosing Spondylitis
Annex 13: Medical codes and ICD-10 codes for ITP
Annex 14: Medical codes and ICD-10 codes for AI haemolytic Anaemia
Annex 15: Medical codes and ICD-10 codes for Type 1 Diabetes Mellitus
Annex 16: Medical codes and ICD-10 codes for AI thyroiditis, including Hashimoto's disease, Graves'/Basedow's diseases
Annex 17: Medical codes and ICD-10 codes for Inflammatory Bowel Diseases
Annex 18: Medical codes and ICD-10 codes for Crohn's diseases
Annex 19: Medical codes and ICD-10 codes for Ulcerative Colitis
Annex 20: Cross tabulation of vaccines
Annex 21: Product codes related to autoimmune diseases
Annex 22: Procedure codes related to autoimmune diseases

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

A medical code reviewed as 'Yes' = direct etiological code

A medical code reviewed as 'Possible' = possible, need to be related to the auto-immune disease but need to be confirmed by other sources

[1] Neuroinflammatory/ophthalmic autoimmune disease

Annex 1: Medical codes and ICD-10 codes for Multiple Sclerosis

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
684	F20..00	Multiple sclerosis	G35	Y
3440	F20..11	Disseminated sclerosis	G35	Possible
40344	F200.00	Multiple sclerosis of the brain stem	G35	Y
69886	F201.00	Multiple sclerosis of the spinal cord	G35	Y
23730	F202.00	Generalised multiple sclerosis	G35	Y
2298	F203.00	Exacerbation of multiple sclerosis	-	Y
96291	F204.00	Benign multiple sclerosis	-	Y
96607	F206.00	Primary progressive multiple sclerosis	-	Y
95972	F207.00	Relapsing and remitting multiple sclerosis	-	Y
96246	F208.00	Secondary progressive multiple sclerosis	-	Y
20493	F20z.00	Multiple sclerosis NOS	G35	Y

CONFIDENTIAL

Annex 2: Medical codes and ICD-10 codes for Transverse Myelitis

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
25252	F03..00	Encephalitis, myelitis and encephalomyelitis	G04	Possible
5061	F03..11	Encephalomyelitis	G04.0	Possible
5060	F03..12	Myelitis	G04	Possible
5463	F03..13	Transverse myelitis	G37.3	Y
6494	F037.00	Transverse myelitis	G37.3	Y
61707	F03y.11	Encephalomyelitis NOS	G04.9	Possible
4364	F03y.12	Myalgic encephalomyelitis	-	Possible
2423	F03z.00	Encephalitis NOS	G04.9	Possible
95623	F03z.11	Encephalomyelitis NOS	G04.9	Possible
44795	F210.00	Neuromyelitis optica	G36.0	Possible
36312	F210.11	Devic's disease	G36.0	Y
62945	F211.11	Balo's concentric sclerosis	G37.5	Possible
97660	Fyu0.00	[X]Inflammatory diseases of the central nervous system	G00-G09	Possible
99730	Fyu0A00	[X]Encephalitis,myelitis+encephalomyelitis/other diseases CE	G05	Possible

CONFIDENTIAL

Annex 3: Medical codes and ICD-10 codes for Optic Neuritis

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
3771	F4H3.00	Optic neuritis	H46	Y
3339	F4H3000	Unspecified optic neuritis	H46.9	Possible
8366	F4H3100	Optic papillitis	H46.0	Y
1870	F4H3200	Acute retrobulbar neuritis	H46.1	Y
26835	F4H3z00	Optic neuritis NOS	H46.9	Possible

CONFIDENTIAL

Annex 4: Medical codes and ICD-10 codes for Guillain Barre Syndrome

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
28294	F326100	Polyneuritis cranialis	G52.7	Possible
44512	F364.00	Idiopathic progressive polyneuropathy	G60.3	Possible
14884	F36y.00	Other idiopathic peripheral neuropathy	G60	Possible
1607	F370000	Guillain-Barre syndrome	G61.0	Y
24216	F370100	Postinfectious polyneuritis	G61.0	Possible
33841	F370200	Miller-Fisher syndrome	G61.0	Y
63555	F374z00	Polyneuropathy in disease NOS	G63	Possible
31551	F37X.00	Inflammatory polyneuropathy, unspecified	G61.9	Possible
69047	F37y000	Serum neuropathy	G61.1	Possible
96256	F37y100	Axonal sensorimotor neuropathy	G60	Y
15481	F37z.00	Toxic or inflammatory neuropathy NOS	G61.9	Possible
24226	F37z.11	Polyneuropathy unspecified	G62.9	Possible
55076	Fyu7.00	[X]Polyneuropathies & other disord of peripheral nerv syst	G60/G64	Possible
97449	Fyu7000	[X]Other hereditary and idiopathic neuropathies	G60.8	Possible
97306	Fyu7200	[X]Other specified polyneuropathies	G62.8	Possible

CONFIDENTIAL

Annex 5: Medical codes and ICD-10 codes for Other Demyelinating Diseases

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 Codes	Review
31510	F034.00	Postimmunisation encephalitis	G04.0	Possible
59723	F034000	Post BCG vaccination encephalitis	G04.0	Possible
97865	F034100	Post typhoid vaccination encephalitis	G04.0	Possible
46467	F034200	Post paratyphoid vaccination encephalitis	G04.0	Possible
66883	F034500	Post tetanus vaccination encephalitis	G04.0	Possible
68395	F034700	Post pertussis vaccination encephalitis	G04.0	Possible
65548	F034800	Post smallpox vaccination encephalitis	G04.0	Possible
97847	F034A00	Post typhus vaccination encephalitis	G04.0	Possible
36707	F034B00	Post yellow fever vaccination encephalitis	G04.0	Possible
67795	F034C00	Post measles vaccination encephalitis	G04.0	Possible
39398	F034D00	Post polio vaccination encephalitis	G04.0	Possible
92431	F034E00	Post mumps vaccination encephalitis	G04.0	Possible
49070	F034G00	Post influenza vaccination encephalitis	G04.0	Possible
31544	F034H00	Post hepatitis A vaccination encephalitis	G04.0	Possible
39346	F034J00	Post hepatitis B vaccination encephalitis	G04.0	Possible
65184	F034x00	Post mixed vaccination encephalitis	G04.0	Possible
96428	F034z00	Postimmunisation encephalitis NOS	G04.0	Possible
53080	F03y.00	Other causes of encephalitis	-	Possible
41655	F0y..00	Inflammatory diseases of central nervous system OS	G00-G09	Possible
44819	F0z..00	Inflammatory diseases of central nervous system NOS	G00-G09	Possible
3464	F21..00	Other central nervous system demyelinating diseases	G37	Possible

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 Codes	Review
28082	F21X.00	Acute disseminated demyelination, unspecified	G36.9	Possible
43583	F21y.00	Other specified central nervous system demyelinating disease	G37.8	Possible
69848	F21y400	Sub-acute necrotizing myelitis	G37.4	Possible
54300	F21yz00	Other specified central nervous system demyelination NOS	G36	Possible
12054	F21z.00	Central nervous system demyelination NOS	G36	Possible
1040	F286.15	Myalgic encephalomyelitis	-	Possible
6552	F286.16	ME - Myalgic encephalomyelitis	-	Possible
16452	F35z.00	Mononeuritis of unspecified site NOS	-	Possible
29515	F36y000	Supranuclear paralysis	G60	Possible
53790	Fyu4.00	[X]Demyelinating diseases of the central nervous system	G35-G37	Possible
53919	Fyu4100	[X]Other specified demyelinating diseases/the CNS	G37.8	Possible

CONFIDENTIAL

Annex 6: Medical codes and ICD-10 codes for Autoimmune Uveitis

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review (Y/N/possible)
50067	C34y300	Gouty iritis	H22	Possible
29420	F400.00	Purulent endophthalmitis	H44.0	Possible
32145	F400100	Acute endophthalmitis	H44	Possible
48128	F400200	Panophthalmitis	H44.0	Possible
49756	F400300	Chronic endophthalmitis	-	Possible
61617	F400400	Vitreous abscess	H44.0	Possible
19986	F400z00	Purulent endophthalmitis NOS	H44.0	Possible
28431	F401.00	Other endophthalmitis	H44.1	Possible
31684	F401000	Sympathetic uveitis	H44.1	Possible
21685	F401100	Panuveitis	H44.1	Possible
68759	F401200	Parasitic endophthalmitis NOS	H44.1	Possible
72273	F401300	Ophthalmia nodosa	H16.2	Possible
18898	F401400	Phacoanaphylactic endophthalmitis	H44.19	Possible
35453	F401z00	Other endophthalmitis NOS	H44.1	Possible
6929	F43..00	Chorioretinal inflammations scars and other disorders	H30-H36	Possible
8278	F43..11	Choroid disorders	H30-H36	Possible
34170	F430.00	Focal chorioretinitis and retinochoroiditis	H30-H36	Possible
16629	F430.11	Retinitis and chorioretinitis	H30-H36	Possible
39964	F430000	Unspecified focal chorioretinitis	H30-H36	Possible
45908	F430100	Focal juxtapapillary choroiditis	H30-H36	Possible
94848	F430200	Other posterior pole focal chorioretinitis	H30-H36	Possible

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review (Y/N/possible)
55978	F430300	Peripheral focal chorioretinitis	H30-H36	Possible
96303	F430400	Focal juxtapapillary retinitis	H30-H36	Possible
27424	F430500	Focal macular retinochoroiditis	H30-H36	Possible
97455	F430600	Other posterior pole focal retinitis	H30-H36	Possible
55312	F430700	Peripheral focal retinochoroiditis	H30-H36	Possible
98959	F430800	Multifocal choroiditis	H30-H36	Possible
32239	F430z00	Focal chorioretinitis or retinochoroiditis NOS	H30-H36	Possible
43496	F431.00	Disseminated chorioretinitis and retinochoroiditis	H30-H36	Possible
66369	F431000	Unspecified disseminated chorioretinitis	H30-H36	Possible
65993	F431300	General disseminated chorioretinitis	H30-H36	Possible
99779	F431z00	Disseminated chorioretinitis and retinochoroiditis NOS	H30-H36	Possible
58055	F432.00	Other chorioretinitis and retinochoroiditis	H30-H36	Possible
4785	F432000	Choroiditis NOS	H30-H36	Possible
16264	F432100	Retinitis NOS	H30-H36	Possible
10999	F432200	Posterior uveitis NOS	H30-H36	Possible
19645	F432300	Posterior cyclitis	H30-H36	Possible
7882	F432311	Pars planitis	H30-H36	Possible
39882	F432400	Harada's disease	H30-H36	Possible
16194	F432z00	Other chorioretinitis or retinochoroiditis NOS	H30-H36	Possible
12249	F433.00	Chorioretinal scars	H30-H36	Possible
48117	F433000	Unspecified chorioretinal scar	H30-H36	Possible
10878	F433200	Other macular scars	H30-H36	Possible

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review (Y/N/possible)
97053	F433300	Other posterior pole scars	H30-H36	Possible
40527	F433400	Peripheral chorioretinal scars	H30-H36	Possible
10305	F433z00	Chorioretinal scars NOS	H30-H36	Possible
38997	F434300	Angioid streaks of choroid	H30-H36	Possible
17329	F44..00	Disorders of iris and ciliary body	H21	Possible
2703	F44..12	Iridocyclitis	H20	Possible
10579	F440.00	Acute and subacute iridocyclitis	H20%	Possible
478	F440.11	Iritis - acute	H20.0	Possible
43159	F440000	Unspecified acute iridocyclitis	H20.9	Possible
63210	F440100	Unspecified subacute iridocyclitis	H20.9	Possible
21685	F401100	Panuveitis	-	Possible
69723	F440200	Primary iridocyclitis	H20.19	Possible
21163	F440300	Recurrent iridocyclitis	H20.029	Possible
92888	F440400	Secondary infected iridocyclitis	H20.049	Possible
54389	F440500	Secondary noninfected iridocyclitis	H20.04	Possible
6849	F440z00	Acute or subacute iritis NOS	H20%	Possible
37140	F441.00	Chronic iridocyclitis	H20.1	Possible
17480	F441.11	Chronic iritis	H20.0	Possible
38716	F441000	Unspecified chronic iridocyclitis	H20.9	Possible
64052	F441100	Chronic iridocyclitis due to disease EC	H20.9	Possible
5556	F441200	Chronic anterior uveitis	H20.0	Possible
64804	F441z00	Chronic iridocyclitis NOS	H20.13	Possible

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review (Y/N/possible)
72481	F442.00	Certain types of iridocyclitis	H20.0	Possible
29764	F442100	Glaucomatocyclitic crises	H40.40x0	Possible
91596	F442z00	Certain types of cyclitis NOS	H20.0	Possible
14731	F443.00	Unspecified iridocyclitis	H20.9	Possible
477	F443.11	Uveitis NOS	H20.0	Possible
2245	F443000	Anterior uveitis	H20.0	Possible
98094	F443100	Iritis	H20.0	Possible
58719	F444.00	Iris and ciliary body vascular disorders	H21.1	Possible
3401	F444000	Hyphaema	H21.0	Possible
11131	F444100	Rubeosis iridis	H21.1x9	Possible
60634	F444z00	Iris and ciliary body vascular disorders NOS	H21.1	Possible
34986	F445.00	Iris and ciliary body degenerations	H21.2	Possible
62499	F445000	Progressive iris atrophy	H21.269	Possible
30413	F445100	Iridoschisis	H21.259	Possible
44798	F455000	Phacolytic glaucoma	H40.5	Possible
72913	FyuD100	[X]Other superficial keratitis without conjunctivitis	H16.1	Possible
72393	FyuD200	[X]Other keratitis	H16.8	Possible
97336	FyuD300	[X]Other central corneal opacity	H17.1	Possible
85982	FyuD400	[X]Other corneal scars and opacities	H17.8	Possible
73260	FyuD600	[X]Other corneal deformities	H18.7	Possible
72926	FyuD700	[X]Other specified disorders of cornea	H18.8	Possible
101293	FyuD800	[X]Scleritis+episcleritis in diseases CE	H19.0	Possible

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review (Y/N/possible)
101578	FyuDA00	[X]Keratitis+keratoconjunctivitis in other diseases CE	H16.2	Possible

[2] Other autoimmune diseases

Annex 7: Medical codes and ICD-10 codes for Systemic Lupus Erythematosus

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
44095	F371000	Polyneuropathy in disseminated lupus erythematosus	G63	Y
31564	H57y400	Lung disease with systemic lupus erythematosus	M32.1	Y
47672	K01x400	Nephrotic syndrome in systemic lupus erythematosus	M32.1	Y
22205	K01x411	Lupus nephritis	M32.1	Y
7871	N000.00	Systemic lupus erythematosus	M32	Y
20007	N000000	Disseminated lupus erythematosus	M32	Y
57675	N000100	Libman-Sacks disease	M32.1	Y
36942	N000200	Drug-induced systemic lupus erythematosus	M32.0	Possible
29519	N000300	Systemic lupus erythematosus with organ or sys involv	M32.1	Y
11920	N000400	Systemic lupus erythematosus with pericarditis	M32.1	Y
101433	N000600	Cerebral lupus	M32.1	Y
42719	N000z00	Systemic lupus erythematosus NOS	M32	Y
12177	N006.00	Antiphospholipid syndrome	D68.6	Possible
52860	Nyu4.00	[X]Systemic connective tissue disorders	M30-M36	Possible
62652	Nyu4200	[X]Other specified necrotizing vasculopathies	M31.8	Possible
58706	Nyu4300	[X]Other forms of systemic lupus erythematosus	M32.8	Y

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Annex 8: Medical codes and ICD-10 codes for Rheumatoid Arthritis

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
4502	43F1.00	Rheumatoid factor positive	-	Possible
62401	F371200	Polyneuropathy in rheumatoid arthritis	M05.3/G63.6	Y
47465	F371100	Polyneuropathy in polyarteritis nodosa	M30.0/G63.5	Possible
31209	F396400	Myopathy due to rheumatoid arthritis	M05.3/G73.7	Y
49787	G5y8.00	Rheumatoid myocarditis	M05.3	Possible
43816	G5yA.00	Rheumatoid carditis	M05.3	Possible
27603	N04..00	Rheumatoid arthritis and other inflammatory polyarthropathy	M06	Y
844	N040.00	Rheumatoid arthritis	M06.8	Y
44743	N040000	Rheumatoid arthritis of cervical spine	M06.8	Y
44203	N040100	Other rheumatoid arthritis of spine	M06.8	Y
21358	N040200	Rheumatoid arthritis of shoulder	M06.8	Y
100914	N040400	Rheumatoid arthritis of acromioclavicular joint	M06.8	Y
59738	N040500	Rheumatoid arthritis of elbow	M06.8	Y
63365	N040600	Rheumatoid arthritis of distal radio-ulnar joint	M06.8	Y
48832	N040700	Rheumatoid arthritis of wrist	M06.8	Y
42299	N040800	Rheumatoid arthritis of MCP joint	M06.8	Y
41941	N040900	Rheumatoid arthritis of PIP joint of finger	M06.8	Y
63198	N040A00	Rheumatoid arthritis of DIP joint of finger	M06.8	Y
49067	N040B00	Rheumatoid arthritis of hip	M06.8	Y
100776	N040C00	Rheumatoid arthritis of sacro-iliac joint	M06.8	Y

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
50863	N040D00	Rheumatoid arthritis of knee	M06.8	Y
51239	N040F00	Rheumatoid arthritis of ankle	M06.8	Y
73619	N040G00	Rheumatoid arthritis of subtalar joint	M06.8	Y
70658	N040H00	Rheumatoid arthritis of talonavicular joint	M06.8	Y
71784	N040J00	Rheumatoid arthritis of other tarsal joint	M06.8	Y
51238	N040K00	Rheumatoid arthritis of 1st MTP joint	M06.8	Y
99414	N040L00	Rheumatoid arthritis of lesser MTP joint	M06.8	Y
20615	N04..11	Inflammatory polyarthropathy	M06.4	Possible
6916	N040P00	Seronegative rheumatoid arthritis	M06.0	Y
30548	N040N00	Rheumatoid vasculitis	M05.2	Possible
18155	N040Q00	Rheumatoid bursitis	M06.2	Possible
31054	N040S00	Rheumatoid arthritis - multiple joint	M06.8	Y
8350	N040T00	Flare of rheumatoid arthritis	-	Y
23552	N041.00	Felty's syndrome	M05.0	Y
53621	N040R00	Rheumatoid nodule	M06.3	Possible
49227	N042.00	Other rheumatoid arthropathy + visceral/systemic involvement	M05.8	Possible
8583	N042000	Rheumatic carditis	M05.3	Possible
46436	N042100	Rheumatoid lung disease	M05.1	Possible
5723	N042200	Rheumatoid nodule	M06.3	Possible
37431	N042z00	Rheumatoid arthropathy + visceral/systemic involvement NOS	M05.6	Possible
3944	N044.00	Chronic post-rheumatic arthropathy	M12.0	Possible
40841	N044.11	Jaccoud's syndrome	M12.0	Possible

CONFIDENTIAL

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
58543	N044.12	Nodular fibrositis of chronic rheumatic disease	M12	Possible
9707	N047.00	Seropositive errosive rheumatoid arthritis	M05.8	Y
12019	N04X.00	Seropositive rheumatoid arthritis, unspecified	M05.9	Y
36597	N04y.00	Other specified inflammatory polyarthropathy	M06	Possible
31724	N04y000	Rheumatoid lung	M05.1	Possible
56838	N04y011	Caplan's syndrome	M05.1	Possible
28853	N04y012	Fibrosing alveolitis associated with rheumatoid arthritis	M06	Possible
4578	N04y100	Sero negative arthritis	M06.0	Possible
10919	N04y111	Sero negative polyarthritis	M06	Possible
23833	N04yz00	Other specified inflammatory polyarthropathy NOS	M06	Possible
24747	N04z.00	Inflammatory polyarthropathy NOS	M06	Possible

CONFIDENTIAL

Annex 9: Medical codes and ICD-10 codes for Juvenile Rheumatoid Arthritis

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
4186	N043.00	Juvenile rheumatoid arthritis - Still's disease	M08.2	Y
47831	N043100	Acute polyarticular juvenile rheumatoid arthritis	M08	Y
21533	N043200	Pauciarticular juvenile rheumatoid arthritis	M08.4	Y
36276	N043300	Monarticular juvenile rheumatoid arthritis	M08	Y
27557	N043z00	Juvenile rheumatoid arthritis NOS	M08	Y
50644	N043000	Juvenile rheumatoid arthropathy unspecified	M08.9	Possible
31181	N045100	Juvenile seronegative polyarthritis	M08.3	Y
7196	N045.00	Other juvenile arthritis	M08.8	Possible
31360	N045500	Juvenile rheumatoid arthritis	M08.0	Y
46622	N045600	Pauciarticular onset juvenile chronic arthritis	M08.4	Y
28456	N045200	Juvenile arthritis in psoriasis	M09.0	Possible

Annex 10: Medical codes and ICD-10 codes for Still's disease

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
32001	N04y200	Adult-onset Still's disease	M06.1	Y
23834	N005.00	Adult Still's Disease	M08.2	Possible

CONFIDENTIAL

Annex 11: Medical codes and ICD-10 codes for Psoriatic Arthritis

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
21503	M160200	Arthritis mutilans	L40.5	Possible
96880	M160.11	Psoriatic arthritis	L40.5	Y
26368	M160000	Psoriasis spondylitica	L40.5	Y
32149	M160100	Distal interphalangeal psoriatic arthropathy	L40.5	Y
12500	M160z00	Psoriatic arthropathy NOS	L40.5	Y

Annex 12: Medical codes and ICD-10 codes for Ankylosing Spondylitis

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review (Y/N/possible)
34880	N10..00	Inflammatory spondylopathies	M45-M49	Possible
2184	N100.00	Ankylosing spondylitis	M45	Y
40946	N100.11	Marie - Strumpell spondylitis	M45	Possible
44026	N10y.00	Other inflammatory spondylopathies	M46	Possible
58683	N10y000	Inflammatory spondylopathies in diseases EC	M46	Possible
37892	N10yz00	Other inflammatory spondylopathies NOS	M46	Possible
1963	N10z.00	Spondylitis NOS	M46	Possible
42405	N045000	Juvenile ankylosing spondylitis	M08.1	Y

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Annex 13: Medical codes and ICD-10 codes for ITP

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
16420	42P2.11	Auto-immune thrombocytopenia	D69.3	Y
46754	D313.11	Evan's syndrome	D69.3	Y
5144	D313.12	Idiopathic thrombocytopenic purpura	D69.3	Y
71257	D313.13	Idiopathic purpura	D69	Possible
65723	D313.14	Megakaryocytic hypoplasia	D69	Possible
877	D313.15	Thrombocytopenic purpura	D69.3	Possible
12234	D313000	Idiopathic thrombocytopenic purpura	D69	Y
21604	D313011	Idiopathic purpura	D69	Possible
5181	D313012	ITP - idiopathic thrombocytopenic purpura	D69	Y
54005	D313y00	Other specified primary thrombocytopenia	D69	Possible
57456	D313z00	Primary thrombocytopenia NOS	D69.4	Possible
45698	D313z11	Essential thrombocytopenia NOS	D69	Possible
62795	Dyu3200	[X]Other primary thrombocytopenia	D69.4	Possible

Annex 14: Medical codes and ICD-10 codes for AI haemolytic Anaemia

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
3818	D110.00	Autoimmune haemolytic anaemias	D59.1	Y
21369	D110.11	Coombs positive haemolysis	-	Possible
39944	D110000	Primary cold-type haemolytic anaemia	D59.1	Possible
49182	D110100	Primary warm-type haemolytic anaemia	D59.1	Possible
57575	D110200	Secondary cold-type haemolytic anaemia	D59.1	Possible
31734	D110400	Drug-induced autoimmune haemolytic anaemia	D59.0	Possible
39876	D110z00	Autoimmune haemolytic anaemia NOS	D59.1	Y
48096	D112z11	Cold haemoglobinuria	D59.1	Possible
100388	Dyu1500	[X]Other autoimmune haemolytic anaemias	D59.1	Y

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Annex 15: Medical codes and ICD-10 codes for Type 1 Diabetes Mellitus

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
711	C10..00	Diabetes mellitus	E10	Possible
24490	C100000	Diabetes mellitus, juvenile type, no mention of complication	E10	Y
1038	C100011	Insulin dependent diabetes mellitus	E10	Y
1682	C101.00	Diabetes mellitus with ketoacidosis	E10	Possible
53200	C101000	Diabetes mellitus, juvenile type, with ketoacidosis	E10	Y
54856	C101100	Diabetes mellitus, adult onset, with ketoacidosis	E10	Possible
38617	C101y00	Other specified diabetes mellitus with ketoacidosis	E13	Possible
42505	C101z00	Diabetes mellitus NOS with ketoacidosis	E10	Possible
21482	C102.00	Diabetes mellitus with hyperosmolar coma	E10	Possible
40023	C102000	Diabetes mellitus, juvenile type, with hyperosmolar coma	E10	Y
43139	C102100	Diabetes mellitus, adult onset, with hyperosmolar coma	E10	Possible
72345	C102z00	Diabetes mellitus NOS with hyperosmolar coma	E10	Possible
15690	C103.00	Diabetes mellitus with ketoacidotic coma	E10	Possible
42567	C103000	Diabetes mellitus, juvenile type, with ketoacidotic coma	E10	Y
68843	C103100	Diabetes mellitus, adult onset, with ketoacidotic coma	E10	Possible
59288	C103y00	Other specified diabetes mellitus with coma	E13	Possible
65062	C103z00	Diabetes mellitus NOS with ketoacidotic coma	E10	Possible
16502	C104.00	Diabetes mellitus with renal manifestation	E10	Possible
2475	C104.11	Diabetic nephropathy	E08	Possible
93922	C104000	Diabetes mellitus, juvenile type, with renal manifestation	E10	Y

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
35105	C104100	Diabetes mellitus, adult onset, with renal manifestation	E10	Possible
13279	C104y00	Other specified diabetes mellitus with renal complications	E13	Possible
35107	C104z00	Diabetes mellitus with nephropathy NOS	E10	Possible
33254	C105.00	Diabetes mellitus with ophthalmic manifestation	E10	Possible
69748	C105000	Diabetes mellitus, juvenile type, + ophthalmic manifestation	E10	Y
41389	C105100	Diabetes mellitus, adult onset, + ophthalmic manifestation	E10	Possible
47377	C105y00	Other specified diabetes mellitus with ophthalmic complicatn	E13	Possible
34283	C105z00	Diabetes mellitus NOS with ophthalmic manifestation	E10	Possible
16230	C106.00	Diabetes mellitus with neurological manifestation	E10	Possible
59903	C106.11	Diabetic amyotrophy	E10	Possible
7795	C106.12	Diabetes mellitus with neuropathy	E10	Possible
16491	C106.13	Diabetes mellitus with polyneuropathy	E10	Possible
67853	C106000	Diabetes mellitus, juvenile, + neurological manifestation	E10	Y
39317	C106100	Diabetes mellitus, adult onset, + neurological manifestation	E10	Possible
61523	C106y00	Other specified diabetes mellitus with neurological comps	E13	Possible
22573	C106z00	Diabetes mellitus NOS with neurological manifestation	E10	Possible
35399	C107.00	Diabetes mellitus with peripheral circulatory disorder	E10	Possible
32403	C107.11	Diabetes mellitus with gangrene	E10	Possible
32556	C107.12	Diabetes with gangrene		Possible
70448	C107000	Diabetes mellitus, juvenile +peripheral circulatory disorder	E10	Y
63357	C107100	Diabetes mellitus, adult, + peripheral circulatory disorder	E10	Possible

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
33807	C107200	Diabetes mellitus, adult with gangrene	E10	Possible
69124	C107300	IDDM with peripheral circulatory disorder	E10	Y
65025	C107z00	Diabetes mellitus NOS with peripheral circulatory disorder	E10	Possible
1647	C108.00	Insulin dependent diabetes mellitus	E10	Y
18505	C108.11	IDDM-Insulin dependent diabetes mellitus	E10	Y
17858	C108.12	Type 1 diabetes mellitus	E10	Y
24423	C108.13	Type I diabetes mellitus	E10	Y
46963	C108000	Insulin-dependent diabetes mellitus with renal complications	E10	Y
61344	C108011	Type I diabetes mellitus with renal complications	E10.2	Y
21983	C108012	Type 1 diabetes mellitus with renal complications	E10.2	Y
49276	C108100	Insulin-dependent diabetes mellitus with ophthalmic comps	E10	Y
52283	C108200	Insulin-dependent diabetes mellitus with neurological comps	E10	Y
49146	C108211	Type I diabetes mellitus with neurological complications	E10	Y
61829	C108212	Type 1 diabetes mellitus with neurological complications	E10	Y
52104	C108300	Insulin dependent diabetes mellitus with multiple complicatn	E10.7	Y
26855	C108400	Unstable insulin dependent diabetes mellitus	E10	Y
60107	C108411	Unstable type I diabetes mellitus	E10	Y
97474	C108412	Unstable type 1 diabetes mellitus	E10	Y
44443	C108500	Insulin dependent diabetes mellitus with ulcer	E10	Y
51957	C108511	Type I diabetes mellitus with ulcer	E10	Y
68390	C108512	Type 1 diabetes mellitus with ulcer	E10	Y

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
60499	C108600	Insulin dependent diabetes mellitus with gangrene	E10	Y
6509	C108700	Insulin dependent diabetes mellitus with retinopathy	E10	Y
38161	C108711	Type I diabetes mellitus with retinopathy	E10	Y
41049	C108712	Type 1 diabetes mellitus with retinopathy	E10	Y
6791	C108800	Insulin dependent diabetes mellitus - poor control	E10	Y
46850	C108811	Type I diabetes mellitus - poor control	E10	Y
45914	C108812	Type 1 diabetes mellitus - poor control	E10	Y
31310	C108900	Insulin dependent diabetes maturity onset	E10	Y
63017	C108911	Type I diabetes mellitus maturity onset	E10	Y
97446	C108912	Type 1 diabetes mellitus maturity onset	E10	Y
56448	C108A00	Insulin-dependent diabetes without complication	E10	Y
95992	C108A11	Type I diabetes mellitus without complication	E10	Y
24694	C108B00	Insulin dependent diabetes mellitus with mononeuropathy	E10	Y
99231	C108B11	Type I diabetes mellitus with mononeuropathy	E10	Y
41716	C108C00	Insulin dependent diabetes mellitus with polyneuropathy	E10	Y
57621	C108D00	Insulin dependent diabetes mellitus with nephropathy	E10	Y
66872	C108D11	Type I diabetes mellitus with nephropathy	E10	Y
44440	C108E00	Insulin dependent diabetes mellitus with hypoglycaemic coma	E10	Y
42729	C108E11	Type I diabetes mellitus with hypoglycaemic coma	E10	Y
70766	C108E12	Type 1 diabetes mellitus with hypoglycaemic coma	E10	Y
44260	C108F00	Insulin dependent diabetes mellitus with diabetic cataract	E10	Y

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
17545	C108F11	Type I diabetes mellitus with diabetic cataract	E10	Y
64446	C108G00	Insulin dependent diab mell with peripheral angiopathy	E10	Y
65616	C108H00	Insulin dependent diabetes mellitus with arthropathy	E10	Y
62352	C108H11	Type I diabetes mellitus with arthropathy	E10	Y
39809	C108J00	Insulin dependent diab mell with neuropathic arthropathy	E10	Y
60208	C108J11	Type I diabetes mellitus with neuropathic arthropathy	E10	Y
18230	C108J12	Type 1 diabetes mellitus with neuropathic arthropathy	E10	Y
46290	C108y00	Other specified diabetes mellitus with multiple comps	E13	Possible
64449	C108z00	Unspecified diabetes mellitus with multiple complications	E14	Possible
52236	C10A.00	Malnutrition-related diabetes mellitus	E12	Possible
66675	C10A000	Malnutrition-related diabetes mellitus with coma	E12	Possible
33969	C10A100	Malnutrition-related diabetes mellitus with ketoacidosis	E12	Possible
100347	C10A500	Malnutritn-relat diabetes melitus wth periph circul complctn	E12	Possible
43453	C10C.00	Diabetes mellitus autosomal dominant		Possible
46624	C10C.11	Maturity onset diabetes in youth	-	Possible
98392	C10C.12	Maturity onset diabetes in youth type 1	-	Possible
1549	C10E.00	Type 1 diabetes mellitus	E10	Y
12455	C10E.11	Type I diabetes mellitus	E10	Y
51261	C10E.12	Insulin dependent diabetes mellitus	E10	Y
47582	C10E000	Type 1 diabetes mellitus with renal complications	E10	Y
47649	C10E100	Type 1 diabetes mellitus with ophthalmic complications	E10	Y

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
99311	C10E111	Type I diabetes mellitus with ophthalmic complications	E10	Y
98071	C10E112	Insulin-dependent diabetes mellitus with ophthalmic comps	E10	Y
42831	C10E200	Type 1 diabetes mellitus with neurological complications	E10	Y
101735	C10E212	Insulin-dependent diabetes mellitus with neurological comps	E10	Y
47650	C10E300	Type 1 diabetes mellitus with multiple complications	E10	Y
91942	C10E311	Type I diabetes mellitus with multiple complications	E10	Y
45276	C10E312	Insulin dependent diabetes mellitus with multiple complicat	E10	Y
43921	C10E400	Unstable type 1 diabetes mellitus	E10	Y
49949	C10E411	Unstable type I diabetes mellitus	E10	Y
54600	C10E412	Unstable insulin dependent diabetes mellitus	E10	Y
18683	C10E500	Type 1 diabetes mellitus with ulcer	E10	Y
93878	C10E511	Type I diabetes mellitus with ulcer	E10	Y
98704	C10E512	Insulin dependent diabetes mellitus with ulcer	E10	Y
69993	C10E600	Type 1 diabetes mellitus with gangrene	E10	Y
102112	C10E611	Type I diabetes mellitus with gangrene	E10	Y
18387	C10E700	Type 1 diabetes mellitus with retinopathy	E10	Y
95343	C10E711	Type I diabetes mellitus with retinopathy	E10	Y
93875	C10E712	Insulin dependent diabetes mellitus with retinopathy	E10	Y
35288	C10E800	Type 1 diabetes mellitus - poor control	E10	Y
72702	C10E812	Insulin dependent diabetes mellitus - poor control	E10	Y
40682	C10E900	Type 1 diabetes mellitus maturity onset	E10	Y

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
96235	C10E911	Type I diabetes mellitus maturity onset	E10	Y
97849	C10E912	Insulin dependent diabetes maturity onset	E10	Y
69676	C10EA00	Type 1 diabetes mellitus without complication	E10	Y
62613	C10EA11	Type I diabetes mellitus without complication	E10	Y
99719	C10EA12	Insulin-dependent diabetes without complication	E10	Y
68105	C10EB00	Type 1 diabetes mellitus with mononeuropathy	E10	Y
46301	C10EC00	Type 1 diabetes mellitus with polyneuropathy	E10	Y
91943	C10EC11	Type I diabetes mellitus with polyneuropathy	E10	Y
101311	C10EC12	Insulin dependent diabetes mellitus with polyneuropathy	E10	Y
10418	C10ED00	Type 1 diabetes mellitus with nephropathy	E10	Y
102163	C10ED12	Insulin dependent diabetes mellitus with nephropathy	E10	Y
39070	C10EE00	Type 1 diabetes mellitus with hypoglycaemic coma	E10	Y
99716	C10EE12	Insulin dependent diabetes mellitus with hypoglycaemic coma	E10	Y
49554	C10EF00	Type 1 diabetes mellitus with diabetic cataract	E10	Y
100770	C10EF12	Insulin dependent diabetes mellitus with diabetic cataract	E10	Y
93468	C10EG00	Type 1 diabetes mellitus with peripheral angiopathy	E10	Y
18642	C10EH00	Type 1 diabetes mellitus with arthropathy	E10	Y
54008	C10EJ00	Type 1 diabetes mellitus with neuropathic arthropathy	E10	Y
30323	C10EK00	Type 1 diabetes mellitus with persistent proteinuria	E10	Y
30294	C10EL00	Type 1 diabetes mellitus with persistent microalbuminuria	E10	Y
102620	C10EL11	Type I diabetes mellitus with persistent microalbuminuria	E10	Y

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
10692	C10EM00	Type 1 diabetes mellitus with ketoacidosis	E10	Y
62209	C10EM11	Type I diabetes mellitus with ketoacidosis	E10	Y
40837	C10EN00	Type 1 diabetes mellitus with ketoacidotic coma	E10	Y
66145	C10EN11	Type I diabetes mellitus with ketoacidotic coma	E10	Y
22871	C10EP00	Type 1 diabetes mellitus with exudative maculopathy	E10	Y
97894	C10EP11	Type I diabetes mellitus with exudative maculopathy	E10	Y
55239	C10EQ00	Type 1 diabetes mellitus with gastroparesis	E10	Y
95636	C10ER00	Latent autoimmune diabetes mellitus in adult		Y
43857	C10M.00	Lipoatrophic diabetes mellitus	-	Possible
22487	C10N.00	Secondary diabetes mellitus	E13	Possible
33343	C10y.00	Diabetes mellitus with other specified manifestation	E10	Possible
63371	C10y100	Diabetes mellitus, adult, + other specified manifestation	E10	Possible
10098	C10yy00	Other specified diabetes mellitus with other spec comps	E13	Possible
70821	C10yz00	Diabetes mellitus NOS with other specified manifestation	E10	Possible
45491	C10z.00	Diabetes mellitus with unspecified complication	E10	Possible
68792	C10z000	Diabetes mellitus, juvenile type, + unspecified complication	E10	Y
63762	C10z100	Diabetes mellitus, adult onset, + unspecified complication	E10	Possible
64283	C10zy00	Other specified diabetes mellitus with unspecified comps	E13	Possible
64357	C10zz00	Diabetes mellitus NOS with unspecified complication	E10	Possible
17067	F171100	Autonomic neuropathy due to diabetes	E10.4/G99.0	Possible
17247	F35z000	Diabetic mononeuritis NOS	G57	Possible

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
31790	F372.00	Polyneuropathy in diabetes	G61	Possible
5002	F372.11	Diabetic polyneuropathy	G62	Possible
2342	F372.12	Diabetic neuropathy	G63	Possible
48078	F372000	Acute painful diabetic neuropathy	G64	Possible
35785	F372100	Chronic painful diabetic neuropathy	G65	Possible
24571	F372200	Asymptomatic diabetic neuropathy	G66	Possible
39420	F381300	Myasthenic syndrome due to diabetic amyotrophy	G70	Possible
2340	F381311	Diabetic amyotrophy	G70	Possible
37315	F3y0.00	Diabetic mononeuropathy	-	Possible
1323	F420.00	Diabetic retinopathy	-	Possible
7069	F420000	Background diabetic retinopathy	-	Possible
3286	F420100	Proliferative diabetic retinopathy	-	Possible
2986	F420200	Preproliferative diabetic retinopathy	-	Possible
10099	F420300	Advanced diabetic maculopathy	-	Possible
3837	F420400	Diabetic maculopathy	-	Possible
47584	F420500	Advanced diabetic retinal disease	-	Possible
10755	F420600	Non proliferative diabetic retinopathy	-	Possible
30477	F420700	High risk proliferative diabetic retinopathy	-	Possible
65463	F420800	High risk non proliferative diabetic retinopathy	-	Possible
11626	F420z00	Diabetic retinopathy NOS	-	Possible
17313	F440700	Diabetic iritis	-	Possible

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
2471	K01x100	Nephrotic syndrome in diabetes mellitus	-	Possible

Annex 16: Medical codes and ICD-10 codes for AI thyroiditis, including Hashimoto's disease, Graves'/Basedow's diseases

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
1882	C0...00	Disorders of thyroid gland	E06	Possible
45587	C0...11	Struma - goitre	E04	Possible
677	C02..00	Thyrotoxicosis	E05	Possible
1472	C02..11	Hyperthyroidism	E05	Possible
10760	C02..12	Toxic goitre	E05	Possible
23315	C020.00	Toxic diffuse goitre	E05	Possible
44405	C020.11	Basedow's disease	E05	Y
5257	C020.12	Graves' disease	E05	Y
26702	C020000	Toxic diffuse goitre with no crisis	E05	Possible
57011	C020100	Toxic diffuse goitre with crisis	E05	Possible
100476	C020200	Thyroid-associated dermopathy	E05	Possible
49334	C020z00	Toxic diffuse goitre NOS	E05	Possible
53280	C021.00	Toxic uninodular goitre	E05	Possible
26869	C021000	Toxic uninodular goitre with no crisis	E05	Possible
61498	C021z00	Toxic uninodular goitre NOS	E05	Possible
11426	C022.00	Toxic multinodular goitre	E05	Possible
46985	C022000	Toxic multinodular goitre with no crisis	E05	Possible
53981	C022z00	Toxic multinodular goitre NOS	E05	Possible
15790	C023.00	Toxic nodular goitre unspecified	E05	Possible
68512	C023000	Toxic nodular goitre unspecified with no crisis	E05	Possible

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
100004	C023100	Toxic nodular goitre unspecified with crisis	E05	Possible
49361	C023z00	Toxic nodular goitre NOS	E05	Possible
49508	C024.00	Thyrotoxicosis from ectopic thyroid nodule	E05	Possible
64656	C024000	Thyrotoxicosis from ectopic thyroid nodule with no crisis	E0	Possible
56270	C024z00	Thyrotoxicosis from ectopic thyroid nodule NOS	E05	Possible
43136	C02y.00	Thyrotoxicosis of other specified origin	E05	Possible
58138	C02y.11	Factitia thyrotoxicosis	E05	Possible
51273	C02y000	Thyrotoxicosis of other specified origin with no crisis	E05	Possible
64856	C02y200	Thyrotoxicosis factitia	E05	Possible
19205	C02y300	Thyroid crisis	E05	Possible
34220	C02yz00	Thyrotoxicosis of other specified origin NOS	E05	Possible
15565	C02z.00	Thyrotoxicosis without mention of goitre or other cause	E05	Possible
26701	C02z000	Thyrotoxicosis without mention of goitre or cause no crisis	E05	Possible
3194	C02z100	Thyrotoxicosis without mention of goitre, cause with crisis	E05	Possible
26699	C02zz00	Thyrotoxicosis NOS	E05	Possible
3290	C04..00	Acquired hypothyroidism	E03	Possible
1619	C04..11	Myxoedema	E03	Possible
14704	C04..12	Thyroid deficiency	E03	Possible
273	C04..13	Hypothyroidism	E03	Possible
31971	C046.00	Autoimmune myxoedema	E03	Possible
95830	C047.00	Subclinical hypothyroidism	E02	Possible
24748	C04y.00	Other acquired hypothyroidism	E03	Possible

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
3941	C04z.00	Hypothyroidism NOS	E03	Possible
20310	C04z.11	Pretibial myxoedema - hypothyroid	E03	Possible
23014	C04z.12	Thyroid insufficiency	E03	Possible
18282	C04z.13	Hypothyroid goitre, acquired	E03	Possible
56722	C04z000	Premature puberty due to hypothyroidism	-	Possible
59702	C04z100	Myxoedema coma	E03	Possible
1346	C05..00	Thyroiditis	E06	Possible
4898	C050.00	Acute thyroiditis	E06	Possible
67972	C050000	Acute nonsuppurative thyroiditis	E06	Possible
70773	C050100	Acute suppurative thyroiditis	E06	Possible
29296	C050200	Abscess of thyroid	E06	Possible
42323	C050z00	Acute thyroiditis NOS	E06	Possible
30799	C051.00	Subacute thyroiditis	E06	Possible
21747	C051.11	De Quervain's thyroiditis	E06.1	Y
26833	C052.00	Chronic lymphocytic thyroiditis	E06	Possible
3857	C052.11	Autoimmune thyroiditis	E06.3	Y
3436	C052.12	Hashimoto's disease	E06.3	Y
70244	C053.00	Chronic fibrous thyroiditis	E06.5	Possible
53667	C053.11	Riedel's thyroiditis	E06.5	Y
65444	C05y.00	Other and unspecified chronic thyroiditis	E06	Possible
65907	C05y400	Chronic thyroiditis with transient thyrotoxicosis	E06.2	Possible
20909	C05z.00	Thyroiditis NOS	E06.9	Possible

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
52843	C182.00	Autoimmune polyglandular failure	E31.0	Possible
47658	F11x500	Cerebral degeneration due to myxoedema	-	Possible
61069	F381400	Myasthenic syndrome due to hypothyroidism	E05	Possible
47695	F381600	Myasthenic syndrome due to thyrotoxicosis	E05	Possible
51416	F395300	Myopathy due to myxoedema	E03	Possible
48167	F395400	Myopathy due to thyrotoxicosis	E05	Possible
68626	FyuBD00	[X]Dysthyroid exophthalmos	H06.2/E05	Possible

Annex 17: Medical codes and ICD-10 codes for Inflammatory Bowel Diseases

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
5473	J4...00	Noninfective enteritis and colitis	K51	Possible
1561	J4...11	Colitis - noninfective	K51	Possible
1796	J4...12	Inflammatory bowel disease	K50-K52	Y
6016	J4...13	Noninfective diarrhoea	K52.9	Possible
42477	J4z..00	Non-infective gastroenteritis NOS	K52.9	Possible
30321	J4z..11	Presumed noninfectious diarrhoea	K52.9	Possible
6420	J4z0.00	Non-infective gastritis NOS	K52.9	Possible
59908	J4z1.00	Non-infective jejunitis NOS	K52.9	Possible
9788	J4z2.00	Non-infective ileitis NOS	K52.9	Possible
8301	J4z3.00	Non-infective colitis NOS	K52.9	Possible
15674	J4z4.00	Non-infective sigmoiditis NOS	K52.9	Possible
30662	J4z5.00	Exacerbation of non-infective colitis	K52.9	Possible
96976	J4z6.00	Indeterminate colitis	K52.3	Possible
12377	J4zz.00	Non-infective gastroenteritis NOS	K52.9	Possible
17017	J4zz.11	Diarrhoea - presumed non-infectious	K52.9	Possible
63580	Jyu4.00	[X]Noninfective enteritis and colitis	K52.9	Possible

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Annex 18: Medical codes and ICD-10 codes for Crohn's diseases

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
11286	J40..00	Regional enteritis - Crohn's disease	K50	Y
593	J40..11	Crohn's disease	K50	Y
51578	J40..12	Granulomatous enteritis	K50	Possible
51576	J400.00	Regional enteritis of the small bowel	K50.1	Possible
71945	J400000	Regional enteritis of the duodenum	K50.1	Possible
63036	J400100	Regional enteritis of the jejunum	K50	Possible
28476	J400200	Crohn's disease of the terminal ileum	K50	Y
66238	J400300	Crohn's disease of the ileum unspecified	K50	Y
39278	J400400	Crohn's disease of the ileum NOS	K50	Y
36913	J400500	Exacerbation of Crohn's disease of small intestine	K50	Y
9359	J400z00	Crohn's disease of the small bowel NOS	K50	Y
44426	J401.00	Regional enteritis of the large bowel	K50.1	Possible
62628	J401000	Regional enteritis of the colon	K50.1	Possible
64773	J401100	Regional enteritis of the rectum	K50.1	Possible
39037	J401200	Exacerbation of Crohn's disease of large intestine	K50	Y
20688	J401z00	Crohn's disease of the large bowel NOS	K50	Y
6538	J401z11	Crohn's colitis	K50	Y
15773	J402.00	Regional ileocolitis	K50.1	Possible
52449	J40z.00	Regional enteritis NOS	K50.1	Possible
59994	J40z.11	Crohn's disease NOS	K50	Y

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
69959	Jyu4000	[X]Other Crohn's disease	K50	Y
20480	N031100	Arthropathy in Crohn's disease	K50	Y
12575	N045300	Juvenile arthritis in Crohn's disease	K50	Y

CONFIDENTIAL

Annex 19: Medical codes and ICD-10 codes for Ulcerative Colitis

Medcode (CPRD-GOLD Medical Code Events)	Read Code	Read Description	ICD-10 codes	Review
5133	J41..00	Idiopathic proctocolitis	K51	Possible
23950	J41..11	Mucous colitis and/or proctitis	K59.9	Possible
1784	J41..12	Ulcerative colitis and/or proctitis	K51	Y
6650	J410.00	Ulcerative proctocolitis	K51	Y
48732	J410000	Ulcerative ileocolitis	K51	Y
704	J410100	Ulcerative colitis	K51	Y
24858	J410200	Ulcerative rectosigmoiditis	K51	Y
8347	J410300	Ulcerative proctitis	K51.2	Y
22516	J410400	Exacerbation of ulcerative colitis	K51	Y
33456	J410z00	Ulcerative proctocolitis NOS	K51	Y
30433	J411.00	Ulcerative (chronic) enterocolitis	K51	Y
42822	J412.00	Ulcerative (chronic) ileocolitis	K51	Y
24550	J41y.00	Other idiopathic proctocolitis	K51	Possible
16463	J41y000	Pseudopolyposis of colon	K51.4	Possible
26422	J41y100	Toxic megacolon	K59.3	Possible
43090	J41yz00	Other idiopathic proctocolitis NOS	K51	Possible
15207	J41z.00	Idiopathic proctocolitis NOS	K51	Possible
53743	Jyu4100	[X]Other ulcerative colitis	K51	Y
17641	N031000	Arthropathy in ulcerative colitis	K51	Y
71083	N045400	Juvenile arthritis in ulcerative colitis	K51	Y

Annex 20: Cross tabulation of vaccines

Available upon request

Annex 21: Product codes related to autoimmune diseases

Available upon request

Annex 22: Procedure codes related to autoimmune diseases

Available upon request

Annex 6 Example of table and figure templates

Annex 6 Version 2.0 (11 FEB 2013)

Example of table and figure templates

eTrack study number and Abbreviated Title	116239 (EPI-HPV-040 VS UK)
Title	An observational cohort study to assess the risk of autoimmune diseases in adolescent and young adult women aged 9 to 25 years exposed to Cervarix® in the United Kingdom

CONFIDENTIAL

TABLE OF CONTENTS

PAGE

LIST OF TABLES

PAGE

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

1. DEMOGRAPHICS AND BASELINE CHARACTERISTICS

Table 1 Subjects disposition

	Exposed F cohort		Unexposed F		Concurrent M		Historical M	
	n	%	n	%	n	%	n	%
Number of screened subjects Female subjects vaccinated with Cervarix								
Number of non-included subjects (1)								
First dose of Cervarix vaccination outside 01Sep2008 and 31Dec2010								
Subjects not actively registered with the practice during the study period								
Subjects not aged 9 to 25 years at the reference date								
Recorded in the CPRD for less than 12 months at reference date								
Subjects defined as non- acceptable in CPRD								
Diagnostic code of NOAD during the year prior to the reference								
Subjects who received at least one dose of unspecified HPV vaccine or Gardasil at any time before the reference date								
Any dose of Cervarix before the reference date								
Number of eligible subjects (2)								
Number of included subjects (3)								

1. Percent = 100 * n of non-included subjects / total number of screened subjects
2. Percent = 100 * n of eligible subjects / total number of screened subjects
3. Percent = 100 * n of included subjects / total number of eligible subjects

CONFIDENTIAL

Table 2 Demographic characteristics (all included subjects)

		Exposed F N= XXX		Unexposed F N= XXX		Concurrent M N= XXX		Historical M N= XXX		Overall N=xxx		p-value
Characteristics	**Parameters or Categories**	**Value or n**	**%**	**Value or n**	**%**	**Value or n**	**%**	**Value or n**	**%**	**Value or n**	**%**	
Age [years] at reference date	Mean											
	SD											
	Median											
	Minimum											
	Maximum											
	Missing											
Age Group [years]	9											
	10											
	etc.											
	25											
Region	<region 1>											
	<region 2>											
	<region 3>											
	Etc											
	Missing											

CONFIDENTIAL

Table 3 History of auto-immune disease (all included subjects)

		Exposed F N= XXX		Unexposed F N= XXX		Concurrent M N= XXX		Historical M N= XXX		Overall N=xxx		p-value
Characteristics	**Parameters or Categories**	**Value or n**	**%**	**Value or n**	**%**	**Value or n**	**%**	**Value or n**	**%**	**Value or n**	**%**	
Diagnosis of any auto-immune disease	Any time before ref date											
	1 – 2 years before ref date											
	3 - 5 years before ref date											
	> 5 years before ref date											
Diagnosis of neuro-inflammatory auto-immune disease	Any time before ref date											
	1 – 2 years before ref date											
	3- 5 years before ref date											
	> 5 years before ref date											
Diagnosis of other auto-immune disease	Any time before ref date											
	1 – 2 years before ref date											
	3 - 5 years before ref date											
	> 5 years before ref date											

CONFIDENTIAL

Table 4 Healthcare Resources utilization (all included subjects)

		Exposed F N= XXX		Unexposed F N= XXX		Concurrent M N= XXX		Historical M N= XXX		Overall N=xxx		p-value
Characteristics	Parameters or Categories	Value or n	%	Value or n	%	Value or n	%	Value or n	%	Value or n	%	
Number of GP consultation/Primary care during the year previous the reference date	Mean											
	SD											
	Median											
	Minimum											
	Maximum											
	Missing											
Number of GP consultation/Primary care during the year previous the reference date	0											
	1-2											
	3-5											
	> 5											

CONFIDENTIAL

Table 5 Data availability (all included subjects)

		Exposed F N= XXX		Unexposed F N= XXX		Concurrent M N= XXX		Historical M N= XXX		Overall N=xxx		p-value
Characteristics	Parameters or Categories	Value or n	%	Value or n	%	Value or n	%	Value or n	%	Value or n	%	
Number of years in CPRD	Mean											
	SD											
	Median											
	Minimum											
	Maximum											
	Missing											
Number of year in CPRD	[1-2]											
	[3-5]											
	[5-10]											
	[10+]											

CONFIDENTIAL

2. EXPOSURE TO OTHER VACCINES

Table 6 Exposure to other vaccines (all included subjects)

		Exposed F N= XXX		Unexposed F N= XXX		Concurrent M N= XXX		Historical M N= XXX		Overall N=xxx		p-value
Characteristics	Parameters or Categories	Value or n	%	Value or n	%	Value or n	%	Value or n	%	Value or n	%	
Any vaccines	During the previous year											
	During the one-year follow-up period											
<vaccine name 1>	During the previous year											
	During the one-year follow-up period											
<vaccine name 2>	During the previous year											
	During the one-year follow-up period											
etc.	During the previous year											
	During the one-year follow-up period											

CONFIDENTIAL

3. AUTOIMMUNE DISEASES CHARACTERISTICS

3.1. Main analysis

Table 7 Frequency of Autoimmune Diseases during the one-year follow-up period (all included subjects)

	Exposed F N= XXX		Unexposed F N= XXX		Concurrent M N= XXX		Historical M N= XXX		Overall N=xxx	
Outcomes in GPRD	Value or n	%	Value or n	%	Value or n	%	Value or n	%	Value or n	%
Primary										
Neuroinflammatory AIDs										
Other AIDs										
Individual diseases										
Multiple sclerosis										
Transverse myelitis										
Guillain-Barré Syndrome										
etc.										

CONFIDENTIAL

Table 8 Frequency of Autoimmune Diseases during specific risk period (all included subjects)

	Exposed F N= XXX		Unexposed F N= XXX		Concurrent M N= XXX		Historical M N= XXX		Overall N=xxx	
Outcomes	**Value or n**	**%**	**Value or n**	**%**	**Value or n**	**%**	**Value or n**	**%**	**Value or n**	**%**
AIDs within 2 months of reference date										
Guillain-Barré Syndrome										
Etc.										
Secondary										
AIDs within 6 months of reference date										
Idiopatihic thrombocytopenic purpura										
AIDs within 1 year of reference date										
Multiple sclerosis										
Transverse myelitis										
Optic neuritis										
etc.										

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Table 9 Incidence Rates, Incidence Rate Ratios, and 95% Confidence Intervals (CIs) of Confirmed Autoimmune Disease Events in Females cohorts

Autoimmune Disease Events	Unexposed F Cohort			Exposed F Cohort						
	Person-years (/100,000)	No. of Events	Incidence Rate	Person-years (/100,000)	No. of Events	Incidence Rate	Crude Incidence Rate Ratio	Crude Incidence Rate Ratio 95% CI	Adjusted Incidence Rate Ratio	Adjusted Incidence Rate Ratio 95% CI
Primary Endpoints[1]										
Neuroinflammatory Autoimmune Diseases										
Other Autoimmune Diseases										
Secondary Endpoints										
Multiple sclerosis										
Transverse myelitis										
Guillain-Barré Syndrome										
etc.										

Table 10 Incidence Rates, Incidence Rate Ratios, and 95% Confidence Intervals (CIs) of Confirmed Autoimmune Disease (AID) Events in Male cohorts

<Same temple as above table>

3.2. Sensitivity analyses

The same table template will be used for sensitivity analysis.

CONFIDENTIAL

Table 11 Relative incidences in the case series analysis during the risk period after vaccination with Cervarix

				95% CI		
Period	**Episodes of AIDs N=xx**	**Total person-years**	**Relative incidence**	**LL**	**UL**	**P-value**
Risk						
Control						

N = number of episodes
LL = lower limit
UL = upper limit
p-value = p-value of Wald test

The same table template will be used for the various SCCS analyses.

CONFIDENTIAL

Annex 7 TRADEMARKS

The following trademarks are used in the present study outline. Note: In the remainder of the document, the names of the vaccines will be written without the superscript symbol ™ or ®.

Trademarks of the GlaxoSmithKline group of companies	**Generic description**
Cervarix®	Bivalent human papillomavirus (types 16, 18) recombinant vaccine

Trademarks not owned by the GlaxoSmithKline group of companies	**Generic description**
Gardasil® (Merck & Co. Inc.)	Recombinant human papillomavirus quadrivalent (Types 6, 11, 16 and 18) vaccine

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Annex 8 Protocol Sponsor Signatory Approval

Protocol Sponsor Signatory Approval

eTrack study number and Abbreviated Title	116239 (EPI-HPV-040 VS UK)
Date of protocol	FDA – EMA PASS Final Version 1: 09 July 2013
Title	An observational cohort study to assess the risk of autoimmune diseases in adolescent and young adult women aged 9 to 25 years exposed to Cervarix® in the United Kingdom
Sponsor signatory	[redacted] Director, Head of Global Epidemiology Vaccine Value & Health Sciences (VVHS), GlaxoSmithKline Biologicals
Signature	_____
Date	_____

CONFIDENTIAL

116239 (EPI-HPV-040 VS UK)
Protocol FDA - EMA PASS Final Version 1

Annex 8 Protocol Sponsor Signatory Approval

Protocol Sponsor Signatory Approval

eTrack study number and Abbreviated Title	116239 (EPI-HPV-040 VS UK)
Date of protocol	FDA – EMA PASS Final Version 1: 09 July 2013
Title	An observational cohort study to assess the risk of autoimmune diseases in adolescent and young adult women aged 9 to 25 years exposed to Cervarix® in the United Kingdom
Sponsor signatory	[redacted] Director, Head of Global Epidemiology Vaccine Value & Health Sciences (VVHS), GlaxoSmithKline Biologicals
Signature	[redacted] [redacted] DIRECTOR GLOBAL EPIDEMIOLOGY.
Date	19/07/2013.

For internal use only

- ----------Checksum----------!Ver.!Created On - -
4193ee712ff403a2684a6c3246e5aed2c54b6da5 2.0 2/14/2013 11:18:59 AM - -
17ab5eb69fb64c2c680cb2c4d5e93cbe835d0e13 3.0 2/27/2013 11:49:46 AM - -
d48af067664a55c0a166b1fb9c3cc10d270e950b 2.0 2/27/2013 9:27:27 AM - -
28d06e92fa3d5ca2e79820f7d7508b9c9e99040d 3.0 2/27/2013 10:50:39 AM - -
d50f35fb1c34727873a6e5e7ab98aa61d0bd6db8 2.0 7/19/2013 1:55:13 PM - -
- -

List of Independent Ethics Committees /Institutional Review Boards

The study protocol and other information that required pre-approval were reviewed and approved by the Independent Scientific Advisory Committee (ISAC) for Medicines and Healthcare products Regulatory Agency (MHRA) database research.

Representative written information for patient and sample consent forms

No patient informed consent was needed, because the patient information in the Clinical Practice Research Datalink General Practice OnLine Database (CPRD GOLD) is fully coded and GlaxoSmithKline (GSK) Biologicals personnel were not able to make a link between the data and specific individuals.

The CPRD GOLD has obtained ethical approval from a National Research Ethics Service Committee (NRES) for purely observational research (i.e. studies that do not include patient involvement).

Important publications referenced in the report

Angelo MG, David MP, Zima J, Baril L et al. Pooled analysis of large and long-term safety data from the human papillomavirus-16/18-AS04-adjuvanted vaccine clinical trial programme. *Pharmacoepidemiology and Drug Safety* 2014a;23:466-479.

Angelo M G, Zima J, Tavares Da Silva F, Baril L, Arellano F. Post-licensure safety surveillance for human papillomavirus-16/18-AS04-adjuvated vaccine: more than 4 years of experience. *Pharmacoepidemiology and Drug Safety* 2014b: **DOI**: 10.1002/pds.3593.

Chao C, Klein NP, Velicer CM, Sy LS, Slezak JM, Takhar H, et al. Surveillance of autoimmune conditions following routine use of quadrivalent human papillomavirus vaccine. *Journal of Internal Medicine*. 2011; doi: 10.1111/j.1365 2796.2011.02467.x271(2):193-203. (Article also published in 2012 with the journal reference: 271(2):193-203).

Chao C & Jacobsen S. Evaluation of autoimmune safety signal in observational vaccine safety studies. *Human Vaccines & Immunotherapeutics* 2012;8(9):1302-1304. URL: http://www.landesbioscience.com/journals/vaccines/2012HV0049.pdf?nocache=2093950 154. Last accessed on 05-September-2014.

Signatures of principal or coordinating investigator(s) or sponsor's responsible medical officer

GlaxoSmithKline Biologicals
Vaccine Value and Health Science
Sponsor Signatory Approval Page

Please note that by signing this page, you take responsibility for the content of the Study Report, including appendices

STUDY TITLE: An observational cohort study to assess the risk of autoimmune diseases in adolescent and young adult women aged 9 to 25 years exposed to Cervarix® in the United Kingdom

Study: 116239 (EPI-HPV-040 VS UK) Development Phase: NA

I have read this report and confirm that to the best of my knowledge it accurately describes the conduct and results of the study.

Name of Sponsor Signatory: [redacted]

Title of Sponsor Signatory: Director, Head of Global Epidemiology

Signature: [redacted]

Date: March 23, 2015

For internal use only

- ----------Checksum----------!Ver.!Created On - -
a99c0822764d6722cc690ff0b555f7dc268a6994 1.0 3/17/2015 12:12:51 PM - -
8b4b89677732f882781da00c24191f72ea419be2 1.0 3/17/2015 12:12:47 PM - -
09595272375de6806df57a4e3d34b7dbbb9133e8 1.0 3/17/2015 12:12:56 PM - -
f3d04cd369336cb9e2236612784ea3d360e40fac 1.0 3/17/2015 12:12:53 PM - -
b5aabf5e91c4dcf9a3881e9ee9f06313256cf877 1.0 3/17/2015 12:12:44 PM - -
a0b79f264d60b5759046106e5c2bcb2da48cef91b 1.0 3/17/2015 12:22:27 PM - -
0ed51ea08beb7a70640a0fdf7e54f7f8ca22e9c1 3.0 3/23/2015 2:55:45 PM - -
---------- -