NOVEL Trial

A randomised controlled trial of Gardasil

Statistical Analysis Plan

Version 1.0 based on V10.0 of the Protocol dated 22/05/2024

Version 0.1 started: 26/08/2020

ClinicalTrials.gov Identifier: NCT03979014

REVISION HISTORY:

Version	Date Approved	Reason for Change	Author
V1.0		n/a	Bernard North

LIST HERE ANY AMENDMENTS TO THE SAP THAT WERE MADE AFTER THE SAP WAS SIGNED OFF BYTHE TSC

Approval Page

Trial Name	NOVEL
Sponsor	Imperial College, London
Chief Investigator	Professor Maria Kyrgiou
Version number and date	1.0
Effective Date	6.12.2024

	Name	**Position (on study)**
Author	Bernard North	Independent Statistician
	Signature	Date
Reviewed by	Shama Sheikh	Independent Statistician
	Signature	Date
Approved by	Prof Peter Sasieni	Chief Investigator (and Senior Statistician)
	Signature	Date
Approved by	Prof Maria Kyrgiou	Chief Investigator
	Signature	Date

CONTENTS

This document contains up to date statistical analysis plans (with version numbers and dates).

A) Quantitative Analysis Plan
B) Economic Analysis Plan (not applicable)
C) Schedule of Assessments and Measures

(MAY ALSO HAVE OTHERS SUCH AS QUALITATIVE ANALYSIS PLAN)

LIST OF ABBREVIATIONS & GLOSSARY

AE	Adverse Event
AGC	Abnormal Glandular Cells
AIS	Adenocarcinoma in situ
ASCH	Abnormal Squamous Cells, cannot exclude high-grade squamous intraepithelial lesions
BGSC	British Gynaecological Cancer Society
ccfDNA	Circulating Cell-free DNA
CGIN	Cervical glandular intraepithelial neoplasia
CIN	Cervical Intraepithelial Neoplasia
CIN1+	Cervical Intraepithelial Neoplasia grade 1 or higher, (i.e. low or high-grade CGIN, CIN1, CIN2, CIN3 or invasive cervical cancer)
CIN2+	Cervical Intraepithelial Neoplasia grade 2 or higher, (i.e. high-grade CGIN, CIN2, CIN3 or invasive cervical cancer)
CIN3+	Cervical Intraepithelial Neoplasia grade 3 or higher, (i.e. high-grade CGIN, CIN3 or invasive cervical cancer)
cGIN	Cervical Glandular Intraepithelial neoplasia
Colposcopy	A procedure which uses a more detailed view of the cervix using a special microscope to see the changes at high magnification with good lighting
CPTU	Cancer Prevention Trials Unit
CTCAE	Common Terminology Criteria for Adverse Events
CTU	Clinical Trials Unit
DNA	Deoxyribonucleic acid
eCRF	Electronic Case Report Form
EDC	Electronic Data Collection
HIV	Human Immunodeficiency Virus
HPV	Human Papillomavirus
HSIL	High-grade Squamous Intraepithelial Lesion
ICF	Incomplete Failure – an HPV sample that cannot be analysed usually due to insufficient sample. There is usually a repeat attempt to analyse such a sample.
IMP	Investigational Medicinal Product
ITT	Intention to Treat
LBC	Liquid-Based Cytology
MAR	Missing at random
NCRI	National Cancer Research Institute
NICE	National Institute for health and Care Excellence
NIHR	National Institute for Health Research
P	The imputation probability in section 3.1.2 (different for each participant and zero for those with no uncertain persistence)
P0	In the imputation section 3.1.2 the probability of persistence given a single positive.
QA	Quality Assurance
QM/QMUL	Queen Mary, University of London
REC	Research Ethics Committee
RCT	Randomised Controlled Trial
SAE	Serious Adverse Effect
SAP	Statistical Analysis Plan

SAR	Serious Adverse Reaction
T0	In 3.1.2 where persistence is imputed in uncertain persistence this is the date of the first positive persistent test where there is uncertain persistence from 1 or 2 positive tests.
Test of Cure	An HPV test taken after treatment for cervical abnormalities to confirm if a woman has been successfully treated
TMG	Trial Management Group
TN	In 3.1.2 where persistence is imputed in uncertain persistence following one or two positive tests this is the time of a potential later negative test (more than 6 months after the first or only positive test).
TP	In 3.1.2 where persistence is imputed in uncertain persistence this is the date of the second positive persistent test where there is uncertain persistence from 2 positive tests.
TSC	Trial Steering Committee

A) QUANTITATIVE ANALYSIS PLAN

Investigators

Professor Maria Kyrgiou
Consultant Gynaecologic Oncologist
Level 2, Hammersmith House
Imperial College London
Hammersmith Hospital Campus
Du Cane Road
London W12 0HS
United Kingdom
Email: m.kyrgiou@imperial.ac.uk
Tel: +44 (0)20 8383 5290
Fax: +44 (0)02 8383 8065

Associate Professor Bjorn Strander
Department of Obstetrics and Gynaecology,
Institute of Clinical Sciences, Sahlgrenska Academy,
University of Gothenburg,
Sweden

Associate Professor Pekka Nieminen
Department of Obstetrics and Gynaecology, University Hospital Helsinki
Box 615 00029 HUS
Finland

Professor Peter Sasieni
Centre for Cancer Screening,
Prevention and Early Diagnosis,
Wolfson Institute of Population Health,
Queen Mary University of London,
Whitechapel Campus,
67-75 New Road,
E1 1HH
Email: p.sasieni@qmul.ac.uk

Chief investigator

Professor Maria Kyrgiou
Consultant Gynaecologic Oncologist
Level 2, Hammersmith House
Imperial College London
Hammersmith Hospital Campus
Du Cane Road
London W12 0HS

United Kingdom
Email: m.kyrgiou@imperial.ac.uk
Tel: +44 (0)20 8383 5290
Fax: +44 (0)02 8383 8065

Trial manager

Áine Hogan
Imperial College London
3rd Floor Radiotherapy Building
Hammersmith Hospital Campus
Du Cane Road
London
W12 0NN
Tel: +44 (0) 203 311 5203
Fax: +44 (0) 203 311 7443

Trial statistician

Professor Peter Sasieni
Centre for Cancer Screening,
Prevention and Early Diagnosis,
Wolfson Institute of Population Health,
Queen Mary University of London,
Whitechapel Campus,
67-75 New Road,
E1 1HH
Email: p.sasieni@qmul.ac.uk

Independent statistician / Author

Bernard North
Centre for Cancer Screening,
Prevention and Early Diagnosis,
Wolfson Institute of Population Health,
Queen Mary University of London,
Whitechapel Campus,
67-75 New Road,
E1 1HH
Email: b.v.north@qmul.ac.uk

Independent statisticians

Nefeli Kouppa
Centre for Cancer Screening,
Prevention and Early Diagnosis,
Wolfson Institute of Population Health,
Queen Mary University of London,
Whitechapel Campus,

67-75 New Road,
E1 1HH
Email: n.kouppa@qmul.ac.uk

Shama Riaz Sheikh
Centre for Cancer Screening,
Prevention and Early Diagnosis,
Wolfson Institute of Population Health,
Queen Mary University of London,
Whitechapel Campus,
67-75 New Road,
E1 1HH
Email: s.sheikh@qmul.ac.uk

1. Description of the trial

This document contains the **Statistical Analysis Plan** (SAP) for the NOVEL study. The SAP is required by the National Institute of Health to improve reproducibility, transparency and validity of clinical trials.
The table of contents of this SAP follows the one recommended in SOP KCTU ST 02 “Statistical Analysis Plan”, version 3.0.
The SAP is based on the NOVEL protocol V10 dated 22 May 2024
NOVEL Protocol v10_22May2024.docx.

1.1 Principal research objectives to be addressed

Primary Objective

The primary objective is to demonstrate whether the Gardasil®9 nonavalent HPV vaccine initiated at the time of local cervical treatment reduces persistent HPV infection (type specific positive tests ≥5.5 months / >166 days apart without intervening type-specific negative test) by the 7 oncogenic vaccine HPV types (of the 9 types targeted by the vaccine i.e. the oncogenic types 16/18/31/33/45/52/58 but not 6 and 11) in women treated for high-grade CIN. This is performed via assessment of a reduction in the vaccine arm of a composite endpoint that is a weighted sum of 3 types of persistent infection (incident, recurrent, prevalent) defined via HPV test results over the 24-month follow-up with a potential extra 30 month test for those who have uncertain persistence at the 24 month test e.g. they have a first positive at the 24 month test.

Secondary Objectives

- To evaluate the efficacy of the vaccine against the three components (incident, recurrent, prevalent) separately of the composite endpoint based on the 7 oncogenic vaccine HPV types (16/18/31/33/45/52/58).
- To evaluate the efficacy of the vaccine against post-treatment persistent cervical infections (incident, recurrent, prevalent) with any of the 14 oncogenic HPV types (HPV 16/18/31/33/35/39/45/51/52/56/58/59/66/68).
- To evaluate the efficacy of the vaccine against any cervical infection, not necessarily persistent as overall composite and incident, recurrent, and prevalent infection with any of the 7 oncogenic vaccine HPV types (16/18/31/33/45/52/58).
- To evaluate the efficacy of the vaccine against any cervical infection, not necessarily persistent (overall and incident, recurrent, prevalent) with any of the 14 oncogenic HPV types 16/18/31/33/35/39/45/51/52/56/58/59/66/68.
- To evaluate the efficacy of the vaccine against new, post-treatment, CIN2+ lesions (high-grade squamous intra-epithelial lesions (HSIL))

associated with the 7 oncogenic vaccine HPV types (HPV 16/18/31/33/45/52/58).

- To evaluate the efficacy of the vaccine against new, post-treatment, CIN2+ lesions (HSIL) associated with any of the 14 oncogenic HPV types 16/18/31/33/35/39/45/51/52/56/58/59/66/68.
- To monitor the safety of Gardasil®9 with the first dose given at localised cervical treatment.

Tertiary Objectives

- To evaluate the efficacy of the vaccine against new, post-treatment, CIN1+ lesions (high-grade squamous intra-epithelial lesions (HSIL)) associated with the 7 oncogenic vaccine HPV types (HPV 16/18/31/33/45/52/58).
- To evaluate the efficacy of the vaccine against new, post-treatment, CIN1+ lesions (HSIL) associated with any of the 14 oncogenic HPV types.

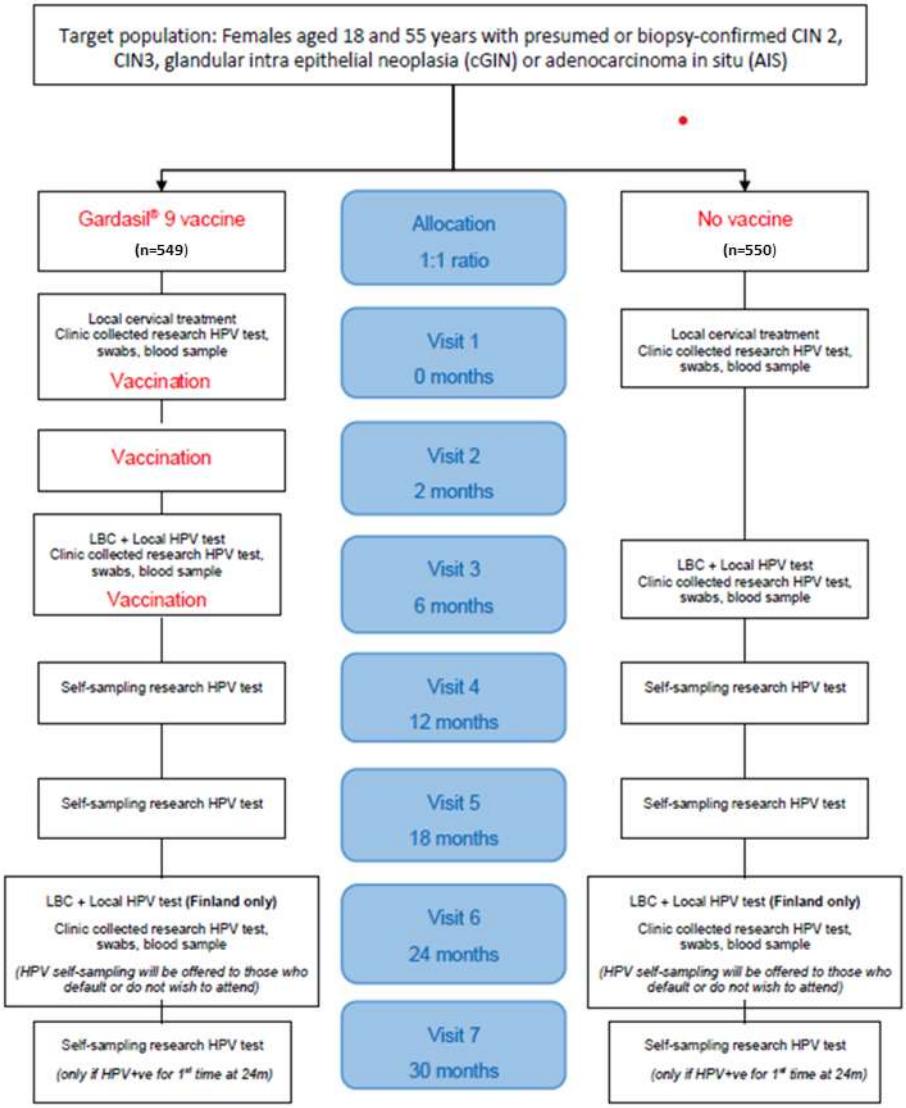
1.2 Trial design including blinding

NOVEL is a phase III, observer-blind controlled, randomised multicentre study to investigate if the nonavalent prophylactic HPV vaccine Gardasil®9 reduces persistent HPV infection in women treated for high-grade CIN.
The trial consists of two parallel arms to which patients will be randomised 1:1: Gardasil®9 vaccine versus No vaccine. Patients are females aged between 18 and 55 years with presumed or biopsy-confirmed CIN2, CIN3, glandular intra epithelial neoplasia (cGIN) or adenocarcinoma in situ (AIS). The patients and clinicians will not be blinded but the laboratory staff performing the HPV assays on vaccinated and non-vaccinated arms will be blinded. The trial will be performed at approximately 16 investigational sites in the UK, Finland and Sweden. Each arm will enrol around 550 patients, for a total of 1099 patients in the entire study.

- Arm 1 – Vaccine: Gardasil®9 vaccine (N=549): administered at 0, 2 and 6 months.
- Arm 2 – Control: no vaccine (N=550).

The trial randomisation is stratified by country and study site (clinic).

Figure 1. Trial design flow diagram



1.3 Method of allocation of groups

After eligibility has been confirmed, patients will be randomised to one of the treatment arms. Randomisation will be done in a 1:1 ratio. Randomisation will be performed centrally using the InForm eCRF; there is no option available for manual randomisation. Upon randomisation each patient will be allocated a unique ID which should be used in all future correspondence. Randomisation will be stratified by country and study site (clinic) with variable block sizes to ensure that equal numbers of patients are allocated to the 2 arms within each stratum.
The procedure is as follows:

1. On receipt of the baseline questionnaire, the Trial Co-ordinator electronically submits details of each participant to the CTU. This includes participant ID number, site, initials and date of birth.
2. The system immediately notifies the relevant site and records the randomisation outcome. However, the Trial Co-ordinator does not receive the randomisation outcome.

Please refer to the eCRF Completion Manual for further details on patient randomisation.

1.4 Duration of the treatment period

The duration of the trial is 51 months, with patient recruitment of 12 months. The vaccination in the active arm occurs at 0-, 2- and 6-months post randomisation. Both arms receive local cervical treatment at 0 months.

1.5 Frequency and duration of follow-up

Participants will receive HPV tests at 0, 6, 12, 18 and 24 months (plus potential additional HPV test at 30 months described below). The HPV tests at 0, 6 and 24 months are clinic based, others being self-sampling, and are accompanied by swabs and blood samples. There will be no additional colposcopies beyond clinical indications and guidelines.
Some participants will be offered an additional 30-months test if they meet the following criteria:

- Participants who are positive for the first time at 24 months (for a particular type).
- Participants who have two positive tests of the same type, less than 5.5 months apart so not sufficient for persistence, with the second occurring at 24 months and no negative tests for that type in between the two tests.

In the UK, all women post-treatment have a cytology and HPV test (‘test of cure’) at 6m. In Finland, all women post-treatment have a cytology and HPV test at 6m.

1.6 Visit windows

When assessment of a patient should take place are outlined below.

- Assessment days are relative to the start of visit 1 at 0 months i.e. baseline, the date of local treatment. With the exception of the vaccine visits at 2 and 6 months scheduled follow-up assessments should take place only after a 5.5 month (>166 days) period has passed since the previous assessment.
- If a positive test occurs within 5.5 months from a previous positive test for the same HPV type then it can't be considered the 2nd test in an identification of persistence.
- If there is a negative test for the same HPV type as a previous positive test, then if there is also a later positive test for that HPV type that is more than 5.5 months from the positive test preceding that negative test then the intervening negative test will rule out the two flanking positive tests from defining persistence.
- If a patient is less than 8 weeks late for a test, the next test should be delayed ensuring 5.5 months between tests. However, if a patient is more than 8 weeks late, they will return to their normal schedule. If the last test (i.e. 24 months) is positive for a particular HPV type and the previous non-missing test was negative for that type or the previous test was positive for that type but less than 5.5 months earlier, then there should be an additional test (nominally at 30 months).
- If any test is out of sync, then samples should be taken as close to the original planned test date as possible. For the vaccine visits, the second dose should be administered at least one month after the first dose and the third dose should be administered at least 3 months after the second dose. All three doses should be given within a 1-year period.

Persistence will be defined as two post-baseline (>121 days from randomisation) tests at least 5.5 months apart (>166 days) that are both positive for the same HPV type and with no intervening negative (for that type) test.

1.7 Data collection

Eligibility screening

- Inclusion criteria
    - Female (18-55y) attending for local treatment for presumed CIN2 (cytological and colposcopy impression) OR presumed CIN3 (cytological and colposcopy impression) OR presumed cGIN/AIS (cytological and colposcopy impression) OR biopsy-

confirmed CIN2 OR biopsy-confirmed CIN3 OR biopsy-confirmed CGIN/AIS.

- Written informed consent obtained from the subject prior to enrolment.
- Free of other relevant health problems as established by medical history and clinical examination, e.g. immunosuppression.
- Patients who the investigator believes can and will comply with the protocol requirements (e.g. attendance at clinic appointments and return for follow-up visits).

- Exclusion criteria
  - Use of other investigational/non-registered product within 30 days preceding the 1st vaccine dose.
  - Continuous administration of immunosuppressants.
  - Previous vaccination against HPV.
  - Cancer or autoimmune disease under treatment. Patients who have a history of cancer or autoimmune disease but are not currently being treated for the condition will be included.
  - Any confirmed or suspected immunosuppressive condition, including HIV infection.
  - History of allergic disease or any neurologic disorders likely to interact with study vaccination.
  - Acute febrile disease at enrolment (will be postponed).
  - Pregnant women or women intending to get pregnant during the next 6 months (if pregnant during follow-up, remaining doses will be delayed until after delivery).

Measures

- Baseline
  - HPV presence and type.
  - CIN grade: CIN2/3, cGIN/AIS (cytology, colposcopy or biopsy).

- Primary outcome measures
  - Weighted composite of the following 3 endpoints concerning the efficacy of the Gardasil®9 HPV vaccine as compared to no vaccine, all of which will be evaluated at 24 months after the first dose in female patients aged 18-55 years at the baseline local treatment:

• Persistent incident (I) (≥5.5 month interval i.e. >166 days) cervical infections with 7 oncogenic vaccine HPV types 16/18/31/33/45/52/58. Incident infection is defined as an HPV type not detected at baseline but detected post-baseline i.e. at 4+ months, >121 days, post randomisation.
• Persistent recurrent (R) (≥5.5 month interval) cervical infections with vaccine HPV types 16/18/31/33/45/52/58. Recurrent is infection with a type present at baseline, but not detected at a post-treatment (4+ months ie >121 days post-randomisation date) study test prior to a persistent (2 tests 5.5+ months i.e. >166 days apart with no intervening test negative for that type) detection of the same type.
• Persistent prevalent (P) (≥5.5 month interval) cervical re - infections with vaccine HPV types 16/18/31/33/45/52/58. Prevalent infection is defined as an HPV type present at baseline and at 2 post-treatments (4+ months. >121 days, post-baseline i.e. randomisation and ≥5.5 months, >166 days, apart) tests with no intervening or preceding test negative for that type.

Each participant will have a value of an occurrence of one of the 3 persistence class categories for each HPV type or not persistently infected if not persistent for that HPV type. Each participant will be assigned an overall persistence status based on worst case of persistence over the set of HPV types they are persistent for (with worst based on Incident being higher than Recurrent which is higher than Prevalent) The composite endpoint will be a weighted sum: 6 x “incident” + 3 x “recurrent” + 1 x “prevalent infection” where “incident” means proportion of those whose worst persistent status, over all HPV types, is incident etc

- Secondary outcome measures
    - The three separate components (incident, recurrent, prevalent) of the composite endpoint for any of the 7 oncogenic vaccine HPV types 16/18/31/33/45/52/58.
    - Post-treatment persistent cervical infections (composite of incident, recurrent, prevalent) with any of the 14 oncogenic HPV types 16/18/31/33/35/39/45/51/52/56/58/59/66/68.
    - Post-treatment not necessarily persistent cervical infections (composite of incident, recurrent, prevalent) with any of the 7 oncogenic vaccine HPV types 16/18/31/33/45/52/58.
    - Post-treatment not necessarily persistent cervical infections (composite of incident, recurrent, prevalent) with any of the 14 oncogenic HPV types 16/18/31/33/35/39/45/51/52/56/58/59/66/68.
    - New, post-treatment, CIN2+ lesions (high-grade squamous intra-epithelial lesions (HSIL)) associated with the 7 oncogenic vaccine HPV types 16/18/31/33/45/52/58.

- New, post-treatment, CIN2+ lesions (HSIL) associated with any of the 14 oncogenic HPV types 16/18/31/33/35/39/45/51/52/56/58/59/66/68.

- To monitor the safety of Gardasil®9 with the first dose given at localised cervical treatment.

- Tertiary outcome measures

  To tabulate
  - CIN1+ associated with 7 oncogenic vaccine types.
  - CIN1+ associated with any of the 14 oncogenic HPV types.

- Mediators of treatment
  Comparison of the composite between the control arm and active arm women receiving 0,1,2 or all 3 vaccine doses will be undertaken.

- Moderators of treatment
  1. We shall examine prevalence of incident, prevalent and recurrent HPV according to local cervical treatment results including age, depth of excision, incomplete excision/ involved margins and HPV type, type 16 vs other, at baseline (not for incident). We will tabulate proportions of worst case (over type) of incident, prevalent and recurrent HPV against age (<35, 35-44, >44 years at randomisation), binary occurrence of incomplete excision, involved margins, mean excision depth, and baseline HPV type (HPV 16 vs other types combined).
  2. The mean, SD and quartiles of the composite endpoint will be explored for incomplete excision and incomplete margins, against excision depth (relative to its median cut-off value) and by baseline and follow up HPV type.

- Process indicators
  1. There will be a tabulation of the numbers and percentages of randomised participants returning an analysable (i.e. not ICF) HPV test at each timepoint. We shall also tabulate those with an ICF (incomplete failure) sample and those who have withdrawn prior to that timepoint.
  2. Recruitment numbers (and percentage) for each month per site and treatment arm will be tabulated.

- Adverse events

Reportable AEs and SAEs with CTCAE grade, seriousness, causality and outcome will be recorded throughout the study, from the point of consent until the end of follow-up; they will be followed up according to local practice until the event has stabilised or resolved, or the Follow-up Visit, whichever is the sooner.

- Additional pre-randomisation (baseline) measures

  Conisation depth and positive/negative cone margins.

- Additional post-randomisation (follow-up) measures

  Liquid based cytology measures (LBC).

1.8 Sample size estimation (including clinical significance)

Consider three types of HPV infection with the 7 high-risk vaccine types: prevalent (P), incident (I) and recurrent (R). In the paper by Soderlund-Strand et al (2014), there were 25 prevalent, 17 Incident and 9 recurrent. Thus the percentages of infections P:I:R were 49%:33%:18%.

In England, data from colposcopy clinics consulted for this study suggest that 15% of treated women will be HPV positive at 6 months post infection. We should also allow for 3% to gain (persistent) infections after 6 and before 24 months. If we take the proportion with persistent infection 6-24 months post treatment to be 18%, then we would expect 9% to have the same types as present at diagnosis (P), 6% to have a new infection (I) and 3% to have a recurrence of an infection present at diagnosis but absent at 6 months post treatment (R).

We assume that the efficacy of vaccination will be 80% for I, 20% for P and 50% for R. Thus, in the two arms the percentages will be 6.0% and 1.2% for I, 9.0% and 7.2% for P and 3.0% and 1.5% for R. The optimal weighting (to maximise the power) for a composite outcome gives weights proportional to the expected difference in the proportions divided by the variance of that difference. The formula for this is $(p1-p0)/\{p1(1-p1)+p0(1-p0)\}$. Which yields: 0.703 for I, 0.121 for P and 0.342 for P. This is approximately 6:1:3. Hence we take the optimal composite score, C to be the trial endpoint defined as $C = 6I+P+3R$ where I,P,R are the proportions of women with incident, prevalent and recurrent HPV in each treatment arm.

Now the expected value of the composite in control patients is 0.54 and in vaccinated patients it is 0.189 and the variance of the composite is 2.37 (SD=1.54) in controls and 0.627 (SD = 0.792) in vaccinated patients. Treating these as normal random variables, the sample size required for 90% power is 256 per arm.

If, however, the proportion of women with *persistent* infection is only 12% (rather than 18%) then the numbers needed per arm are 391 for 90% power. Allowing for drop-out / non supply of self-samples means that we would want 830-975 in total depending on how much power and the level of drop out.

Given the low proportions of test kit returns from the UK sites in Version 6 of the protocol we increased the number recruited to 1099 (549 and 550 per arm) to further ensure power allowing for 22% dropout/non-supply of self-samples. Note that if the proportions (In the absence of vaccination) with persistent infection is rather greater than 12%, then we will have good power even with greater dropout.

1.9 Brief description of proposed analyses and any pre-analysis statistical checks required

Analyses will be carried out by the trial statistician. In the first instance data will be analysed under a modified intention-to-treat assumptions. Under ITT we would analyse all those with data in groups as randomised irrespective of treatment received based on a treatment policy estimand.
For the primary analysis we exclude participants who do not return any post-baseline (>121 days post randomisation date) HPV tests or, for the main primary analysis, those who do not return a baseline HPV test and so therefore cannot have any post-baseline persistence characterised as incident, recurrent or prevalence.
The anticipated intercurrent events are discontinuation from study medication (i.e. before month 6 when the final vaccine dose is administered to the vaccine arm) due to:

- Patient decision.
- Significant adverse events or unacceptable toxicities.
- Severe non-compliance to this protocol as judged by the Investigator.
- Allergic reaction to study medication.
- If the Investigator considers that a patient's health will be compromised due to adverse events or concomitant illness that develop after entering the study.
- Use of any investigational or non-registered product other than the study vaccine during the study.
- Patients with history of cancer or autoimmune disease who has relapsed following vaccination.
- Continuous administration of immune-suppressants.
- Newly diagnosed immunosuppressive condition, including HIV infection.

Other intercurrent events can result in patient withdrawal from the study resulting in non-availability of HPV tests for those HPV test occasions scheduled after the withdrawals

- Patient decision
- Loss to follow-up
- Death
- Investigator decision
- Diagnosis and treatment of cancer

Patients are classified into the 3 categories of persistent HPV infection endpoint: Incident (I), Recurrent (R), Prevalent (P) or Not Persistent (no infection or a non-persistent infection). For the primary analysis this is based on HPV positivity with the 7 oncogenic HPV types 16/18/31/33/45/52/58 at baseline and months 6 ,12, 18 and 24 and, potentially month 30. Each participant will have either one of the 3 persistence categories or be non-persistent for each of the HPV types (7 for the primary analysis). Each participant will be assigned an overall persistence status based on worst case of persistence over the set of HPV types they are persistent for (with worst based on Incident being higher than Recurrent which is higher than Prevalent). The composite endpoint will be a weighted sum: 6 x “incident” + 3 x “recurrent” + 1 x “prevalent infection” where “incident” means proportion of those whose worst persistent status is incident etc.

The estimand is the difference between treatment arms of the trial endpoint defined as C = 6I+P+3R where I, P, R are the proportions of women with incident, prevalent and recurrent HPV in each treatment arm. This is compared by estimating the difference between arms of the trial endpoint C = 6I+P+3R i.e. between C_1 and C_2 divided by its standard error of treatment effect: this expression is denoted Z (1= control, 2=vaccine, so higher, positive values of Z indicate higher persistence levels in the control arm)

$$Z = (C_1 - C_2) / \text{ sqrt } ((6^2 I_1 (1-I_1)+P_1 (1-P_1)+3^2 R_1 (1-R_1)) / n_1 + (6^2 I_2 (1-I_2)+P_2 (1-P_2)+3^2 R_2 (1-R_2))/ n_2)$$

and comparing the result to a z-statistic using a two-sided alpha level of 0.05. This difference is calculated separately for each site/country used in the stratified randomisation to obtain a statistic Zi for each site i and the test is based on an overall test statistic from the Zi’s weighted by site sample size.

2. Data analysis plan – Data description

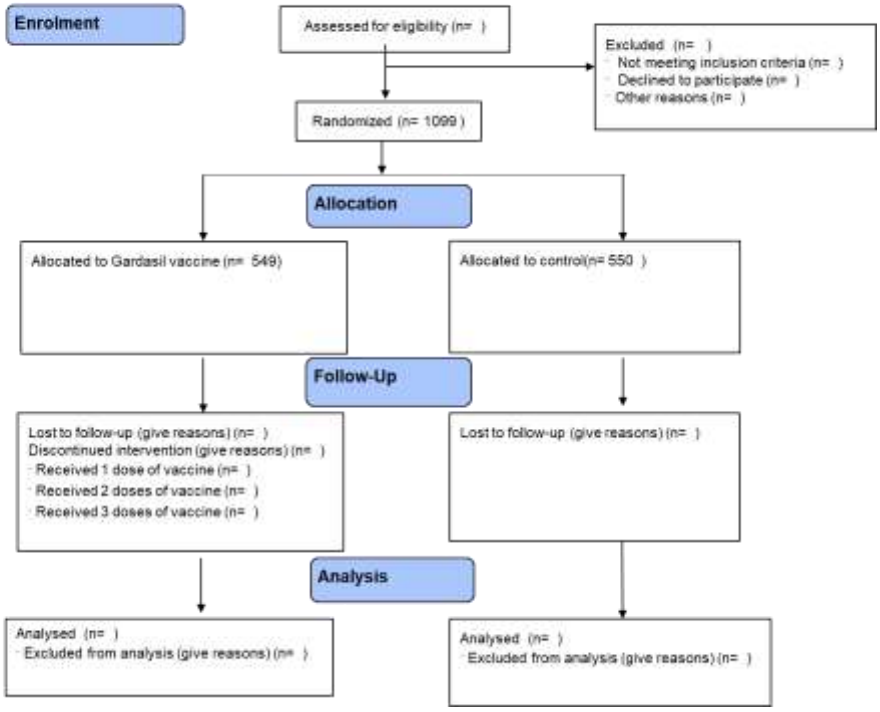
2.1 Recruitment and representativeness of recruited patients

CONSORT flow chart will be constructed (Moher, 2001) – see Figure 2. This will include the number of eligible patients, number of patients agreeing to enter the trial, number of patients refusing; then, by treatment arm: the

number of patients not/inadequately/adequately treated OR compliant/non-compliant, the number continuing through the trial, the number withdrawing, the number lost to follow-up and the numbers excluded/analysed.

Patients inadequately treated in the vaccine arm are those who, due to intercurrent events leading to study withdrawal, do not receive all 3 doses of vaccine.

Figure 2. CONSORT diagram for NOVEL trial



2.2 Baseline comparability of randomised groups

The following baseline descriptions of participants by treatment arm and overall will be presented:

- For continuous variables means and standard deviations.
- For categorical variables numbers and proportions.

The variables listed will be age, proportions with baseline HPV positive vs negative and biopsy/presumed baseline CIN2 and CIN3. No significance testing between arms will be performed.

2.3 Adherence to allocated treatment and treatment fidelity

1. Compliant versus non-compliance with the vaccine treatment will be described in terms of baseline variables.

2. The proportions receiving 1,2 or 3 doses of vaccine will be summarised by proportions overall and by age groups. The reasons for withdrawal from treatment will be summarised.

2.4 Loss to follow-up and other missing data

1. The proportions who have withdrawn prior to each timepoint will be summarised as a count and percentage of the randomised arm.
2. The proportions of participants missing each test kit return will be summarised in each arm and at each time point as a count and percentage of the randomised arm.
3. The baseline characteristics of those missing follow up i.e. who are withdrawals will be compared to those with complete follow up.
4. The reasons for withdrawal from the trial will be summarised.

2.5 Adverse event reporting

Adverse events (AE), adverse reactions (AR), related and unexpected serious adverse events (SAE) and serious adverse reactions (SARs, based on investigator causality rating, yes/no) will be summarised by

1. Treatment arm and by site.
2. Treatment arm by CTCAE maximum grade and seriousness.

2.6 Assessment of outcome measures (unblinding)

Patients and clinicians are not blinded in this study.

2.7 Descriptive statistics for outcome measures

Patients will be classified into the 3 categories Incident (I), Recurrent (R), Prevalent (P) or Not Persistent (no HPV infection or a non-persistent infection) based on HPV positivity with HPV types 16/18/31/33/45/52/58 at visits baseline and months 6 ,12, 18 and 24 and, potentially, month 30, or classified as no persistent infection detected. These proportions will be presented by treatment arm with Wilson's confidence intervals.
The main study endpoint C = 6I+P+3R, where I, P, R are the proportions of women with incident, prevalent and recurrent HPV in each treatment arm, will be presented in each treatment arm, combined over sites, with Wald confidence intervals.

2.8 Description of therapists/therapies

The intervention in the active arm 1 is Gardasil®9 vaccine (N=549): administered at 0, 2 and 6 months. There is no vaccine or dummy vaccine in the control arm.

3. Data analysis plan – Inferential analysis

3.1 Main analysis of treatment differences

The main statistical analyses will estimate a treatment policy estimand for the difference in the composite outcome between patients randomised to vaccine and to no vaccine by intention to treat restricting to those who return at least one post-baseline HPV test result and those with a baseline HPV test result. For the main primary analysis of the composite persistence outcome a baseline HPV test result is also required. In addition, the secondary analyses that involve a composite endpoint from a categorisation into I, P or R (either with the 14 HPV types or for not necessarily persistent infections) also require a non-missing baseline HPV test result. Group difference estimate with p-value will be reported.

3.1.1 Analysis of primary outcomes

Main Treatment Policy Estimand: The composite measure is 6I+ P+3R where P, I and R are the proportions of patients with prevalent, incident and recurrent vaccine HPV-type infections, respectively.

As stated in 3.1 patients who do not return at least one post-baseline HPV test result will not be analysable and will be excluded even from ITT analysis. Patients who are missing the baseline HPV test will also be excluded from the primary analysis even from the ITT analysis. This is because those with a missing baseline may still have a later persistence but it will not be categorisable into the 3 classes of persistence so will not be analysable leading to their being excluded from the analysis of the primary composite outcome. Further those with a missing baseline test will be excluded from the primary analysis (and, in fact, those secondary analyses that involve I,P,R categorisation) whether or not they have a later persistence. However, since persistence itself can be identified for individuals without a baseline test a secondary analysis described below will include all ITT individuals, with a post-baseline test, even without the baseline test, and will assess the effect of treatment on a single binary outcome of persistence without prevalent/incident/recurrent categorisation.

We define non-persistence and the 3 categories of persistence with each patient having a classification for each of the 7 possible vaccine HPV types.

- Patients who are not defined by the subsequent criteria as persistent of type incident, recurrent or prevalent for a vaccine HPV type are classified as Not Persistent i.e. they do not have 2 positive tests for a particular HPV type more than 5.5 months apart without an intervening test negative for that type. For example, a patient who is HPV negative at all recorded post-baseline tests, one who has a single positive test for a type followed by a type-specific negative test, and a patient with two positive tests that are either less than 5.5 months apart or have an intervening type-specific negative test are Not Persistent.
- Prevalent patients are positive for the same HPV type at baseline for two tests, at 4+ months / >121 days post-baseline, that are greater

than 5.5 months apart (i.e., the second at least 167 days after the first) with no preceding or intervening samples with a type-specific negative HPV test even if later tests are negative.

- Incident patients are HPV-type negative at baseline but later become HPV-type persistent i.e. HPV type Q positive at two samples, at 4+ months, >121 days post-baseline (baseline in this context being the randomisation date as this is present for all participants), and greater than 5.5 months apart (i.e., the second at least 167 days after the first , >166 days apart) with no intervening samples with a negative HPV test (there can be an intervening positive test that is within 5.5 months each of the flanking tests).
- Recurrent patients for a particular type Q are HPV positive at baseline, negative for that same type Q at a test at 4+ months, >121 days post-baseline, and then, following that type-specific negative test, they are HPV positive, for the same type Q as baseline, at two tests whose occurrence is ≥5.5 months with no intervening sample with a negative HPV type Q test. So, for example, if the baseline HPV test is positive for a type Q and then the 6 month test is missing and the 12 month test is negative then recurrent persistence can occur with 18 and 24 month both being positive.
  If patients have a persistence prior to any type-specific post-baseline negative test and following this, have a type-specific negative test and then a second occurrence of persistent tests they are classified as recurrent based on the 2nd occurrence of persistence and not prevalent based on the first occurrence of persistence. This would occur with the following sequence (and in general requires a 30 month test result): Baseline=POS, 6 months = POS, 12 months = POS, 18 months = NEG, 24 months = POS, 30 months = POS.

If a patient is HPV positive at 24 months for a type not present at 18 months or positive at 18 and 24 month samples but with a <5.5 months interval between them, then the patient will also need to be positive at a repeat 30 month sample (ideally scheduled > 5.5 months after the 24 month sample) in order to establish persistence if it has not been established by previous successive tests that are ≥ 5.5 months apart.
If a participant's last test is less than 24 months post randomisation and is negative then the patient's follow-up will be censored at the time of that test.

The HPV test will report which type a woman is infected by and a woman could be infected by more than one type during the study.
The preceding scoring will be applied separately for each HPV type so that a patient potentially could, for example, be Incident for one type but Prevalent for another.
For those patients persistent for more than 1 HPV type each will receive a single overall persistence status, contributing to the composite endpoint, derived from her individual persistence statuses for each HPV type and based on ranking the 3 types of persistence status with Incident being higher than Recurrent which is higher than Prevalent. For example, a woman who is Incident for one (or more) HPV types and possibly Recurrent and/or Prevalent

for other types will count as overall Incident irrespective of the status of those other types and a woman who is not Incident for any types but is Recurrent for at least one type will count as Recurrent irrespective of any Prevalent types.

The estimand is the difference between treatment arms of the trial endpoint defined as the composite C = 6I+P+3R where I, P, R are the proportions of women with incident, prevalent and recurrent HPV in each treatment arm, the denominator being all randomised women with a baseline test result and at least one post-baseline test result. This is calculated by estimating the difference between arms of the trial endpoint C = 6I+P+3R divided by its SD, the square root of the sum over the two arms of the variance.

Var=
$6^2I_1 (1-I_1)+P_1 (1-P_1)+3^2R_1 (1-R_1)) / n_1 + (6^2I_2 (1-I_2)+P_2 (1-P_2)+3^2R_2 (1-R_2))/ n_2$

To obtain

$$Z = (C_1 - C_2) / \text{ sqrt } ((6^2I_1 (1-I_1)+P_1 (1-P_1)+3^2R_1 (1-R_1)) / n_1 + (6^2I_2 (1-I_2)+P_2 (1-P_2)+3^2R_2 (1-R_2))/ n_2)$$

where C_1 and C_2 are the values of the composite in the 2 arms (1= control, 2=vaccine, so higher, positive, values of Z indicate higher persistence levels in the control arm) and P_1, R_1 , I_1 are the proportions of the composite P,R,I components in the control arm and P_2, R_2, I_2 for the vaccine arm.
The NOVEL randomisation is stratified by country and study site and this will require the analysis of the primary endpoint to also be stratified. However, we will stratify the analysis just by the 3 countries because some of the sites within countries have a relatively low number of participants. So the difference in value of the composite divided by its standard deviation i.e. the standard error of the estimate of the difference in the value of the composite (the square root of the sum of the variances in each arm) will be calculated for each country, I, individually and the sum of the weighted Z_i statistics for the 3 countries, is taken as the overall test statistic.
The weight for test statistic Z_i from country i is
$w_i = [(n_{1i}n_{0i}) / (n_{1i} + n_{0i})] /\Sigma_j [(n_{1j}n_{0j}) / (n_{1j} + n_{0j})]$

Where n_{1i} and n_{0i} are the number in each arm for each country. We then obtain a normalised test statistic Z over the 3 countries k=3.

$$Z = \frac{\sum_{i=1}^{k} w_i \ Z_i}{\sqrt{\sum_{i=1}^{k} w_i^2}}$$

We will obtain the p-value from comparing the normalised test statistic to the distribution of a standard normal deviate z statistic and take significance as a two-sided alpha level of 0.05 (i.e. significant if the absolute value of the test statistic exceeds 1.96).

3.1.2 Secondary Analysis of primary outcome with imputation in cases of uncertain persistence

We will perform a secondary analysis allowing imputation of persistence in uncertain cases:
The analysis of 10 imputed data sets will follow that of the main analysis for the 7 oncogenic vaccine types except that there will be no stratified analysis by country. As with the main primary analysis this imputed analysis is restricted to those participants with a baseline HPV sample result and at least 1 post-baseline sample result.

This procedure of imputation is done, as is the main unimputed analysis, for each HPV type individually, in order to obtain 10 data sets with imputed persistence for each type where persistence is uncertain, in a sense defined later, based on the actual sequence of self-sample test results. A persistence class, Prevalent or Recurrent or Incident, is assigned to the imputed persistence occurrences based on the value of the baseline test for that type and prior occurrence of a type-specific negative test.
Then, for each of 10 imputed data sets, the worse persistence class (over HPV types with Incident worse than Recurrent worse than Prevalent) is again obtained for each participant (for those where a participant has persistence, either based on actual test results or possibly as a result of imputation, for more than one HPV type or has multiple occurrences of persistence for the same type), and then from the proportions of each class of worst persistence in each treatment arm we calculate the composite value for each treatment arm and then the z-statistics to compare persistence between vaccine and control arms.

Imputation is performed separately for each treatment group. In practice this specifically means P0, the probability of persistence following a single positive (described more below) is calculated separately for each treatment group as the other imputation steps are unaffected by whether they are performed for each group separately or overall.

The sequence of imputing is

1. To identify participants with uncertain persistence for each HPV type in turn. This is elaborated below as individuals with a single type-specific positive with no preceding or succeeding positive that would indicate persistence (i.e. ≥5.5m, >166 days apart) or 2 type-specific positives less than 167 days apart). There may be a later type-specific negative test.
2. For each participant with uncertain persistence for an HPV type we assign an imputation probability i.e. a probability of persistence based on if there are one or two positive tests, the period between the two

positives if there are two and the timing of any later negative relative to the positive test/s (described in more detail below).

3. We simulate 10 imputed data sets. We obtain each data set by simulating, for each participant with uncertain persistence for a particular HPV type, a binary variable, using the imputation probability for that HPV type in that participant, that will assign to them an imputed persistence status i.e. either (imputed) persistent or not persistent with the location of the (imputed) persistent sequence starting at the first (or only) positive test denoting the start of the uncertain persistence.
4. For those with an imputed persistence we assign an imputed persistence class of incident, recurrent or prevalent depending on their type-specific HPV result at baseline and whether there was a post-baseline type-specific negative prior to the timepoints relating to the imputed persistence.
5. For individuals with an imputed persistence but who also have a certain persistence class based on actual HPV results (perhaps for a different type or perhaps for the same type but earlier or later in the follow-up and so preceding the later uncertain persistence situation) we will use the actual HPV persistence class and type in the imputed datasets.

So in the initial step for each HPV type we consider participants who have otherwise no other determination of persistence of any type but with either a single positive test (i.e. no immediately preceding positive test of the same type) of some HPV type or 2 positive HPV tests less than 5.5 months apart without an intervening negative test and, in each case, either A) no subsequent test at all or B) a negative test, or several, after the 5.5 months. There is no subsequent positive test as that would indicate persistence i.e. there would be persistence (not uncertain persistence) if there was a 3rd or 2nd positive test more than 5.5 months from the first (or only) test without an intervening negative test.
For example, if those visits are the previously described “18” and “24” month samples with <5.5 months (< 167 days) between them, then the intended repeat 30-month sample did not occur or was negative. They therefore either have no later tests or a 3rd test that is negative more than 6 months from the first (or only) test leading to unclear information on persistence at 5.5 months that will be imputed by the following scheme.

We consider separately the two cases where there are two positive tests within 5.5 months (or just one positive test with no preceding positive test of the same type) as A) no subsequent test after the 5.5 months or B) a negative test more than 6 months from the initial or only positive HPV test (a 3rd case of a subsequent positive test would indicate persistence).

- Case A: The first case is where there is one post-baseline positive test or two positive tests within 5.5 months (<167 days) and, in each case, no subsequent negative test.

If there are two positive tests (within 5.5 months and no subsequent negative) we define them as being at times T0 and TP then if (TP-T0)/5.5 = x where TP-T0 is the difference between the 2 positive tests (in months i.e. the difference in days multiplied by 12/365) we will estimate the probability of persistence at 6 months as P= x + (1-x)P0 (P0 defined below) and use this probability P to impute these uncertain cases. The special case of a single positive (and no other positive within 5.5 months and no subsequent negative) is treated with the same formula with x=0. In this case P is P0 the probability of persistence given a single positive described below. In addition, if the last test is a single positive and it occurs at month 30 then this is not considered an occurrence of uncertain persistence and no imputation is done as we are only concerned with possible persistence situations occurring in the time scale of the trial up to 30 months.

We define the probability of persistence P following a single positive test as P0 (except for first positive tests at 30m) and estimate by considering all occurrences of a positive test, no other tests within 5.5 months and a single test between 5.5 and 8 months as P equal to P0 i.e. the proportion of such occurrences where that next non-missing test between 5.5 and 8 months (>166 days and <244 days) is positive for the same type.
This quantity P0 is calculated on a by arm and by HPV type basis although due to data sparsity we will calculate it separately for type 16 and type 18 (each of those 2 types individually) and then a pooled estimate for all other types i.e. the estimate of P0 for non-16/18 types is the proportion of the total number of single positives that have a subsequent positive test.

- Case B: In the second case there is (a in Case B) one positive test or two positive tests within 5.5 months (< 167 days) but there is also a negative test at TN (TN>5.5months from the first or only positive test giving uncertain persistence at 6 months itself). Again we consider the first positive test to have occurred at time T0 and the potential second positive test at TP (TP < 6 months). If there is not a 2nd positive test, we take TP=T0.

  If a patient has 2 occurrences of uncertainty for the same type then they will be imputed based only on the second occurrence i.e. each patient has a single uncertain status.

  We calculate the following quantities taking values between 0 and 1

  x2 = exp(-(6-(TP-T0))) : increases from 0.0025 when TP-T0 is close to 0 to 1 when TP-T0 is close to 6 (this is unconnected to x=(TP-T0)/5.5 above)
  y2 = exp(6-(TN-T0)) : decreases from 1 when TN-T0 is close to 6 to 0.0025 when TN-t) is close to 12

a = 10% * (TP-T0+TN-T0)/12, large values favour a high imputation probability as they indicate that TP-T0 is high (i.e. close to 5.5 so nearly enough to indicate persistence) and/or TN-T0 is high (i.e. close to 12 months past the initial or only positive – or even higher if tests for several successive timepoints are missed before that negative - and a late negative test therefore provides low evidence that the infection was already cleared by 6 months). The range of a is from around 5% or 6% (when TP is close to T0 and TN close to 6)

b = x2/(x2+y2): varies from 0 to 100% and gives the relative nearness to 6 of TP over TN

w = min(max(0,max(TP-T0,12-(TN-T0))-2.5),2)/2
i.e. the distance between the closest to 6 of TP and TN and either the 2.5 month or 9.5 month time point depending if its TP or TN that is closest (if both are more than 3.5 months from 6 then zero is taken as the distance). This distance is rescaled between 0 and 1 with distances above 2 being assigned distance 1.
So, w indicates how close to 6 the nearest of TN and TP are. If either TP or TN are between 4.5 months or 7.5 months from T0 then the distance is 1. If TP or TN are before 2.5 months and after 9.5 months from T0 respectively then the distance is 0.

The imputation probability is obtained by weighting a and b since large values of both indicate a higher chance of positivity at 6 months. The weighting is based on w:

$$P = w * b + (1-w) * a$$

So if either TP or TN are near 6 months from T0 then w is high and the imputed probability is close to b (the relative nearness to 6 of TP over TN). If either TP or TN are between 4.5 months or 7.5 months from T0 then w = 1 and P=b
Alternatively, if both TP or TN are near 12 then w is 0 or low and the imputed probability is close to a.

The diagram below illustrates the definition of T0, TP and TN as the initial positive at T0 and potential later positive tests at TP and TN and their position relative to 5.5 months.

Figure 3. Timing of Positive and Negative tests relative to 5.5 months in uncertain persistence situations

T0 *TP* *5.5* *6* *TN*

Analysis of Imputed Data sets and pooling via Rubin's rules.

After estimation of P for each case of uncertain persistence, case A or case B, 10 multiply imputed data sets will be produced each with persistence imputed for each occurrence of uncertain persistence due to a single positive test or 2 positive tests < 5.5 months without a later positive test apart as a randomly generated Bernoulli variable using the relevant P from above procedure for each participant. For each imputed data set the same calculation of the endpoint will be performed as for the unimputed case with the 7 oncogenic vaccine types except there is no stratification by country. For each participant I/R/P persistence will be calculated for each HPV type and the worst-case class of persistence (I>R>P) assigned to each participant allowing proportions of I, P and R to be calculated for each worst class of persistence.
The m=10 imputed data sets will be combined using Rubin's rules where for each site a pooled estimate of Z_j is calculated for each imputed data set j=1-10.
We calculate the mean (over imputations) of the calculated difference in the values of the composite in each imputation j=1,2,3 … 10: $D_j = C_1 - C_2$

$$\overline{D} = \sum_{j=1}^{m=10} D_j$$

Then the between imputation variance of the composite is

$$V_B = \frac{\sum_{j=1}^{m=10}(D_j - \overline{D})^2}{m-1}$$

For the within imputation variance, we find the mean of the variances of the composite for imputed data set j i.e.

Var_j =
$6^2 I_1 (1-I_1)+P_1 (1-P_1)+3^2 R_1 (1-R_1)) / n_1 +$
$(6^2 I_2 (1-I_2)+P_2 (1-P_2)+3^2 R_2 (1-R_2))/ n_2$

where P_1, R_1, I_1 are the proportions of the composite P, R, I components in the control arm and P_2, R_2, I_2 for the vaccine arm in each imputed data set.

So, the within imputation variance for imputed data set j is

$$\left(V_W = \sum_{j=1}^{m=10} \boldsymbol{Var_j}\right)/\boldsymbol{m}$$

The variance of the pooled estimate of the composite is

$$V_T = V_W + \left(1 + \frac{1}{m}\right) V_B$$

For the imputed analysis $\overline{D}/\sqrt{V_T}$ will be compared to a standard z statistic with 2-sided significance obtained at the 5% level i.e. absolute value above 1.96.

Per protocol Analysis
A per protocol analysis will compare the composite calculated in 3.1.1 but, in the vaccine arm, only those who take at least 1 dose of the vaccine will be included (descriptive analysis will compare the levels of the composite for those with 1, 2 or 3 vaccine does as a mediator).

3.1.3 Sensitivity Analysis of primary outcome treated with 1) 5 months for identifying persistence and 2) as binary persistence

- A sensitivity analysis will repeat the main primary analysis with the interval between tests required for persistence being 5 months (not 5.5 months) i.e. >151 days.
- A sensitivity analysis will assess the difference in the effect of treatment group on persistence treated as a single binary endpoint ignoring incident, recurrent and prevalent status. The analysis is a single analysis for of persistence for any HPV type. Persistence itself can be identified for individuals with or without a baseline test so this secondary analysis will include all ITT individuals with a post-baseline test result.
  The comparison of persistence between arms will be performed using a Fisher exact test.

3.1.4 Analysis of secondary outcomes

- The separate worst case (again according to with Incident being higher than Recurrent which is higher than Prevalent) from the 7 oncogenic vaccine type HPV outcomes in the composite score i.e. proportions incident, proportions prevalent and proportions recurrent will be individually compared by Fisher exact test between treatment groups using a two-sided alpha level of 0.05 adjusted for multiple testing by Bonferroni correction i.e. with significance being 0.05/3. The endpoint is occurrence of the component of the composite e.g. Incident versus no persistence with those participants with (worse-case) of Prevalent or Recurrent being excluded when the Incident endpoint is considered as an individual analysis.
- We shall repeat the analysis of 3.1.1 comparing the composite measure P+6I+3R where P, I and R are the proportions of patients with prevalent, incident and recurrent respectively for any of the 14 oncogenic HPV-type infections, types 16/18/31/33/35/39/45/51/52/56/58/59/66 and 68, not just the 7 oncogenic vaccine types

- We will also examine any HPV infection occurring from any post-baseline test (months 6 to 30),i.e. not necessarily persistent but based on at least 1 post treatment positive test with any of the 7 oncogenic vaccine HPV types 16/18/31/33/45/52/58 comparing the composite measure P+6I+3R where P, I and R are the proportions of patients with prevalent, incident and recurrent respectively.
- We will also examine any HPV infection, i.e. not necessarily persistent but based on at least 1 post treatment positive test with any of the 14 oncogenic HPV-type infections, types 16/18/31/33/35/39/45/51/52/56/58/59/66 and 68 comparing the composite measure P+6I+3R where P, I and R are the proportions of patients with prevalent, incident and recurrent respectively
-

In each case infection is defined as.

  - Prevalent as a single positive HPV test in the first test post baseline (>121 days post randomisation) of the same type as a positive HPV test at the baseline timepoint (any prevalent infection) i.e. there is no type-specific negative test preceding the first positive test.
  - Incident as a positive test for any HPV type post-baseline (>121 days post randomisation) that follows a negative baseline test or is different to any type found at the baseline timepoint if positive. Recurrent as a positive test at the baseline timepoint and at a post baseline (>121 days post randomisation) timepoint for the same HPV type but with the post-baseline positive test following a negative for that type.
  - As with the persistent case a participant can experience both a prevalent and a recurrent infection (not necessarily persistent) for the same HPV type. This would happen if, following a type-specific positive baseline test, the first recorded post-baseline test was positive meeting the definition of a prevalent infection and then following a later type-specific negative test an infection of the same type recurred. In this case the participant would be considered to have a recurrent infection. For example, where the recurrence occurs at month 24 after two negatives: Baseline=POS, 6 months = POS, 12 months = NEG, 18 months = NEG, 24 months = POS. This will still be considered a recurrent infection even if the earlier prevalent infection was persistent.

- Using all randomised participants (including those with missing baseline HPV tests or missing post-baseline tests) the proportions of CIN2+ lesions (HSIL) recorded at a post-baseline visit, i.e. by 24 months, will be compared by Fisher exact test between treatment groups using a two-sided alpha level of 0.05 for those participants:
    i) Who are diagnosed with a persistent infection of a vaccine HPV type 16/18/31/33/45/52/58.

ii) who are diagnosed with a persistent infection of any oncogenic HPV type 16/18/31/33/35/39/45/51/52/56/58/59/66 and 68.

The occurrence of CIN2 is based on there being some occurrence of CIN2+ at a post-baseline assessment i.e. if the worst-case CIN (if multiple assessments) is CIN2/CIN3, SMILE, high grade CGIN (AIS) or cancer. This CIN is considered simultaneous with an HPV type if the previous non-missing HPV result prior to any CIN2 assessment (including potentially the baseline HPV result) is one of the i) 7 vaccine HPV types ii) one of the 14 oncogenic types. Those who have no assessment for CIN will be considered CIN negative (and therefore CIN2+ negative). Those who have an occurrence of CIN not associated with an HPV type of the set under consideration will also be considered negative for this HPV associated CIN endpoint.

- Safety Analyses
  Occurrence of reportable AEs and SAEs will be reported. Safety data analysis will be conducted on all patients receiving at least 1 dose of Gardasil®9. Analyses will consist of data summaries for clinical parameters, and for AEs. The number and percentage of patients experiencing 1 or more AEs will be summarised by the relationship to study vaccine and severity. AEs will be coded using Medical Dictionary for Regulatory Activities (MedDRA) terminology. All available safety data will be provided to an IDMC at the interim and final efficacy analyses (see below). Periodic safety reviews will be conducted by the IDMC as described in the IDMC Charter.

3.1.5 Analysis of tertiary outcomes

- The proportions of CIN1+ lesions recorded at a post-baseline visit will be compared by Fisher exact test between treatment groups using a two-sided alpha level of 0.05 for those participants:
  i) Who are diagnosed with a persistent infection of a vaccine HPV type 16/18/31/33/45/52/58.
  ii) Who are diagnosed with a persistent infection of any oncogenic HPV type. 16/18/31/33/35/39/45/51/52/56/58/59/66 and 68

  Assessment of CIN1+ and HPV involvement will be performed as with the CIN2+ analysis of the secondary endpoint except that CIN1 and also low-grade CIN is also included as CIN1+ i.e. if the worst-case CIN (if multiple assessments) is CIN1/CIN2/CIN3, SMILE, low grade or high grade CGIN (AIS) or cancer

3.1.6 Statistical considerations

Time points
Randomisation and visit 1 with initial HPV test is at 0 months. Visit 2 at 2 months. Timepoints (visits or self-tests) 3-7 with HPV tests occur at

6,12,18,24 and 30 months. Liquid based cytology (LBC) tests for CIN2+ occurrence are at 6 and (Finland only) 24 months.

Stratification and clustering
The main primary analysis and composite based -secondary analyses are stratified by country. There are also subgroup analyses described in 3.1.8 and moderator and mediator analyses described in 3.3.

Missing items in scales and subscales

The number (%) with data of each outcome including positive and negative HPV results and CIN status at each timepoint will be reported.

Missing baseline data

Missing data should not be an issue for the primary analysis.

Missing outcome data

Where there are missing post-baseline/randomisation HPV assessments the primary analysis above is based on the available HPV test results. As described above a secondary analysis is based on imputation for situations with uncertain persistence. This assumes the missing HPV test results are missing at random (MAR). Individuals with no post-baseline HPV assessments are excluded from all analyses. If post treatment variables such as compliance with treatment are found to be predictive of drop out, multiple imputation will be considered.

Method for handling multiple comparisons
There are no multiple comparisons for the main analysis. For the secondary analysis of individual incident, recurrent and prevalent components of the composite the Bonferroni correction will be applied with a significant p-value being less than 0.05/3.

Method for handling non-compliance (per protocol/CACE analyses)
In addition to the primary intention-to-treat analysis the effect of actually receiving treatment as defined in the protocol will also be estimated.

Model assumption checks
The models assume normally distributed outcome of the composite; this will be checked when describing the data and if substantial departures from normality occur, transformations will be considered.

Analysis timings
We will carry out the main analysis of study data once all participants have reached the 24-month timepoint (main dataset). Once the remaining eligible participants have completed their 30-month (updated dataset) follow up, a second data lock will take place and an updated analysis of study data will take place.

3.1.7 Sensitivity analyses

Principal Stratum Sensitivity Analysis:
There will be an analysis of efficacy of the composite score using the per cohort of all evaluable (per protocol) patients (i.e. those meeting all eligibility criteria, complying with the procedure defined in the protocol in receiving at least 1, 2 or all 3 vaccine doses, with no elimination criteria during the study) for whom data concerning efficacy endpoint measures are available.
This will include patients for whom assay results are available for HPV DNA.

Sensitivity analysis of the definition of persistence between successive HPV test visits will be performed by re-calculating the primary endpoints using a definition of persistence defined as at least 5 months (152 days) and at least 6 months (182 days) between tests (the main analysis has persistence defined as at least 5.5 months between tests).
.

3.1.8 Planned subgroup analyses

In secondary endpoint 1 we plan to perform a subgroup analysis to explore the efficacy of vaccination against (worst case) prevalent, incident and recurrent infections separately as 3 separate binary outcomes where each persistence class is compared to the non-persistent group by the Fisher Exact test.
This will allow investigation of the subgroup that benefits the most from vaccination.
We will further perform analysis in subgroups of women according to, grade of CIN treated and HPV type (3 comparisons of individual occurrence of any persistence, not necessarily worse type, of each incident, recurrent or prevalent type versus no-persistence). We will further perform subgroup analyses comparing the primary value of the composite according to age (<35, 35-44, >44 years of age at randomisation).

3.2 Exploratory analyses

The preceding subgroup analyses according to resection margins, grade of CIN treated, HPV type and age are exploratory.

3.3 Exploratory mediator and moderator analysis

- Mediator of treatment

1. We shall examine prevalence of incident, prevalent and recurrent HPV according to local cervical treatment results including age, depth of excision, incomplete excision/ involved margins and HPV type, type 16 vs other, at baseline (not for incident) or at follow-up.
2. We will tabulate proportions of worst case (over type) of incident, prevalent and recurrent HPV against age (<35, 35-44, >44 years at randomisation), binary occurrence of incomplete excision, involved margins, mean excision depth, and baseline HPV type (HPV 16 vs other types combined).The mean, SD and quartiles of the composite

endpoint will be explored for incomplete excision and incomplete margins, against excision depth (relative to its median cut-off value) and by follow up HPV type.

- Moderators of treatment
  1. We shall examine prevalence of incident, prevalent and recurrent HPV according to local cervical treatment results including age, depth of excision, incomplete excision/ involved margins and HPV type, type 16 vs other, at baseline (not for incident). We will tabulate proportions of worst case (over type) of incident, prevalent and recurrent HPV against age (<35, 35-44, >44 years at randomisation), binary occurrence of incomplete excision, involved margins, mean excision depth, and baseline HPV type (HPV 16 vs other types combined).
  2. The mean, SD and quartiles of the composite endpoint will be explored for incomplete excision and incomplete margins, against excision depth (relative to its median cut-off value) and by baseline HPV type.

3.4 Interim analysis

Periodic safety reviews will be conducted by the IDMC as described in the IDMC Charter. There are no planned periodic efficacy analyses.

4. Software

Data management: The Investigator (or delegated member of the site study team) must record all data relating to protocol procedures, IMP administration, laboratory data, safety data and efficacy data into the trial InForm electronic data collection (EDC) system. The database was later transferred to the OpenClinica data base system.
Data will be uploaded to the QMUL CPTU Data Safe Haven for analysis at 2 time points each corresponding to a data lock. This will be done by email transfer of password protected data and via the AIMES/ARO airlock.
We will carry out an initial main analysis of study data once all participants have reached the 24-month timepoint (main dataset). Once the remaining eligible participants have completed their 30-month (updated dataset) follow up, a second data lock will take place and an updated analysis of study data will take place.

Analysis software: Data analysis will be performed using Stata or R within the QMUL CPTU Data Safe Haven.

C) SCHEDULE OF ASSESSMENTS AND MEASURES

Amendments to version 1.0

LIST HERE ANY AMENDMENTS TO THE SAP THAT WERE MADE AFTER THE SAP WAS SIGNED OFF BYTHE TSC

This is version 1.0

Reference list

(1) Moher D, Schulz KF, Altman DG. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 2001 Apr 14;357(9263):1191-4.

(2) Söderlund-Strand, A., Kjellberg, L., & Dillner, J. (2014). Human papillomavirus type-specific persistence and recurrence after treatment for cervical dysplasia. *Journal of medical virology*, *86*(4), 634-641.