HIPvac: a trial of vaccination and cream treatment in patients with anogenital warts

Human papillomavirus infection: a randomised controlled trial of Imiquimod cream (5%) versus Podophyllotoxin cream (0.15%), in combination with quadrivalent human papillomavirus or control vaccination in the treatment and prevention of recurrence of anogenital warts

UCL CCTU ID:	**CCTU/2012/009**
Protocol version:	**6.0 dated 04 May 2017**
International Standard RCT Number (ISRCTN):	**68222801**
European Clinical Trials Database Number (EudraCT):	**2013-002951-14**
MHRA Clinical Trials Authorisation Number (CTA):	**20363/0338/001-0001**
Research Ethics Committee Number:	**13.SC.0638**

STATISTICAL ANALYSIS PLAN (SAP)

VERSION 1.3, 17 NOVEMBER 2017

Authorised by:

	Signature	Date
Dr Richard Gilson, Chief Investigator Email: r.gilson@ucl.ac.uk Tel: 020 3108 2103	 24th January 2018	
Prof Caroline Doré, Statistical Oversight Email: Caroline.Dore@ucl.ac.uk Tel: 020 7907 4673	 24th January 2018	

Prepared by:
Kate Bennett, Trial Statistician
Email:
k.bennett@ucl.ac.uk
Tel: 020 7907 4667

24th January 2018

Table of contents

1 ABBREVIATIONS

Acronyms	Meaning
AUC	Area Under the Curve
CI	Confidence Interval
CRF	Case Report Form
HR	Hazard Ratio
IRR	Incidence Rate Ratio
ITT	Intention-To-Treat
MAR	Missing At Random
MNAR	Missing Not At Random
MI	Multiple Imputation
MICE	Multiple Imputation by Change Equations
OR	Odds Ratio
qHPV	quadrivalent Human Papilloma Virus
QoL	Quality of Life
SAE	Serious Adverse Event
SAP	Statistical Analysis Plan
eSMF	Statistical Master File stored electronically
TMF	Trial Master File

2 ABSTRACT – BACKGROUND AND DESIGN

Aim and objectives: The aim of this study is to assess the clinical effectiveness and cost-effectiveness of topical therapy with imiquimod or podophyllotoxin in clearing anogenital warts, and of quadrivalent human papillomavirus (qHPV) vaccination to provide additional benefits in terms of either higher clearance or reduced recurrence rates.

The primary objectives of the trial are:

1. To compare the effectiveness of imiquimod 5% cream versus podophyllotoxin 0.15% cream in the treatment of external anogenital warts. The primary objective will be to compare the proportions of participants receiving each treatment who have complete resolution of warts at 16 weeks and remain free of warts up to 48 weeks after starting treatment.
2. To compare the effectiveness of a course of quadrivalent HPV vaccine started at the same time as topical wart treatment with saline placebo, in improving wart clearance at 16 weeks and preventing recurrence up to 48 weeks.
3. To estimate the cost-effectiveness of the two topical treatments, taking into account treatment, staff and other healthcare costs of initial and recurrent warts, and reduction in participants' quality of life due to warts.
4. To estimate the cost-effectiveness of a course of quadrivalent HPV vaccine compared with placebo control, taking into account treatment, staff and other healthcare costs of initial and recurrent warts, and reduction in participants' quality of life due to warts.
5. Given that the cost of imiquimod is currently higher than podophyllotoxin but likely to decrease in future, the price at which it is likely to become cost-effective in comparison to podophyllotoxin will be estimated. If the HPV vaccine is more effective than control vaccine, an additional economic evaluation will be carried out to estimate the price at which it will become cost-effective.

Secondary objectives:

6. To compare wart clearance rate at interim time points corresponding to the end of the prescribed treatment course.
7. To compare the time to wart clearance in those treated with podophyllotoxin versus imiquimod.
8. To compare the proportion experiencing wart recurrence/relapse (after wart clearance) at 48 weeks.
9. To compare the tolerability of all treatments as measured by reported local and systemic reactions and other adverse events, and adherence to treatment.
10. To compare health-related quality of life, as measured by the Area Under the Curve for EQ-5D.
11. To compare the requirements for additional therapy, including extension of the initial topical treatment course, treatment with cryotherapy, or recourse to other agents.
12. To collect and store blood samples (at 0 and 48 weeks for serum; at 0, 4, 8, 16 and 48 weeks for peripheral blood mononuclear cells [PBMC]) and swab samples from genital wart lesions (at baseline and in the event of recurrence) for laboratory sub-studies, including comparison of HPV types at recurrent disease with initial lesions, and anti-HPV antibody and cell-mediated immune responses. Separate funding will be sought for this work which would be further defined according to the outcome of the main trial. Samples will be collected at specific HIPvac trial sites only, where resource is sufficient to enable correct sampling and storage processes.

Population studied: 500 patients presenting with external anogenital warts, aged 18 years or over, males and females, with either a first or subsequent episode of anogenital warts which, in the opinion of the investigator, could be appropriately treated with either self-administered imiquimod or podophyllotoxin creams.

Trial design: HIPvac is a randomised, controlled partially blinded 2 x 2 factorial design trial of the treatment of anogenital warts, with an accompanying economic analysis.

Patients allocated to imiquimod will receive a 16 weeks treatment course. Patients allocated to podophyllotoxin will received a 4-week treatment course, which will be extended if there is a partial response to therapy. All patients will receive a qHPV vaccine or saline placebo injection at months 0, 2 and 6.

Sample size: The trial was originally designed with a sample size of 1000 participants with equal numbers randomised to each of the two topical cream arms and each of the two vaccine groups in a 2x2 factorial design, so that allowing for 20% loss to follow-up 800 participants will contribute primary outcome data. The anticipated proportion achieving the primary endpoint in the less favourable topical treatment group is 35%, assuming a wart clearance rate of 50% within 16 weeks and a 30% subsequent recurrence rate. This sample size provided 80% power (at the 5% significance level) to detect an increase to 45% with the better

treatment. It also provided 80% power to detect an increase from 35% to 45% in the primary endpoint from vaccination, as would arise if vaccination reduces the recurrence rate from 30% to 10% whilst leaving the wart clearance rate unchanged at 50%.
Owing to lack of feasibility to achieve the proposed recruitment target of 1000, a revised sample size of 500 participants was proposed in February 2016. With 15% loss to follow-up, this would provide 52% power (at the 5% significance level) to detect the pre-specified difference in the combined primary endpoint.

Randomisation: Participants will be allocated in equal proportions to the two topical cream arms and the two vaccine arms, creating four groups using minimisation with a random element, with gender (male vs. female), previous occurrences of warts (no previous occurrences vs. one or more previous occurrences), HIV-status and site as stratification factors. HIV-status has been included as a stratification factor since the decision to include HIV-positive participants in the trial was taken in December 2015.

Blinding: Placebo injection syringes will **not** be identical to qHPV vaccine syringes. To ensure blinding is maintained as far as possible, each site will ideally have an unblinded team member who will administer the vaccine. The unblinded team member will not have any other involvement in the treatment or assessment of HIPvac trial participants.
The packaging of the vaccine/placebo (a sealed carton and an inner plastic pouch) will be fully blinded, to prevent inadvertent unblinding of other members of the site trial team or the participant at any stage. This labelling strategy will also ensure that the unblinding of one participant will not unblind the entire trial arm.
The podophyllotoxin vs. imiquimod portion of the trial will be open-label.

3 OUTCOME MEASURES

3.1 Primary outcome

The primary outcome is a composite endpoint of wart clearance 16 weeks after starting treatment and remaining wart-free between 16 and 48 weeks. This captures both the initial clearance efficacy as well as the impact on relapse or recurrence.

3.2 Clinically important factor-specific outcomes

The two components of the composite primary end-point will be assessed as clinically important outcomes for each factor:

1. The clinically important outcome for the topical treatment factor is the proportion of patients wart-free at 16 weeks.

2. The clinically important outcome for the vaccine treatment factor, is the proportion experiencing complete wart clearance at 48 weeks, in those with wart clearance at 16 weeks.

3.3 Secondary outcomes

3.3.1 Clinical

1. Proportion wart-free at the end of the assigned treatment course (4 or 16 weeks)
2. Proportion wart-free at the end of the assigned treatment course (4 or 16 weeks) without receiving additional treatment
3. Quantity of additional treatment (number of cryotherapy applications, additional weeks of podophyllotoxin or imiquimod) required to achieve clearance by 16 weeks
4. Proportion wart-free at 16 weeks without receiving additional treatment
5. Proportion experiencing complete wart clearance
6. Proportion experiencing wart recurrence/relapse after complete wart clearance
7. Time to complete wart clearance
8. Time from complete wart clearance to recurrence/relapse

3.3.2 Safety

1. Proportion of patients reporting at least one Adverse Event (AE) throughout the trial.
2. Proportion of patients reporting at least one Serious Adverse Event (SAE) throughout the trial.
3. Severity of most severe side effects measured by a self-rated 5-point Likert scale (symptom scores) at 4, 8, 16, 24 and 48 weeks

3.3.3 Health Economics

1. Health-related quality of life, as measured by the Area Under the Curve for EQ-5D-5L
2. Additional outcomes will be collected for a detailed cost and cost-effectiveness analysis; however this analysis will not be performed by the Trial statistician so these outcomes are not described here. Further details on the cost effectiveness analysis can be found in the Health Economic Analysis Plan.

3.4 Scoring the EQ-5D-5L

The EQ-5D-5L consists of a self-reported matrix comprising 5 items or dimensions (i.e., mobility, self-care, usual activities, pain-discomfort and anxiety-depression) rated on 5-point scales ranging from 0 to 4 and a self-rated health state 100mm visual analogue scale (VAS). Respondents' ratings can be combined into a single health utility score (see below).

Scoring: Value sets based on preferences directly elicited from representative general population samples to derive the EQ-5D-5L health utility score have recently been made

available [10]. The EQ-5D-5L value set has not yet been verified by NICE but has been compared to both the EQ-5D-3L and the EQ-5D-5L Crosswalk Index Value Calculator (CIVC)2, and a "crosswalk" between the EQ-5D-3L value sets and the new EQ-5D-5L descriptive system derived as an interim measure [11].

Missing items: If the instrument has one or more dimension scores missing, the EQ-5D-5L index score will be set to missing.

4 DATA

4.1 CRF and variables

Full details of data collection and timing are described in the trial protocol (6.0 dated 04 May 2017). Copies of CRFs are included in the Trial Master File (TMF).

4.2 Management of datasets

At the time of analysis:

- A copy of each dataset will be prepared by the Trial or Delegated Statistician (frozen dataset) and saved in section *3. Analysis* of the Statistical Master File stored electronically (eSMF).
- If necessary, data can be added to or amended in the main, unfrozen copy of the dataset.
- If any outstanding queries are resolved during the analysis that relate to data in the frozen dataset (e.g. problems that are found during analysis or amended CRFs that are return to CCTU), the main and frozen dataset should both be altered.
- If any outstanding data queries are resolved while the analysis files are being prepared (when only a practice dataset has so far been copied), the changes need only be made to the main dataset and an updated frozen copy made available in section 3 of the eSMF.

4.3 Data completion and schedule

The last patient for the HipVAC trial was randomised on 06 January 2017. All forms for 48-week follow-up visits should therefore be available by 30 November 2017.

4.4 Data verification

Data verification, consistency and range checks are performed during data entry, as well as checks for missing data (copies of these checks can be found in the TMF). Additional range, consistency and missing data checks will be performed when the datasets for analysis are constructed, as appropriate, before the statistical analysis is performed. All variables will be examined for unusual, outlying, unlabelled or inconsistent values.

Any problems with trial data will be queried with the Trial Manager or Data Manager as appropriate. If possible, data queries will be resolved; although it is accepted that due to administrative reasons and data availability a small number of problems will continue to exist. These will be minimised.

4.5 Data coding

Details of the variables, including variable coding lists are included in the metadata which forms part of the TMF.

5 SAMPLE SIZE ESTIMATION

5.1 Primary outcome

The two main objectives of the HIPvac trial are to compare the efficacy of the two topical treatments for wart clearance and to assess the efficacy of quadrivalent human papillomavirus vaccination on wart recurrence. The trial was originally designed with a target sample size of 1000 participants with equal numbers randomised to each of the two topical cream arms and each of the two vaccine groups in a 2x2 factorial design, so that allowing for 20% loss to follow-up 800 participants would contribute primary outcome data. The anticipated proportion achieving the primary endpoint in the less favourable topical treatment group was 35%, assuming a wart clearance rate of 50% within 16 weeks and a 30% subsequent recurrence rate. This sample size provided 80% power (at the 5% significance level) to detect an increase to 45% with the better treatment. It also provided 80% power to detect an increase from 35% to 45% in the primary endpoint from vaccination, as would arise if vaccination reduces the recurrence rate from 30% to 10% whilst leaving the wart clearance rate unchanged at 50%.

Owing to slow recruitment and the infeasibility of achieving the original recruitment target of 1000 participants within the funding available, a revised sample size of 500 participants was agreed in January 2017. A sample size of 500 patients assuming a 15% loss to follow-up, will provide only 52% power at the 5% significance level to detect the pre-specified clinically important difference in the composite primary end-point.

However, 500 patients will provide 80% power at the 5% significance level to evaluate clinically relevant differences for each of the two clinically important components of the composite primary outcome:

i. Proportion wart-free at 16 weeks with or without cryotherapy
ii. Proportion remaining wart-free at 48 weeks among participants wart-free at 16 weeks with or without cryotherapy

For the 16-week topical treatment outcome, a difference of 14% between topical arms (57% versus 43%) could be detected. For the 48-week vaccine outcome, a difference of 16% between vaccine arms (12% versus 28%) could be detected. These differences would be considered to be clinically relevant and may influence management guidelines. The 5% significance level has been used for both calculations as there is a different outcome for each of the two factors to answer two independent questions. It is expected that the primary effect of the topical treatment will be on initial wart clearance rates, while the vaccine is expected to act primarily to reduce wart recurrence rates.

5.2 Secondary outcomes

The trial is not powered to detect differences between the treatment groups for any of the other secondary outcomes.

6 ANALYSIS PRINCIPLES

6.1 Intention-to-treat (ITT) or per-protocol?

All analyses will be conducted on a modified intention-to-treat (mITT) basis. We will include all consented randomised patients for whom at least one follow-up visit is available regardless of their adherence to treatment as HIPvac is a pragmatic study concerned with the effectiveness of topical therapy and qHPV vaccination.

6.2 Significance level of tests

All confidence intervals will be 95% and two-sided. Statistical tests will use a two-sided *p* value of 0.05, unless otherwise specified. There will be no formal adjustment of *p* values for any interim analyses performed.

6.3 Baseline comparability

Baseline characteristics will be summarised by randomised group using the four groups formed from combinations of the two treatments.

6.4 Adjustment for design factors

We will adjust the analyses for the stratification variables HIV-status, gender, and whether the patient had a previous wart recurrence (as recommended in ICH E9, 5.7) by including them as fixed effect covariates. HIV status was introduced as a stratification factor in December 2015 following a protocol amendment to change the entry criteria in to allow the inclusion of HIV-positive participants to the trial. Patients randomised before the addition of HIV status as a stratification factor are known to be HIV negative and will be coded as such for analysis purposes. Site will be included in the mixed effects models as a random effect (random intercept) to account for any possible variation by site. Treatment effects are then estimated conditional on HIV status, gender, and previous occurrence of warts and account for variation between sites.

Adjustment for design factors will not be made for binary safety secondary outcomes since there are likely to be too few events to fit logistic regression models.

6.5 Follow-up and losses to follow-up: handling missing data

In a pragmatic clinical trial over a three-year time frame, some patients are inevitably lost to follow-up. Data for such patients will therefore be only partially observed. This can lead to a loss of power, biased estimates and standard errors, and a loss of efficiency. In order to produce unbiased results, and in order to maximise the power to detect a treatment effect, multiple imputation using chained equations [MICE] will be used to impute data from missing follow-up visits and missing data items from baseline and follow-up visits.

Missing data will be assumed to be missing at random (MAR) conditional on all variables included in the imputation model (see section 7.9) and so independent of the values of the unobserved data themselves. Multiple imputation (MI) using chained equations [1] will be used to impute data from missing follow-up visits. Results will be combined using Rubin's rules [7]. Participants who do not attend any follow-up visits will be excluded from the analysis.

The analyses for all primary and secondary outcomes will be performed on fully imputed datasets (see section 7.9).

Sensitivity analyses will investigate the impact of the MAR assumption and missing data for all patients (see section 7.5).

6.6 Summarising models

Wherever possible, analysis of outcomes will involve a parametric model. Treatment effect estimates will be presented as regression coefficients and 95% confidence intervals.

7 ANALYSIS DETAILS

The results of the analyses will be reported following the principles of the ICH E3 guidelines on the Structure and Content of Clinical Study Reports.

7.1 Recruitment and follow-up patterns

Recruitment will be presented by year and site.
The number of CRFs completed – excluding patients who have been withdrawn from therapy and were unwilling to continue follow-up will be reported by treatment group.
The number of patients who have been withdrawn from therapy, were unwilling to continue follow-up or died while on study will be reported by treatment group.

7.2 Baseline Characteristics

Baseline characteristics will be reported for each of the four treatment arms. Summary measures for the baseline characteristics of each group will be presented as mean and standard deviation for continuous (approximately) normally distributed variables, medians and interquartile ranges for non-normally distributed variables, and frequencies and percentages for categorical variables.

7.3 Trial treatment

Adherence to treatment will be summarised by treatment group.

7.4 Analysis Methods

7.4.1 Primary outcome

The primary outcome (the composite endpoint of wart clearance within 16 weeks of starting treatment and remaining wart-free between 16 and 48 weeks) will be analysed using a mixed effect logistic regression model, and will be adjusted for gender, previous occurrence of warts, and HIV-status as stratification factors [4, 5]. Both treatment factors (topical treatment and vaccination) will be include as covariates in the model. Site will be included as a random factor (random intercept). If there are problems with model convergence with a random site intercept, then site will be included in the model as a fixed effect. Treatment effect estimates will be transformed back from their logistic form and reported as adjusted odds ratios (OR) with their corresponding 95% confidence intervals (CI) and two-sided p-values.

7.4.2 Clinically important factor-specific outcomes

The analysis for both factors (podophyllotoxin vs. imiquimod, and qHPV vs. placebo vaccine) will be based on comparisons at the margins of the 2 x 2 table (Table 1), meaning all participants randomised to podophyllotoxin will be compared with all participants randomised to imiquimod, and all participants randomised to qHPV vaccine will be compared with all participants randomised to saline placebo.

We do not anticipate a substantial interaction between topical treatment and vaccination. However, as a secondary analysis, we will perform an interaction test between the two factors, and present results from a four-arm analysis (where each of the four treatment combinations is regarded as a separate treatment arm), as is recommended for factorial trials [2, 3].

The analysis for each of the two components of the composite primary endpoint:

- Proportion wart-free at 16 weeks with or without cryotherapy (clinically important secondary outcome for topical treatment factor)

- Proportion experiencing complete wart clearance at 48 weeks with or without cryotherapy (clinically important secondary outcome for vaccine factor), in those with wart clearance at 16 weeks

will be performed using a logistic regression model adjusting for gender, previous occurrence of warts, and HIV status as stratification factors and will include both treatment factors (i.e. topical treatment or vaccination). Site will be included as a random factor (random intercept) to account for variation by site. If there are problems with convergence in the mixed model, then site will be included alternatively as a fixed effect. All estimated treatment effects will be reported as adjusted odds ratios (OR) with their corresponding 95% CIs.

Table 1: Interventions received according to the 2 x 2 factorial trial design. Randomisation will be 1:1 between the two topical cream arms and 1:1 between the two vaccine arms

		Topical creams	
		Imiquimod	*Podophyllotoxin*
Vaccines	*qHPV vaccine*	Arm A n=125 Imiquimod cream for 16 weeks; qHPV vaccine at months 0, 2 and 6	Arm B n=125 Podophyllotoxin cream for 4 weeks; qHPV vaccine at months 0, 2 and 6
	Saline, placebo control	Arm C n=125 Imiquimod cream for 16 weeks; placebo vaccine at months 0, 2 and 6	Arm D n=125 Podophyllotoxin cream for 4 weeks; placebo vaccine at months 0, 2 and 6.

7.4.3 Other secondary outcomes

As stated in section 5.2 the trial has not been powered to detect differences between the randomised groups for any of the secondary endpoints. Analyses of secondary outcomes for both factors will be based on comparisons at the margins of the 2x2 table (i.e. podophyllotoxin vs. imiquimod, and qHPV vs. placebo vaccine).

a) Effectiveness

Each of the following binary secondary outcomes:

- Proportion wart-free at the end of the assigned treatment course (4 or 16 weeks)
- Proportion wart-free at the end of the assigned treatment course (4 or 16 weeks) without receiving additional treatment
- Proportion wart-free at 16 weeks without receiving additional treatment
- Proportion experiencing complete wart clearance
- Proportion experiencing wart recurrence/relapse after complete wart clearance

will be analysed using a logistic regression model adjusting for gender, previous occurrence of warts, and HIV status as stratification factors. We will include both treatment factors (i.e. topical treatment and vaccination) as covariates in the regression models. Site will be included as a random factor (random intercept) to account for variation by site. If there are problems with model convergence with a random site intercept, then site will be included in the model as a fixed effect. All

estimated treatment effects will be reported as odds ratios (OR) with their corresponding 95% CIs.

Each of the following time-to-event secondary outcomes:

- Time to complete wart clearance
- Time to complete wart clearance without the use of additional treatment (i.e. cryotherapy)

will be analysed using a Cox proportional hazards model adjusting for stratification factors gender, previous occurrence of warts, site, and HIV status to estimate differences between treatment arms.

If warts have not cleared by the end of the trial (i.e. week 48) or the last available assessment date, they will be treated as censored observations.

For those patients who experience complete wart clearance during the trial, we will estimate:

- Differences in time from complete wart clearance to recurrence/relapse between treatment arms
- Differences in time from complete wart clearance to recurrence/relapse between treatment arms without the use of additional treatment (i.e. cryotherapy)

from a Cox proportional hazards model adjusting for gender, previous occurrence of warts, site and HIV status.

Cox proportional hazard models will include both treatment factors (i.e. topical treatment and vaccination) as covariates in the model. Treatment effects will be reported as hazard ratios (HR) with their corresponding 95%CIs. The assumption of proportional hazards will be assessed graphically using log–log plots of the estimated survivor functions and by examining Schoenfeld residuals. Violations of the proportionality assumption for particular covariates in the model will be addressed through inclusion of time–dependent covariates.

The quantity of additional treatment given will be summarised for each of the four treatment groups. We will report means and standard deviations if the data is approximately normally distributed and medians and interquartile ranges if it does not follow a normal distribution. No statistical analyses will be performed as the number of events are unlikely to be sufficient to fit a regression model.

b) **Safety**

Differences between treatment arms for the proportion of patients with any safety events (e.g., AE or SAE) will be analysed using Fisher's exact tests. Summary measures will be the

number (%) of patients with an event in each group. Treatment effects will be estimated by the difference in event rates and 95% CI for the differences.
To compare severity of (most severe) side effects between treatment groups over the follow-up period, we will use ordinal logistic regression using the proportional odds approach in a mixed effect model. This statistical technique considers every possible way in which an ordinal scale can be dichotomized, assuming that the odds ratio for a better outcome versus a worse outcome is identical wherever the scale is dichotomized. The model will be adjusted for stratification factors (i.e. gender, previous occurrence of warts, and HIV status). We will include both treatment factors (i.e. topical treatment and vaccination) and assessment time-point as covariates in the model. Site will be included as a random factor (random intercept) to account for variation by site. If there are problems with model convergence with a random site intercept, then site will be included in the model as a fixed effect. We will obtain cluster-robust standard errors of the regression coefficients, based on an extension of the Huber-White sandwich estimator, which relaxes the assumption of independence of observations and therefore accounts for the observed correlation between repeated assessments of the same patient. Estimated treatment effects will be reported as odds ratios (OR) and 95% CIs. To check the assumption of proportional odds we will perform the Brant test.

7.5 Sensitivity analyses following multiple imputation

The analyses for the primary and secondary outcomes will be conducted in fully imputed datasets (see section 6.5.) and available case analysis for the primary composite endpoint and its two components will also be reported. The primary analyses assume that the data are missing at random (MAR), i.e. the probability that data are missing depends on the values of the observed data but does not depend on the values of the missing data. Therefore any systematic differences between the observed and unobserved values can be explained by differences in observed variables. The reasons for missing data will be explored and will include the following descriptive analyses: the amount of missing data at each time point by treatment arm, missing data patterns, and associations between missingness and baseline values.
The MAR assumption is not one that can be tested, so to explore the validity of the MAR assumption, sensitivity analyses that assume missing not at random (MNAR) mechanisms will be undertaken as detailed below.

Define π_0 as the proportion experiencing complete wart clearance at 16 weeks and remaining wart-free at 48 weeks in the unobserved individuals and π_1 as the proportion experiencing complete wart clearance at 16 weeks and remaining wart free at 48 weeks in the observed individuals. Define θ as the odds ratio comparing the primary outcome

(complete wart clearance at 16 weeks and remaining wart-free at 48 weeks) in the observed individuals compared to the unobserved individuals.

Under the MAR assumption, the odds ratio comparing the primary outcome (complete wart clearance at 16 weeks and remaining wart-free at 48 weeks) in the observed individuals compared to the unobserved individuals is expected to be 1. It may be reasonable however to expect that those individuals who have a good outcome (wart clearance) are less likely to attend for follow-up visits ($\pi_0 > \pi_1$ and therefore $\theta<1$), but $\pi_1 > \pi_0$ and $\theta>1$ is also plausible, although perhaps somewhat less so. We will therefore generate three sets of imputed datasets for the sensitivity analysis, with values of θ equal to 0.6, 0.8 and 1.25 using Stata's mi impute command for a logistic model with an offset. Each of the three sets of imputed data for the three scenarios outlined above (θ=0.6, 0.8, 1.25) will be generated using a logistic imputation model with the offset equal to $\ln(\theta)$ and will be combined using Rubin's rules. The results will be compared with those from the analysis done under the assumption of MAR. Substantive differences would indicate that our findings are not robust to the MAR assumption, and inconsistencies between the complete case analysis, the MAR analysis, and the MNAR analysis will be reviewed by the trial team and the trial steering committee and reported in the final trial report.

7.6 Subgroup analyses

Three planned subgroup analyses will be performed for gender (male vs. female), previous occurrences of warts (no previous occurrences vs. one or more previous occurrences), and HIV-status (HIV-positive vs. HIV-negative). Subgroup analyses will be performed for the primary outcome only, by adding interaction terms to the model for the primary outcome. This will consist of six interaction terms which will be tested separately:

- topical treatment and gender
- topical treatment and previous occurrences of warts
- topical treatment and HIV-status
- vaccination (HPV versus control) and gender
- vaccination (HPV versus control) and previous occurrences of warts
- vaccination (HPV versus control) and HIV status

We will report separate estimates and confidence intervals for each subgroup if a significant interaction is found ($p<0.05$).

7.7 Quality of Life analysis (QoL)

QoL will be measured by the EQ-5D-5L health utilities index and the EQ-5D-5L VAS (Visual Analogue Scale) measured at each time point including baseline, as calculated by the "eq5d" Stata command. We will estimate the area under the curve (AUC), using the trapezoidal rule, to summarise QoL responses across the 48 weeks (i.e. weeks 0, 4, 8, 16, 24 and 48). Analysis of QoL will then be performed on the AUC using a mixed effects linear regression model adjusting for stratification variables (i.e. gender, previous occurrence of warts, and HIV status) and baseline values. Site will be included as a random factor (random intercept) to account for variation by site. If there are problems with model convergence with a random site intercept, then site will be included in the model as a fixed effect. Estimates of treatment difference with 95% CI will be presented.

7.7.1 Missing EQ-5D-5L data

Missing EQ-5D-5L utility index scores from each time point will be imputed using multiple imputation as described in section 7.9.

The AUC will be calculated using a mixture of observed and imputed data and a complete case analysis of the AUC will be performed as a sensitivity analysis.

7.8 Regression diagnostics

Residual plots will be used to assess the appropriateness of the regression models fitted. We will plot:

- Histograms and probability plots to assess normality
- Scatterplots of residuals against fitted values to assess constant variance and linearity, and to identify potential outliers

For models with two levels different levels or residuals will be assessed; patients (level-1 residuals) within sites (level-2 residuals).

Should the normality assumption be untenable for any continuous outcomes (i.e., EQ-5D-5L and VAS scales) including after log transformation, a non-parametric method will be undertaken using change from baseline as a sensitivity analysis, although covariate adjustments will not be possible.

7.9 Multiple imputation by chained equations (MICE)

To avoid bias in treatment effect estimates and standard errors and loss in efficiency, missing outcome values will be imputed using MICE [1] under the assumption that missing data values are likely to be missing at random (MAR) which means they are dependent on the values of the observed data, but not dependent on the values of the missing data.

Reasons for missingness may be important; these will be investigated using logistic regression analysis to predict missingness from observed covariates and observed

outcomes. Missing data patterns will be explored, as will percentages of missing data for each variable and baseline imbalance among those for whom the outcome is observed.

The number of imputed datasets will be proportional to the study attrition rate. The revised sample size assumed 15% of patients would not contribute to primary outcome data, therefore at least 15 imputed datasets will be drawn. Data imputation will be done separately for each of the four randomised groups before being combined to form a single imputed dataset [6]. Each imputed dataset will be analysed separately and the results combined using Rubin's rules [7] to produce a single treatment effect estimate and 95% confidence interval. If at the end of the trial the attrition rate is higher than 15%, the number of imputed datasets will be increased accordingly.

Missing outcome values will be replaced with simulated values from a set of imputation models containing all potential prognostic baseline covariates (i.e., age, gender, HIV status and previous occurrence of warts), the two components of the primary outcome (i.e., clearance within 16 weeks of starting treatment and remaining wart-free between 16 and 48 weeks), wart clearance at 4, 8, and 24 weeks, quality of life outcomes (i.e. EQ-5D-5L index and VAS at baseline, 4, 8, 16, 24 and 48 weeks), the censoring indicator and Nelson-Aalen estimator of the cumulative hazard of the time-to-event outcomes (i.e. time to complete wart clearance and time from complete wart clearance to recurrence/relapse) plus the following auxiliary variables: indicator of whether any additional treatment therapy was given to achieve clearance by 16 weeks, indicator of recorded compliance problems with the initially allocated treatment at 4, 8 and 16 weeks, indicator of unscheduled follow-up visits and symptom scores at 4, 8, 16, 24 and 48 weeks. We shall initially attempt to apply an imputation process in which missing values in each variable listed above are imputed based on all the other variables. In the event however that this leads to over-fitting problems, for example due to collinearity between some variables or small cell sizes, we shall instead select imputation models for each variable giving priority to measures of the same factor at other time points and to other factors measured at the same time point.

The imputation model specified will reflect the distribution of the missing outcome data; missing values for continuous outcomes will be imputed from linear regression models, missing values for binary variables will be imputed from binary logistic models and missing values for ordinal variables will be imputed from ordinal logistic models.

7.9.1 Multiple imputation model diagnostics

To assess the extent to which imputed values differ from observed values we will produce:

- Histograms of imputed values and observed values to compare the distributions of observed and imputed data
- Summary statistics of the observed and imputed data to explore differences between the observed and imputed data [10].

- Scatterplots of residuals against fitted values of each imputed dataset. Similar patterns across datasets will be an indication of the compatibility of the imputation model(s).

8 TABLES AND GRAPHS

8.1 Tables

Table 1: Number of patients screened but not enrolled and reasons* not enrolled by site

For those not recruited and not randomised	SITES†																						TOTAL
	MM	YK	HO	BR	HE	MA	SO	LI	BO	JC	ME	SG	RH	TC	KI	NE	CF	KC	DC	DB	SM	GL	
Screened																							
Previous wart treatment in the last 3 months																							
Previous quadrivalent HPV vaccine																							
Previous intolerance to either of the topical treatments, vaccines or their constituents																							
Known HIV-positivity																							
Pregnancy or lactating women																							
Women of child bearing potential not willing to use effective contraception																							
Unable or unwilling to complete follow-up																							
Lesion > 4 cm^2 requiring direct supervised treatment																							
Patients who have had topical or systemic steroids applied, or other immunosuppressive agents < 1 month prior to randomisation																							
Patients enrolled in any other trial of an IMP																							
Other																							
Total eligible																							
Eligible not randomised																							
Refused consent																							
Randomised																							

**Only one reason is tabulated for each participant.*

†MM = Mortimer Market Centre, YK = York, HO = Homerton, BR = Brighton, HE = Heartlands, MA = Manchester, SO = Southend, LI = Liverpool, BO = Bournemouth, JC = James Cook, ME = Medway, SG = St George's, RH = Royal Hallamshire, TC = Trafalgar Clinic, KI = Kings, NE = Newcastle, CF = Cardiff, KC = Kent Community Health, DC = Dorset County, DB = Derby, SM = St Mary's, GL = Gloucester

Table 2: Randomisation to HIPvac by month and site

		SITES†																					
Year	**Month**	**MM**	**YK**	**HO**	**BR**	**HE**	**MA**	**SO**	**LI**	**BO**	**JC**	**ME**	**SG**	**RH**	**TC**	**KI**	**NE**	**CF**	**KC**	**DC**	**DB**	**SM**	**GL**	**TOTAL**
2014	Nov																							6
	Dec																							12
2015	Jan																							17
	Feb																							30
	Mar																							53
	Apr																							76
	May																							93
	Jun																							112
	Jul																							140
	Aug																							155
	Sep																							170
	Oct																							189
	Nov																							211
	Dec																							225
2016	Jan																							246
	Feb																							274
	Mar																							292
	Apr																							313
	May																							340
	Jun																							362
	Jul																							369
	Aug																							385

	Sep																							410
	Oct																							444
	Nov																							484
	Dec																							501
2017	Jan																							504
TOTAL																								

†MM = Mortimer Market Centre, YK = York, HO = Homerton, BR = Brighton, HE = Heartlands, MA = Manchester, SO = Southend, LI = Liverpool, BO = Bournemouth, JC = James Cook, ME = Medway, SG = St George's, RH = Royal Hallamshire, TC = Trafalgar Clinic, KI = Kings, NE = Newcastle, CF = Cardiff, KC = Kent Community Health, DC = Dorset County, DB = Derby, SM = St Mary's, GL = Gloucester

Table 3: Demographic and baseline characteristics of the randomised participants by treatment allocated

		Treatment group			
		Imiquimod + qHPV vaccine	**Podophyllotoxin + qHPV vaccine**	**Imiquimod + placebo**	**Podophyllotoxin + placebo**
		n=	**n=**	**n=**	**n=**
Demographics					
Age (years)	Mean (SD)				
Stratification variables					
Gender, n(%)	Male				
	Female				
Previous occurrence of warts, n(%)	None				
	1 or more				
HIV positive, n(%)	Yes				
	No				
Quantity of warts					
Diameter of largest wart (mm)	Mean (SD)				

Total No. of warts, n(%)	1-5				
	6-10				
	11-20				
	>20				
Position of warts, n(%)					
Male	Penile, Shaft				
	Penile, Glans				
	Penile, Foreskin				
	Perineum				
	Anal/perianal				
	Other				
Women	External genitalia				
	Perineum				
	Anal/perianal				
	Other				
Sexual Orientation, n(%)	Heterosexual				
	Homosexual				
	Bisexual				
	Other				
Partners in the last 3 months	Mean (SD)				
Sexual practices in the last three months, n(%)	Vaginal sex				
	Passive oral sex				
	Performed oral sex				

	Anal receptive sex				
	Insertive anal sex				
Current contraception (female), n(%)	Condoms				
	Other barrier contraception e.g. diaphragm, female condom				
	Hormonal contraception e.g. pills, IUS, implant, depo				
	Not sexually active				
	Other				
	None				
	N/A – Not of child-bearing potential				
Health history					
Previous treatment (warts) , n(%)	Yes				
	No				
Wart treatment for last episode, n(%)	Podophyllotoxin				
	Imiquimod				
	Cryotherapy				
	Surgery				
	Other				
Previous bivalent HPV vaccine, n(%)	Yes				

	No				
How many doses	Mean (SD)				
Previous sexually transmitted infection (STI), excluding anogenital warts, n(%)	Yes				
	No				
Type of STI, n(%)	Chlamydia				
	Gonorrhea				
	Syphilis				
	Herpes				
	Other				
Number of STI episodes	Median (IQR)				
Smoking, n(%)	Daily				
	Less than daily				
	Ex-smoker				
	Never smoked				
Quality of life					
EQ-5D-5L: Health Utility	Mean (SD)				
EQ-5D-5L: VAS	Mean (SD)				

Table 4: Measures of treatment exposure and compliance by treatment allocated and topical treatment

		Treatment group				Topical treatment	
		Imiquimod +qHPV vaccine	Podophyllotoxin + qHPV vaccine	Imiquimod + placebo	Podophyllotoxin + placebo	Imiquimod	Podophyllotoxin
		n=	n=	n=	n=	n=	n=
Topical treatment							
Has the patient switched treatments, n(%)	Yes						
	No						
Has the patient discontinued treatment, n(%)	Yes						
	No						
Has the patient had cryotherapy, n(%)	Yes						
	No						
Has the patient had any other treatment at their treatment centre, n(%)	Yes						
	No						
Has the patient had any treatment from a source outside their treatment centre, n(%)	Yes						
	No						
Vaccine	None						
Has the patient missed any vaccine doses, n(%)	1 dose						
	≥2 doses						
	None						
	1 dose						
	≥2 doses						

Table 5: Primary outcome and key secondary outcomes by treatment allocated

	n(%)				Adjusted Coefficient* (95%CI)	
	Imiquimod + qHPV vaccine n=	**Podophyllotoxin + qHPV vaccine n=**	**Imiquimod + placebo n=**	**Podophyllotoxin + placebo n=**	**Topical treatment effect n=**	**Vaccine treatment effect n=**
Primary outcome						
Wart free at 16 weeks and remaining wart-free between 16 and 48 weeks						
Clinically important secondary outcomes Proportion of patients:						
Wart free at 16 weeks						
Remaining wart free at 48 weeks after clearance at 16 weeks						
Secondary effectiveness outcomes						
Proportion of patients wart free:						
At the end of the assigned treatment course (4 or 16 weeks)						
At the end of the assigned treatment course (4 or 16 weeks) without additional treatment during the assigned treatment course†						
Quantity of additional treatment (number of cryotherapy applications, additional weeks of	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)		

podophyllotoxin or imiquimod, or any other treatment not specified in the protocol) required to achieve clearance by 16 weeks						
Proportion wart-free at 16 weeks without receiving additional treatment†						
Proportion experiencing complete wart clearance at any time during the 48 week trial period						
Proportion of patients experiencing wart recurrence/relapse after complete wart clearance						
Time to complete wart clearance	See Kaplan-Meier plots					
Time from complete wart clearance to recurrence/ relapse						
Severity of most severe side effects:	Mean (SD)					
At 4 weeks						
At 8 weeks						
At 16 weeks						
At 24 weeks						
At 48 weeks						
Quality of life						
EQ-5D-5L Health Utility:						
At 4 weeks						
At 8 weeks						
At 16 weeks						
At 24 weeks						
At 48 weeks						
AUC						
EQ-5D-5L Health State: VAS						
At 4 weeks						

At 8 weeks						
At 16 weeks						
At 24 weeks						
At 48 weeks						
AUC						

**Odds ratio (OR) for binary and ordinal outcomes; hazard ratio for time to event data.*

†Additional treatment defined as one or more episodes of cryotherapy, additional weeks of podphyllotoxin or imiquimod, or any other treatment outside the treatment regime described in the protocol

Table 7: Number of participants reporting any events by treatment allocated

	n (%)			
	Imiquimod + qHPV vaccine n=	**Podophyllotoxin + qHPV vaccine n=**	**Imiquimod + placebo n=**	**Podophyllotoxin + placebo n=**
AEs				
ARs				
SAEs				
SARs				
SUSARs				

8.2 Graphs

G1: CONSORT flow chart.

G2: Profile plots with error bars by treatment allocated for effectiveness outcomes.

G3: Kaplan-Meier plots of the survivor function by treatment allocated for time to event outcomes.

9 REFERENCES

[1] White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011 Feb 20;30(4):377-99. doi: 10.1002/sim.4067. Epub 2010 Nov 30.

[2] McAlister FA, Straus SE, Sackett DL, Altman DG. Analysis and reporting of factorial trials: a systematic review. JAMA. 2003 May 21;289(19):2545-53.26.

[3] Kahan BC. Bias in randomised factorial trials, Stat Med. 2013; 32, 4540-4549.

[4] Kahan BC, Morris TP. Improper analysis of trials randomised using stratified blocks or minimisation. Stat Med. 2012 Feb 20;31(4):328-40. doi: 10.1002/sim.4431. Epub 2011 Dec 4.

[5] Kahan BC, Morris TP. Reporting and analysis of trials using stratified randomisation in leading medical journals: review and reanalysis. BMJ. 2012 Sep 14;345:e5840. doi: 10.1136/bmj.e5840.

[6] Sullivan TR, White IR. Should multiple imputation be the method of choice for handling missing data in randomized trials? SMMR.2017. doi:10.1177/0962280216683570

[7] Rubin D. Multiple imputation for nonresponse in surveys. New York: Wiley & Sons 1987

[8] White IT, Thompson SG. Adjusting for partially missing baseline covariates in randomized trials. Statistics in Medicine. 2005; **24**:993–1007

[9] Carpenter JR, Kenward MG. Missing data in randomised controlled trials – a practical guide

[10] Devlin NJ, Shah KK. Valuing health-related quality-of-life: An EQ-5D5L value set for England. Health Economics. 2017 Jun.

[11] Office of Health Economics. Comparing the UK EQ-5D-3L and English EQ-5D-5L Value sets