GlaxoSmithKline	**Statistical Analysis Plan**
Detailed Title:	Epidemiological, prospective, interventional, multi-centre, long term HAV antibody persistence in children vaccinated with 1 dose and those vaccinated with 2 doses of Havrix® in Panama.
eTrack study number and Abbreviated Title	201630 (EPI-HAV-007 BOD PA)
Scope:	All data pertaining to the above study.
Date of Statistical Analysis Plan	Final: 29-Jan-2018
Co-ordinating author:	PPD
Other author(s):	PPD / PPD
Adhoc reviewers:	PPD (Epidemiologist) PPD (Epidemiologist) PPD (Lead Statistician) PPD (Lead Stat Analyst) PPD (Scientific Writer)
Approved by:	PPD (Epidemiologist) PPD (Epidemiologist) PPD (Lead Statistician) PPD (Lead Stat Analyst) PPD (Scientific Writer) PPD (Statistician)

APP 9000058193 Statistical Analysis Plan Template (Effective date: 14 April 2017)

TABLE OF CONTENTS

PAGE

The complete statistical analysis plan and results presentation is divided into 2 parts: the first part detailing the analyses to be performed (known as SAP, current document) and a second part, annex (-es) (called TFL) describing the flow and format of tables, figures and listings to be annexed to the SR.

LIST OF ABBREVIATIONS

AE	Adverse event
ATP	According-To-Protocol
CI	Confidence Interval
CRF	Case Report Form
CTRS	Clinical Trial Registry
Eli Type	Internal GSK database code for type of elimination code
ELISA	Enzyme-linked immunosorbent assay
GMC	Geometric mean antibody concentration
GSK	GlaxoSmithKline
LL	Lower Limit of the confidence interval
MedDRA	Medical Dictionary for Regulatory Activities
N.A.	Not Applicable
SAE	Serious adverse event
SAP	Statistical Analysis Plan
SBIR	GSK Biological's Internet Randomization System
SD	Standard Deviation
SR	Study Report
TFL	Tables Figures and Listing template annexed to SAP
UL	Upper Limit of the confidence interval

1. DOCUMENT HISTORY

Date	Description	Protocol Version
29-JAN-2018	Final version	Amendment 2 – 24 July 2017

2. STUDY DESIGN

This is an epidemiological, serial, cross-sectional, interventional, multi-centre study in Panama.

The following epochs will be analysed:

Group order in tables	Group label in tables	Group definition for footnote
1	Epoch 1 (Year 8)	Subjects with ≥ 7 years and < 10 years between last *Havrix* dose and Persistence visit
2	Epoch 2 (Year 10)	Subjects with ≥ 10years and < 13 years between last *Havrix* dose and Persistence visit

The following dose group names will be used for the statistical analyses:

Group order in tables	Group label in tables	Group definition for footnote
1	1 Dose	Subjects received one dose of *Havrix*
2	2 Doses	Subjects received two doses of *Havrix*

Time interval between the last *Havrix* vaccination dose and the Persistence visit:

Sub-group order in tables	Sub-group label in tables	Sub-group definition for footnote
1	≥7 years and <8 years	≥7 years and <8 years between last *Havrix* vaccination and Persistence visit
2	≥8 years and <9 years	≥8 years and <9 years between last *Havrix* vaccination and Persistence visit
3	≥9 years and <10 years	≥9 years and <10 years between last *Havrix* vaccination and Persistence visit
4	≥10 years and <11 years	≥10 years and <11 years between last *Havrix* vaccination and Persistence visit
5	≥11 years and <12 years	≥11 years and <12 years between last *Havrix* vaccination and Persistence visit
6	≥12 years and <13 years	≥12 years and <13 years between last *Havrix* vaccination and Persistence visit

Age class at the first dose of Havrix:

Sub-group order in tables	Sub-group label in tables	Sub-group definition for footnote
1	[0 – 6 months[	Subjects aged less than 6 months at the time of the first dose of *Havrix*
2	[6 – 12 months[	Subjects aged between 6 months and 12 months at the time of the first dose of *Havrix*
3	[12 – 18 months[	Subjects aged between 12 months and 18 months at the time of the first dose of *Havrix*
…	…	…

3. OBJECTIVES

3.1. Primary objective

- To assess the persistence of anti-hepatitis A virus (HAV) antibodies, ***approximately*** 8 years and 10 years post vaccination with the last received vaccine dose of the complete series of *Havrix* (2 doses) and the partial series completion (1 dose) in Panama.

3.2. Secondary objectives

- To assess geometric mean concentration (GMC) of anti-HAV antibodies, ***approximately*** 8 years and 10 years post vaccination with the last received dose of the complete series of *Havrix* (2 doses) and the partial series completion (1 dose) in Panama.

- To explore the non-inferiority of the 1-dose schedule of *Havrix* when compared to the 2-dose schedule in terms of the percentage of subjects with anti-HAV antibody concentrations ≥ 15 mIU/mL, ***approximately*** 8 years and 10 years after the last received vaccination dose.

Criteria for evaluation: The lower limit of the 2-sided 95% CI for the difference (1-dose group minus the 2-dose group) of percentage of subjects with anti-HAV antibody concentrations ≥ 15 mIU/mL is greater than or equal to the pre-defined clinical non-inferiority limit of -10%.

4. ENDPOINTS

4.1. Primary endpoint

- Anti-HAV seropositivity status at ***approximately*** 8 years and 10 years after administration of the last dose of *Havrix*.

Subjects are defined as being seropositive if their anti-HAV antibody concentration is ≥ 15 mIU/mL.

4.2. Secondary endpoints

- Anti-HAV concentrations (GMC) at ***approximately*** 8 years and 10 years after administration of the last dose of *Havrix*.
- Anti-HAV seropositivity status at **approximately** 8 years and 10 years after administration of the last dose of *Havrix* – corresponds to the secondary objective for non-inferiority.

Subjects are defined as being seropositive if their anti-HAV antibody concentration is ≥ 15 mIU/mL.

5. STUDY POPULATION

The list of applicable elimination codes for each cohort can be found in the study specific form FORM-BIO-CLIN-9004-05 Criteria for eliminating subjects from the analyses.

Cohort	Elimination codes	Eli Type
Total enrolled cohort at Year 8 (Epoch 1)	900	Y8
ATP cohort for analysis for persistence at Year 8 (Epoch 1)	2010, 2100, 2120	Y8
Total enrolled cohort at Year 10 (Epoch 2)	900	Y10
ATP cohort for analysis for persistence at Year 10 (Epoch 2)	2010, 2100, 2120	Y10

6. STATISTICAL METHODS

All statistical analyses will be performed using SAS version 9.3 or later.

6.1. Analysis of demographics/baseline characteristics

All the analyses concerning objectives will be performed on the ATP cohort for each cross-sectional survey (Epoch1, Epoch2).

Demography

The number of subjects enrolled as well as the number excluded from the ATP analysis will be presented.

The distribution of subjects by centre will be tabulated according to *Havrix* dose group.

Demographic characteristics such as age at the Persistence visit, age/age class [months] at the time of first and second dose of *Havrix* administration, time interval [months] between *Havrix* dose 1 and dose 2 administration, time since last *Havrix* dose [years], sex, geographic ancestry, rural and urban residency of the subjects will be summarized using descriptive statistics per dose group and overall:

- Frequency tables will be generated for categorical variables such as sex.
- Mean, median, standard deviation, minimum and maximum will be provided for continuous data such as age.

The analysis will be performed on the Total enrolled and ATP cohorts for the demographic part.

Subjects enrolled in the two epochs

The number and percentage of subjects enrolled in the two epochs will be summarised once the 2 surveys have been completed at Epoch 2, by *Havrix* dose group.

Hepatitis A infection disease history

The percentage of subjects with a history of Hepatitis A infection will be tabulated, the age at the infection [years] and the time between the infection and the last *Havrix* dose [year].

General medical history

The percentage of subjects reporting any pre-existing conditions, signs or symptoms having started before enrollment in the study will be summarized by diagnosis status, current or past.

6.2. Analysis of persistence

If the percentage of enrolled subjects with available serological results excluded from the ATP cohort is more than 5%, a sensitivity analysis based on the Total enrolled cohort will be performed.

Within group analysis

For the subjects who had received either one dose or two doses of *Havrix*, at each serosurvey (***approximately*** 8 years and 10 years after the last vaccine dose):

- Percentage of subjects with concentrations above the cut-off (≥ 15mIU/mL) with exact 95% CI will be calculated [seropositivity rate].
- Geometric mean concentration (GMC) with 95% CI will be tabulated.
- The distribution of antibody concentrations will be presented by histogram of anti-HAV log titer and also using reverse cumulative curves.

The seropositivity rate and the geometric mean concentrations (GMC) with 95% Confidence Intervals (CI) of antibody to hepatitis A virus (Anti-HAV) will be calculated and presented overall, by *Havrix* dose group, by follow up time class period after their last *Havrix* vaccination dose and by age class at the first *Havrix* dose.

Between groups analysis:

An evaluation of the differences in the immune response ***approximately*** 8 years and 10 years after the last vaccination dose with *Havrix* between the subjects who received only one dose of *Havrix* and subjects who received two doses of *Havrix* will be performed in terms of:

- Differences in seropositivity rate (seropositive = *anti-HAV antibody concentrations above the cut-off (≥15 mIU/mL))* between the subjects who received only one dose of *Havrix* and subjects who received two doses of *Havrix* at each serosurvey.

  Computation of the asymptotic standardised 2-sided 95% CIs for the difference in the percentage of subjects with anti-HAV antibody concentrations ≥ 15 mIU/mL at

Year 8 or at Year 10 after the last received *Havrix* vaccination dose (1-dose group minus the 2-doses group).If the lower limit of the 2-sided 95% CI for the difference (1-dose group minus the 2-doses group) is greater than or equal to the pre-defined clinical limit of -10%, non-inferiority of 1 dose will be demonstrated.

- Ratio (1-dose group/2-doses group) of the anti-HAV GMCs between the subjects who received only one dose of *Havrix* and subjects who received two doses of *Havrix* with their 95% CI at each serosurvey.

6.3. Analysis of safety

SAEs related to study procedures reported during the entire study will be listed.

7. STATISTICAL CALCULATIONS

7.1. Methods for calculating Confidence Intervals (CIs),

All CI's will be 2-sided 95% CI.

- The exact 95% CIs for a proportion within a group will be calculated [Clopper, 1934].
- The standardized asymptotic 95% CI for the group difference in proportions will be calculated [Robert,. 1998]: the standardized asymptotic method used within GSK Vaccine is **method 6**
- The 95% CI for Geometric mean concentrations (GMCs) will be obtained within each group separately. The 95% CI for the mean of log-transformed concentrations will be first obtained assuming that log-transformed values were normally distributed with unknown variance. The 95% CI for the GMCs will be then obtained by exponential-transformation of the 95% CI for the mean of log-transformed concentrations.
- The GMC ratio will be obtained using an ANOVA model on the logarithm based 10-transformed concentrations. The ANOVA model will include the dose group as fixed effect. The GMC ratio and its 95% CI will be derived as exponential-transformation of the log-transformed mean difference and its 95% CI.

7.2. Derived and transformed data

7.2.1. Date derivation

- SAS date derived from a character date: In case day is missing, 15 is used. In case day & month are missing, 30 June is used.

7.2.2. Demography

- The continuous variable age [years] at the Persistence visit will be computed as the difference between the informed consent date and the date of birth /365.25.
- The continuous variable age [months] at the time of first or second dose of *Havrix* is computed based on the number of days between the administration of *Havrix* and the date of birth /30.5.
- The time interval [months] between first and second dose of *Havrix* is derived as the difference between date on second dose and date of first dose plus 1 day /30.5.
- Time interval [years] between last *Havrix* vaccination dose and the Persistence visit will be derived using the date of the visit minus the date of the last *Havrix* vaccination plus 1day / 365.25.

7.2.3. Antibody persistence

- For a given subject and given antibody concentrations, missing or non-evaluable measurements will not be replaced. Therefore, we will exclude these subjects with missing or non-evaluable measurements.
- The Geometric Mean Concentrations (GMCs) calculations will be performed by taking the anti-log of the mean of the log concentrations transformations. Antibody concentrations below the cut-off (15 mIU/mL) of the assay will be given an arbitrary value of half the cut-off of the assay for the purpose of GMC calculation.
- A seronegative subject is a subject whose antibody concentration is below the cut-off value of the assay. A seropositive subject is a subject whose antibody concentration is greater than or equal to the cut-off value of the assay.
- The assay cut-off is the value under which there is no quantifiable result available. For an assay with a specific ‘cut_off’ , numerical immuno result is derived from a character field (rawres):
  - If rawres is ‘NEG’ or ‘-’ or ‘(-)’, numeric result= cutt_off/2,
  - if rawres is ‘POS’ or ‘+’ or ‘(+)’, numeric result = cut_off,
  - if rawres is ‘< value’ and value<=cut_off, numeric result =cut_off/2,
  - if rawres is ‘< value’ and value>cut_off, numeric result =value,
  - if rawres is ‘> value’ and value<cut_off, numeric result =cut_off/2,
  - if rawres is ‘> value’ and value>=cut_off, numeric result =value,
  - if rawres is ‘<= value’ or ‘>= value’ and value<cut_off, numeric result =cut_off/2,
  - if rawres is ‘<= value’ or ‘>= value’ and value>=cut_off, numeric result =value,

- if rawres is a value < cut_off, numeric result = cut_off/2,
- if rawres is a value >= cut_off, numeric result = value,
- else numeric result is left blank.

7.3. Number of decimals

The following decimal description will be used for the demography and persistence tables.

Display Table	Parameters	Number of decimal digits
Demographic characteristics	Mean, median age	1
Demographic characteristics	SD (age)	1
Persistence	Ratio of GMC	2
Persistence	Mean, Min, Q1, Median, Q3, Max for GMC	1
All summaries	% of count, including LL & UL of CI	1
All summaries	% of difference, including LL & UL of CI	2

8. CONDUCT OF ANALYSES

8.1. Sequence of analyses

Description	Analysis ID	Disclosure Purpose	Reference for TFL
Analysis of epoch 1 (Year 8)	E1_02	CTRS	All tables from TFL dated 29Jan2018
Analysis of epoch 2 (Year 10)	E1_01	Study report	All tables from TFL dated 29Jan2018

8.2. Statistical considerations for interim analyses

No interim analysis is planned. However the analysis will be done separately for the two epochs. As all analyses are descriptive, no adjustment of type I error is needed.

9. CHANGES FROM PLANNED ANALYSES

The histogram of anti-HAV log titer will be used to explore the possible outliers (subjects with very high concentrations that may have been elicited by natural HAV infection) in the study and if we detect a suspicious situation of natural infection we will do an additional analysis which determines a ‘natural infection cut-off’ by using the method describe in the publication reference [Wiedermann, 2007] or a more appropriate method. We will perform a sensitivity analysis by excluding the potential outliers detected with the ‘natural infection cut-off’.

10. REFERENCES

Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of binomial. *Biometrika*. 1934;26:404-413

McMahon BJ. Et al. Duration of protection against hepatitis A for the current two-dose vaccine compared to a three-dose vaccine schedule in children, Vaccine 31 (2013) 2152-2155

Robert G. Newcombe, Interval estimation for the difference between independent proportions: comparison of eleven methods, *Statist Med.* 1998; 17, 873-890

Wiedermann U. et al. Persistence of seroprotection 10 years after primary hepatitis A vaccination in an unselected study population Vaccine 25 (2007) 927-931