NAXIVA

Phase II **N**eoadjuvant study of **AXI**tinib for reducing extent of venous tumour thrombus in clear cell renal cell cancer with **V**enous inv**A**sion.

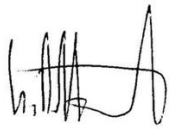
Statistical Analysis Plan

Version 2.0, 20th August 2020

		Previously
Author	**Lisa Hopcroft** Scottish Cancer Trials Research Unit Public Health Scotland Edinburgh EH12 9EB	***Robert Hill*** *Scottish Cancer Trials Research Unit* *NHS National Services Scotland* *Edinburgh EH12 9EB*
Reviewer	**Richard Dobbie** Public Health Scotland Edinburgh EH12 9EB	***James Paul*** *Head of Biostatistics* *Cancer Research UK Clinical Trials Unit* *The Beatson West of Scotland Cancer Centre* *Level 0* *1053 Great Western Road* *Glasgow, G12 0YN*

Trial Coordinating Centre: Scottish Clinical Trials Research Unit, Edinburgh.
Cancer Clinical Trials Unit, Scotland (CaCTUS).

ISRCTN:	ISRCTN96273644
EudraCT Number:	2017-000619-17
REC reference:	17/EE/0240

SAP authorised by			
Name:	Grant D Stewart	**Role in Trial:**	Chief Investigator
Signature:		**Date:**	20/8/2020

DOCUMENT CONTROL

Document Control	
Version	v2.0
Date Issued	20/08/2020
Author(s)	Lisa Hopcroft
Other Related Documents	Protocol (version 2.0); CRF (version 1.0);
Comments to	Lisa.Hopcroft1@nhs.net

Document History				
Version	**Date**	**Author**	**Reviewer**	**Amendments**
0.1	18/1/2018	Jack Vize	Robert Hill	New document started
0.2	22/1/2018	Robert Hill	Jack Vize	Updating missing sections
0.3	10/12/2018	Robert Hill		Updating the statistical methods
0.4	26/02/2019	Robert Hill		Updating following comments by the DMC
0.5	26/03/2019	Robert Hill		Updated following comments from Elaine McCartney and Grant Stewart
0.6	10/04/2019	Robert Hill	Jim Paul	Updated to incorporate comments from Grant Stewart
0.7	22/4/2019	Robert Hill	Jim Paul	Updated to incorporate comments from Jim Paul
1.0	23/4/2019	Robert Hill	Jim Paul	Updated following additional comments from Jim Paul
2.0	20/8/2020	Lisa Hopcroft	Richard Dobbie	Extensive update following identification of inconsistencies with previous, draft versions (specifically 0.5, 0.7); in consistencies between the SAP and protocol; and comments from DMC, notably: • Additional endpoint for RV-only patients is defined, following discussion and agreement with the DMC. • Addition of sensitivity analysis using the ITT population for assessments of endpoints, in addition to the evaluable population. Also updated Sponsor information.

TABLE OF CONTENTS

1 Study Objectives

1.1 Primary endpoint

The primary end-point for this study is the percentage of evaluable patients with an improvement in disease as a result of taking the study drug. The definition of an improvement will vary according to the patient's Mayo Level as captured at screening:

- **For patients presenting at screening with a Mayo Classification of Level 1 or above**, an improvement in disease will be represented by a reduction in their Mayo Classification at week 9.
- **For patients presenting at screening with a Mayo Classification of Level 0**, an improvement in disease will be represented by either:
  - a change of tumor thrombus from main renal vein to branches of the renal vein (on the right); or
  - a change of tumor thrombus from main renal vein to the renal vein lateral to the gonadal vein (on the left).

See **Section 6.1** for further details regarding the definition of improvement in relation to the Mayo classification and the surgical approach classification.

1.2 Secondary endpoints

Secondary endpoints are:

- % change in surgical approach
- % change in VTT length
- Response rate (RECIST)
- Evaluation of surgical morbidity assessed by Clavian-Dindo classification

2 Study design

NAXIVA is a single arm, single agent, open label, phase II feasibility study of axitinib in patients with both metastatic and non-metastatic renal cell carcinoma of clear cell histology. 20 evaluable patients will be recruited from multiple centres within the United Kingdom.

Patients who have signed informed consent and who have met all eligibility criteria will be registered into the trial.

The starting dose of axitinib will be 5mg BID, escalated to 7mg BID and then 10mg BID. A dose modification assessment will take place every 2 weeks in clinic during the 8 week pre-surgical treatment period and will be dependent on tolerability of treatment. Patients will follow an aggressive axitinib dose escalation process

within the 8 week period to a maximum of 10mg BID. Patients should stop axitinib a minimum of 36 hours and a maximum of 7 days prior to surgery in week 9.

Blood, urine and tissue samples will be taken prior to and during therapy to evaluate biomarkers of treatment response. Nephrectomy and IVC tumour thrombectomy will be planned for all patients on the trial.

Where there is diagnostic doubt at site, a central review of week 0, 3 and 9 MRI and CT scans will be conducted to determine if a patient has progressed as per RECIST 1.1 criteria. In addition, those particular patient cases will be discussed on an individual basis by the trial steering committee (TSC). In addition, a central review of scans will be carried out once all scans have been carried out to ensure that all measurements are compliant with the NAXIVA MRI SOP (v2.1).

Patients will be assessed at weeks 1, 3, 5, 7, and 9 before surgery and followed up in clinic at 6 & 12 weeks post-surgery.

3 Reporting outline

This section outlines the planned content of interim, final analyses and, if relevant, additional statistical study reports.

3.1 Interim Analysis

An interim analysis will be performed after thirteen patients have been recruited (recruitment will only be suspended for the purposes of conducting the interim analysis if no improvement in the Mayo Classification has been seen in any patients at the point where the thirteenth patient is recruited). The contents of this report are outlined in this section. Further details regarding the composition of the relevant patient populations are provided in **Section 4**.

3.1.1 Recruitment/accrual

Data describing patient accrual, including comparison to desired targets, will be presented here (**Section 7.1**).

3.1.2 Early stopping rules

Stopping for reasons of futility: if zero evaluable patients have shown an improvement in disease, from screening to week 9, the trial will be stopped for futility. The rule for futility is specified in (**Section 1.1**) with further details set out in the efficiency analysis section (**Section 7.6.1**).

Stopping for reasons of patient safety: completion rates for patient safety data in the ITT population (see **Section 4.2** for population definition) to assess whether the early stopping rules have been violated will be prioritised for review by the DMC. Should data completion be adequate to assess the early stopping rules (i.e., complete data for thirteen patients), each rule will be assessed. If any rules are violated the trial will be stopped for reasons of patient safety. These rules are specified in **Section 5.2** with further details set out in the safety analysis section (**Sections 7.5.1, 7.5.2** and **7.5.3**).

3.1.3 Safety

Data regarding toxicities, adverse events/reactions, serious adverse events (SAEs) and suspected unexpected serious adverse reaction (SUSARs) will be provided here. These analyses will be specified further in **Section 5.1**, with further details set out in the Toxicities (**Sections 7.5.4**) and Serious Adverse Events sections (**Section** Error! Reference source not found.).

3.1.4 Efficacy

Analyses relating to efficacy endpoints will be carried out on the evaluable population (see **Section 4.3** for population definition); these analyses will be further specified in **Section 6**Error! Reference source not found. with further details set out in the Efficacy analysis section (**Section 7.67.6.1**). All endpoints will also be calculated for the ITT population to ensure that the conclusions drawn are the same in the context of patient discontinuation/compliance; this will require a change in denominator where percentages are calculated (**Section 6.1**Error! Reference source not found.).

3.1.5 Baseline demographics

Baseline demographics will be presented for the ITT population (see **Section 7.3** for details).

3.2 Final Analysis

The final analysis will be performed after all data for 20 evaluable patients has been obtained. The analyses performed and the data presented is as described in **Section 3.1** for the interim report.

If analysis findings with regards to efficacy endpoints for the evaluable population and the ITT population conflict (see **Section 3.1.4**), the DMC will discuss and advise with regards to the interpretation of these results.

In addition, 80% two-sided confidence intervals (to correspond to the 10% 1-sided sample size calculation) for the proportions relevant to the efficacy endpoints will be calculated using the approach in Koyama and Chen[5] (again using both the evaluable and ITT populations as alternative contexts).

4 Populations and sample size

4.1 Sample Size Calculations and recruitment schedule

The aim is to recruit 20 patients over a 24 month period. A Simon two stage minimax design to distinguish a <5% from a >25% improvement in the Mayo classification requires 20 evaluable patients (90% power, 10% 1-sided).

Thirteen patients would be recruited in the first stage. If no patients demonstrate an improvement in their Mayo classification between screening and week 9, accrual to the clinical trial would stop (see **Section 3.1.2**). If one or more patients demonstrate an improvement in the Mayo classification between screening and week 9, a further seven patients would be recruited (to the final total of 20 evaluable patients). In order for the clinical trial to be considered a success, at least three evaluable patients should demonstrate an improvement in disease or treatment between screening and week 9 (see **Section 7.6.1**).

4.2 Intention-to-treat (ITT) Population

The Intention-to-treat (ITT) population includes all patients registered onto the study.

4.3 Evaluable Population

The evaluable population includes all patients in the ITT population who have received at least one dose of the study drug (including any patients who were enrolled in error, received study drug and were subsequently found to be ineligible).

4.4 Safety Population

The safety population includes all patients in the ITT population who have received at least one dose of the study drug (including any patients who were enrolled in error, received study drug and were subsequently found to be ineligible). In this study, this population is equivalent to the evaluable population.

5 Safety Evaluations

The study will use the following safety parameters.

5.1 Adverse events, SAEs and SUSARs

All reported adverse events (including adverse reactions, SAEs and SUSARs) will be presented, with events categorised against the National Cancer Institute Common Terminology Criteria for Adverse Events v4.0.

5.2 Early stopping rules

Patient safety will be ensured by applying the following early stopping rules:

(1) If a single M0 (staging data: cancer has not spread to other parts of the body) patient progresses to the point where surgery is no longer possible, then recruitment of new patients will be suspended following a review by the data monitoring committee (DMC)

(2) If a single M0 (staging data: cancer has not spread to other parts of the body) patient becomes a M1 (staging data: cancer has spread to other parts of the body) patient, recruitment of new patients will be suspended following a review by the DMC.

(3) If in 3 patients the VTT extends (as indicated by an increase in Mayo classification) but patients remain surgically resectable the trial will close. If the VTT has extended the patient will be expedited and operated on as soon as possible.

6 Efficacy evaluations

Note that these evaluations will be made for both the evaluable and ITT populations. This represents a sensitivity analysis that assesses whether results are robust in the ITT population.

This study will use the following efficacy parameters:

- At least 3 evaluable patients showing improvement in disease extent between screening and week 9 are required for a positive trial.
- The change in Mayo classification between screening and week 9 in evaluable patients will be presented with the percentage of patients with an improvement, no change and a deterioration in Mayo classification shown. Percentages will be calculated using both the evaluable population and the ITT population.
- The VTT length measurement will be calculated as the sum of: (1) length of IVC tumour thrombus ABOVE renal vein (measured from ostium of RV+IVC to tip of tumour thrombus); (2) length of IVC tumour thrombus BELOW renal vein (measured from ostium of RV+IVC to tip of tumour thrombus) and (3) length of RV thrombus. Those patients who are identified as “main renal vein” will only have the third value populated. Any values that are missing cannot be assumed to represent 0 and must be clarified either via contact with sites or by manual checking of the paper CRF.

- The percentage of patients who have had a change in the surgical approach will be calculated with the percentage with a less invasive and the percentage with a more invasive approach shown separately. Two pieces of data are relevant here: (1) a change from “Open Surgery” to “Minimally invasive surgery”; and (2) a change from a more invasive open to a less invasive open surgical approach (see **Section 6.1** for a definition of surgical approach in this respect). Percentages will be calculated using both the evaluable population and the ITT population.
- The response rate, as determined by RECIST criteria v1.1 and provided by sites will be presented and the percentage of outcomes categorised as complete response, partial response and progressive disease shown; only pre-surgical timepoints (i.e., weeks 3 and 9) will be considered. To provide further quality assurance that these RECIST evaluations are correct, the RECIST response will be recalculated from source data by the study statistician and compared to the evaluations provided provided by sites. Should the evaluations not correspond with each other, a data query will be raised with the sites.
- Morbidity, in terms of the Clavien-Dindo classification system will be presented and percentage of patients with complications of grade I to V will be presented. Percentages will be calculated using both the evaluable population and the ITT population.

6.1 Additional definitions

Definitions of the Mayo classification and surgical approach classification systems that should be used to define improvement are shown below:

Mayo classification (levels are ordered by increasing extensiveness):

- Level 0: thrombus limited to the renal vein
- Level 1: into IVC <2cm from renal vein ostium level
- Level 2: IVC extension >2cm from renal vein ostium and below hepatic vein
- Level 3: thrombus at the level of or above the hepatic veins but below the diaphragm
- Level 4: thrombus extending above the diaphragm.

Tumour thrombus surgical management approach classification (approaches are ordered by increasing invasiveness):

1. Thrombus - Milked back into renal vein and side clamped
2. Infra-hepatic (IVC clamping with no liver mobilisation)
3. Retro-hepatic (liver mobilisation and clamping below hepatic veins)
4. Retro-hepatic (liver mobilisation and clamping above hepatic veins)
5. Supra-hepatic (infradiaphragmatic)

6. Supra-hepatic (supradiaphragmatic)

Note that efficacy outcomes can be expressed as a percentage of the ITT population or the evaluable population (meaning that the denominator should change to be either the size of the relevant population).

7 Tabulations, Figures and Listings

This section is the complete operational specification for the statistical analysis and subsequent presentation of that analysis for the lifetime of this trial. Throughout this section, tabulations consist of counts and percentages per category unless stated otherwise. Listings will include the patient identifier and the specific data for that listing.

7.1 Study Recruitment (ITT population)

- Plot of actual cumulative recruitment against time since study opened will be provided; target recruitment rate line will be superimposed on this plot.
- Table of recruitment by centre showing the date the centre opened, the date the first and last patients were recruited and the number recruited. Centres that have opened but not recruited patients will be included in this table.

7.2 Eligibility/Population membership (ITT population)

- Numbers of patients in each population (**Sections 4.2** and **4.3**) will be presented.
- Ineligible patients will be listed along with the reason for their ineligibility.
- Patients who are recruited but do not receive any treatment will be listed with reasons for not treating.

7.3 Baseline data and demographic characteristics (ITT population)

Discrete variables will be summarised by frequencies and percentages. For continuous variables medians, inter-quartile ranges and ranges will be presented and where appropriate, the results will be presented in suitable categories:

- Registering centre
- Age at registration
- Karnofsky Performance status (ECOG)
- Other medical conditions (frequency and list), with specific reference to cardiovascular, cerebrovascular, thromboembolic, respiratory, gastrointenstinal, endocrine/metabolic,

neurological conditions (further specification of “Other” reasons will be listed). Existing allergies should also be included and specified.

- Heart rate, respiratory rate, systolic and diastolic blood pressure
- Height, Weight, BMI
- Pregnancy test (conducted and result)
- Existing signs and symptoms (with CTCAE gradings)
- Laboratory investigations, including full blood count; white blood cell count; serum biochemistry; thyroid function; urinalysis
- Previous diagnosis of other cancer
- Renal Cell Carcinoma histological type (including whether predominantly clear cell)
- Tumour data, including staging (TNM status); grading (ISUP/Furman); MSKCC model and risk; IMDC model and prognosis
- Surgical assessment data, including whether patient is surgically resectable; whether radiotherapy has been given
- Lesion data: number of target and non-target lesions; sums of longest target lesions diameters
- Thrombus data: IVC and/or MRV involvement; relevant measurements (see **Section 6**)
- Mayo classification

7.4 Delivery of study therapy (evaluable population)

- Total dose (mg) taken per patient separately for axitinib (median, IQ range and range)
- Total dose (mg) missed per patient separately for axitinib (median, IQ range and range)
- Percentage of intended dose taken per patient (taken/(taken+missed)) separately for axitinib (median, IQ range and range)
- Reasons for missed doses will be tabulated; further specification of “other” reasons will be listed.
- The number of dose reductions per patient
- The number of dose escalations per patient
- Reasons why axitinib was reduced to 3mg BID
- Reasons why axitinib was reduced to 2mg BID
- Reasons for stopping trial drug will be tabulated; “other” reasons and “Adverse event/toxicity” reasons will be listed

7.5 Safety Analysis

7.5.1 To inform early stopping rule (1) (Section (1)5) (ITT population)

A listing of:

- Patient label
- Staging (M) data at screening
- Date of Surgery
- Flag indicating whether surgery took place
- Whether the patient belongs to the evaluable population
- Relevant patient notes including any reasons that surgery was not performed or staging has not been provided

7.5.2 To inform early stopping rule (2) (Section 5) (ITT population)

A listing of:

- Patient label
- Staging (M) data at screening
- Flag indicating whether surgery took place
- Staging (M) data at surgery
- Whether the patient belongs to the evaluable population
- Relevant patient notes including any reasons that surgery was not performed or staging has not been provided

7.5.3 To inform early stopping rule (3) (Section 5) (ITT population)

A listing of:

- Patient label
- Mayo classification at screening
- Mayo classification at week 3
- Mayo classification at week 9
- Flag indicating whether surgery took place
- Details of all patients with extension of their VTT will be presented including the planned surgery and procedure performed.

7.5.4 Toxicities (including adverse events and laboratory values) (Safety population)

- All adverse events will be coded according to the CTC toxicity criteria and the worst value over the study drug periods determined for each patient.
- Laboratory values will be coded according to the CTC toxicity criteria and the worst value over the study period determined for each patient.
- Any adverse events or laboratory anomalies which are graded ≥ grade 1 by at least 10% of patients will be tabulated.
- Details of toxicities will be summarised to show (i) grade at trial entry; (ii) the highest grade over the course of treatment; and (iii) the highest grade in follow up. For any toxicity graded ≥ grade 2 at any timepoint, the toxicity grading will be tracked over the course of the trial in a graphical form.

7.5.5 Serious Adverse Events/Reactions and SUSARs (ITT population)

- Details of serious adverse events/reactions and SUSARs will be provided in a separate document output from the SCTRU's pharmacovigilance database.

7.6 Efficacy analysis (evaluable and ITT populations)

Note that these analyses will be carried out for both the evaluable and ITT populations. This represents a sensitivity analysis that assesses whether results are robust in the ITT population.

7.6.1 Primary endpoints

NB. For all primary endpoint assessments, if a patient does not have a week 9 scan, then their week 3 scan (if available) should be used instead.

7.6.1.1 For patients presenting at screening with a Mayo Classification of Level 1 or above

- The number and percentage of patients with an improvement in the Mayo Classification between screening and week 9 (see **Section 6.1** for Mayo classification system).
- The number and percentage of patients with no change in the Mayo Classification between screening and week 9 (see **Section 6.1** for Mayo classification system).
- The number and percentage of patients with deterioration in the Mayo Classification between screening and week 9 (see **Section 6.1** for Mayo classification system).

7.6.1.2 For patients presenting at screening with a Mayo Classification of Level 0

- The number and percentage of patients demonstrating improvement in status, i.e.:

- For patients presenting with right renal vein tumour thrombus: a change of tumour thrombus from main renal vein to branches of the renal vein between screening and week 9.
- For patients presenting with left renal vein tumour thrombus: a change of tumour thrombus from the main renal vein between screening and week 9 as follows:
  - If the patient is lateral to the gonadal vein at screening, a change from the main renal vein into the segmental veins
  - If the patient is medial to the gonadal vein at screening, a change from the main renal vein medial to the gonadal vein to the renal vein lateral to the gonadal vein

- The number and percentage of patients with no change in tumour thrombus status (i.e., tumour still present in the main renal vein and medial to the gonadal vein (on left side)) between screening and week 9.
- The number and percentage of patients demonstrating an extension of the thrombus into the inferior vena cava between screening and week 9.

7.6.2 Secondary endpoints

- % change in surgical approach for all patients comprising:
  - % with less invasive approach
  - % with no change in surgical approach
  - % with more invasive approach
- % change in VTT length across all timepoints for each patient (see **Section 6**)
- % change in VTT length, stratified by pre-surgical RECIST response (as provided by site)
- Waterfall plot of % change in VTT length, coloured by pre-surgical RECIST response (as provide by site)
- Radiological response (RECIST) data (pre-surgical timepoints only), comprising:
  - % with complete response
  - % with partial response
  - % with stable disease
  - % with progressive disease
  - comparison with RECIST values calculated in-house where possible
- Evaluation of morbidity assessed by Clavian-Dindo classification
  - % with grade I to V complications

7.7 Additional analyses

The correlation of the drug exposure with the response of the VTT will be examined by presenting the Pearson's correlation coefficient between the total dose of drug received and the percent reduction in VTT.

8 Statistical Methods for the Analysis of the Study Data

As this is a single arm feasibility study results will be presented as counts, frequencies and percentages. 80% confidence intervals will be provided for the percentage of evaluable patients with an improvement in their Mayo Classification between screening and week 9, percentage change in surgical approach and percentage change in VTT height.

8.1 Methods for handling missing values

As part of the trial management process missing data is closely managed with the intention of minimising the amount of missing data to the lowest level possible. Particular attention will be placed upon primary and secondary end point data.

SCTRU will regularly review the data for compliance with the protocol, and for inconsistent or missing data. Should any missing data or data anomalies be found, queries will be sent to the site for resolution. Sites are expected to respond to queries within a timeous manner.

All missing data is queried with sites and where the data cannot be sources an explanation is required and quality assurance procedures are applied to validate the quality and completeness of the data before any analysis is carried out.

Once the data has been declared as ready for analysis a missing value analysis will be carried out. This will describe the pattern of missing data, identify where the missing data are located and how extensive it is, identify pairs of variables which have missing values in multiple cases and assess if data is missing randomly.

Means, standard deviations, covariates and correlations for different missing value methods (listwise, pairwise, regression and expectation-maximisation) will be produced.

Where missing data is above an acceptable level of 10% for any primary or secondary outcome data or explanatory covariates multiple imputation will be considered.

9 Reference list

1. Eisenhauer EA, *et al*. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur J Cancer* **2009**;45:228-47.
2. National Cancer Institute. Cancer Therapy Evaluation Programme. Common Terminology Criteria for Adverse Events v4.0. *Available at* *http://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm*
3. Mander A.P., Thompson S.G. Two-stage designs optimal under the alternative hypothesis for phase II cancer clinical trials. *Contemp Clin Trials.* **2010** Nov; 31(6): 572–578. Available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3049867/
4. Vollset, S. E. 'Confidence intervals for a binomial proportion', *Statistics in Medicine*, **12**, 809–824 (1993).
5. Koyama, Tatsuki, and Heidi Chen. "Proper inference from Simon's two-stage designs." *Statistics in medicine* vol. 27,16 (**2008**): 3145-54. Available at https://onlinelibrary.wiley.com/doi/10.1002/sim.3123.