NCT Number: NCT02984683

STATISTICAL ANALYSIS PLAN

Open-label Phase 2 study evaluating efficacy and safety of SAR566658 treatment in patients with CA6 positive metastatic Triple Negative Breast Cancer

SAR566658-ACT14884

STATISTICIAN: [redacted]

DATE OF ISSUE: 29-May-2018

Total number of pages: 82

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND DEFINITION OF TERMS

AT:	all-treated
ATA:	antitherapeutic antibodies
[redacted]	[redacted]
CR:	complete response
CT:	computed tomography
DCR:	disease control rate
DFS:	disease free survival
DOR:	duration of response
ECOG:	Eastern Cooperative Oncology Group
e-CRF:	electronic case report form
EOT:	end-of-treatment
FFPE:	formalin-fixed paraffin-embedded
FIH:	first in human
IHC:	immunohistochemistry
IMP:	investigational medicinal product
INR:	international normalized ratio
IVRS:	interactive voice response system
LDH:	lactate dehydrogenase
MedDRA:	Medical Dictionary for Regulatory Activities
NCI-CTCAE:	National Cancer Institute - Common Terminology Criteria for Adverse Events
NE:	not evaluable
ORR:	objective response rate
PCSA:	potentially clinically significant abnormality
PFS:	progression-free survival
PR:	Partial response
PS:	performance status
PT:	preferred term
RECIST:	response evaluation criteria in solid tumors
SD:	stable disease
SOC:	system organ class
TEAE:	treatment emergent adverse event
TNBC:	triple negative breast cancer
TTP:	time to progression
WHO-DD:	World Health Organization Drug Dictionary

1 OVERVIEW AND INVESTIGATIONAL PLAN

1.1 STUDY DESIGN AND RANDOMIZATION

This is an open-label, multicenter, Phase II study evaluating the efficacy and safety of SAR566658 administered as a single agent via intravenous infusion in patients with CA6-positive metastatic triple negative breast cancer (TNBC) tumors after at least 1 prior chemotherapy regimen, but no more than 3 for advanced/metastatic disease.

The study will be performed in 2 parts.

Part1: Selection of the dose

Patients with metastatic TNBC overexpressing CA6 (membrane intensity of 2+, 3+ in ≥30% of tumor cells) will be randomized to either the 90 mg/m^2 or the 120 mg/m^2 cohort, both given at Day 1 and Day 8 of each 3-week cycle. Patients randomized but not treated will be replaced. Study treatment should be initiated no later than 3 calendar days after randomization.

During treatment, safety will be monitored and the risks for predefined safety criteria during the first 2 cycles will be followed and assessed using a Bayesian binomial model. In particular, risks for overdosing and unacceptable toxicity will be assessed as soon as 4 patients treated at the dose level of 120 mg/m^2 become evaluable for the predefined safety criteria.

The Bayesian model assessing the rate of predefined safety criteria will use dose as a parameter. The elicitation of the prior distribution is derived from the data collected in the first in human (FIH) study (TED10499) at various dose levels and schedules of administration (190 mg/m^2 q3w, 240 mg/m^2 q3w, 90 mg/m^2 D1D8 q3w, and 120 mg/m^2 q2w).

The elicitation of the prior distribution for the two dose levels is explained in Appendix A.

The decision to stop 1 arm may be taken on the basis of risk of the assessed overdosing (ie, predefined safety criteria rate above 40%) and unacceptable toxicity (ie, predefined safety criteria rate above 60%). The model recommends stopping 1 arm when its overdosing risk or unacceptable toxicity risk will not be controlled at the levels of 25% and 5%, respectively.

Simulations were performed and are presented in Appendix A.

Randomization will be centralized by IVRS/IWRS, stratified on Eastern Cooperative Oncology Group (ECOG) performance status (PS) (0 versus 1). Randomization will be performed when all eligibility criteria are checked and the patient is deemed eligible. Randomization will be stopped at the end of Part 1 and the recruitment will be authorized in Part 2 only when the selected dose is chosen.

The selection of the dose to be continued in Part 2 will be made at the end of Part 1 following the criteria described below:

A preliminary statistical evaluation of antitumor activity by Response Evaluation Criteria in Solid Tumors (RECIST) 1.1 (1) will be done on the first 14 patients treated in each cohort; if at least 2 confirmed responses, partial response (PR) or complete response (CR), in at least 1 cohort are observed then the study can continue in Part 2, if the safety profile is acceptable.

In addition to the efficacy criteria, safety criteria including the number (%) of patients experiencing an eye disorder leading to dose modification or drug discontinuation, or peripheral neuropathy, or a treatment-emergent adverse event (TEAE) leading to dose reduction or drug discontinuation, will be considered.

In case of the efficacy criteria (<2 PR or CR in 14 patients) not being reached for a given dose, then this dose will not be selected.

In case of neither dose reaching the efficacy criteria (<2 PR or CR in 14 patients), then the study will be stopped.

In case of comparable objective response rate (ORR) and safety profiles between 90 mg/m^2 D1D8 q3w and 120 mg/m^2 D1D8 q3w, then the lower dose (90 mg/m^2 D1D8 q3w) will be chosen.

In case of a higher ORR in 120 mg/m^2 D1D8 q3w and a comparable safety profile to 90 mg/m^2 D1D8 q3w, then 120 mg/m^2 D1D8 q3w will be chosen.

In case of a higher ORR in 120 mg/m^2 D1D8 q3w and a better safety profile with 90 mg/m^2 D1D8 q3w, then the dose will be selected on the best risk/benefit.

If there are 6 responses or more in the first 14 patients treated at 1 dose, then the study could be stopped for overwhelming efficacy.

Part 2: at the selected dose

Part 2a: Overexpression cohort

The expansion/overexpressing CA6 cohort with an additional 15 metastatic TNBC patients with CA6 overexpression (membrane intensity of 2+, 3+ in ≥30% of tumor cells) will start at the selected dose using the same selection criteria as for Part 1. If at least 7 confirmed responses (PR or CR) are seen among the 29 patients, then the null hypothesis of 12% ORR can be rejected. This test has a 1-sided type 1 error of 10% and 80% power for an alternative hypothesis of 30% response rate.

Part 2b: Mild expression cohort

An additional “mild CA6 expression” cohort of 20 patients with metastatic TNBC with at least 1% positive tumor cells at CA6 membrane intensity ≥1+ and <30% of tumor cells at intensity of 2+, 3+, will be assessed at the selected dose.

1.2 OBJECTIVES

1.2.1 Primary objectives

To evaluate the tumor ORR, according to RECIST 1.1 (Section of SAR566658 in patients with CA6-positive metastatic TNBC:

- **Part 1:** To select the SAR566658 dose based on ORR and safety of 2 dose levels of SAR566658 in patients with metastatic TNBC overexpressing CA6 (membrane intensity of 2+, 3+ in ≥30% of tumor cells),
- **Part 2:**
  - ***Part 2a:*** To demonstrate the activity of SAR566658 based on ORR in patients with metastatic TNBC overexpressing CA6 (membrane intensity of 2+, 3+ in ≥30% of tumor cells) treated at the selected dose in an expanded cohort, in addition to the patients treated in Part 1,
  - ***Part 2b:*** To assess the efficacy of SAR566658 based on ORR in patients with metastatic TNBC and mild CA6 expression (with at least 1% positive tumor cells at intensity ≥1+ and <30% of tumor cells at intensity of 2+, 3+) treated at the selected dose in a "mild CA6 expression cohort".

1.2.2 Secondary objectives

Part 1 and Part 2:

- To assess:
  - Disease Control Rate (DCR), Duration Of Response (DOR), Time To Progression (TTP), Progression-Free Survival (PFS),
  - The impact of ocular primary prophylaxis on the incidence of keratopathies,
  - The PK profile of SAR566658,
  - The potential immunogenicity of SAR566658,
  - The relationship between CA6 expression level in the tumor, and circulating CA6 in blood at baseline, and efficacy outcomes.
- To evaluate the global safety profile

1.2.3 Exploratory objectives

[redacted]

[redacted]

1.3 DETERMINATION OF SAMPLE SIZE

1.3.1 Sample size for Part 1 and Part 2a

It is anticipated that approximately 63 patients will be enrolled into this study.

This study is separated into 2 parts:

Part 1:

Each dose level 90 mg/m^2 D1D8 q3w and 120 mg/m2 D1D8 q3w will include in parallel 14 patients overexpressing CA6.

Based on efficacy and safety (including predefined safety criteria), 1 dose will be selected during the interim analysis at the end of the Part 1 if the futility criteria on ORR is not met and will be continued in Part 2.

Part 2:

Part 2a: An additional 15 patients overexpressing CA6 will be treated in the expansion cohort at the selected dose.

Part 2b: In parallel, 20 patients will be treated at the selected dose in a mild CA6 expression cohort.

1.3.2 Sample size calculation

The sample size calculation is based on efficacy criteria (ORR).

A 2-stage design will be used for the selected dose. It is expected that a beneficial ORR under SAR566658 is of at least 30% (H1). An ORR of 12% (or less) will be considered as clinically non relevant (H0). On the basis of these assumptions, 29 treated patients at the selected dose at the end of Part 2a would be necessary to reject the null response rate of 12% with a global power of 80% and a 1-sided alpha level of 10%. The calculation takes into account 1 interim analysis using a gamma (-2) beta spending function for futility analysis and a Lan-DeMets (OF) alpha spending function in a 1 sample test for a binomial proportion (East version 6.3 using exact computations, Cytel Software, Cambridge, MA).

An interim analysis is planned in Part 1 after 14 patients are treated at each dose. If at least 2 responses over the first 14 patients are observed, then 15 other patients will be treated in Part 2a at the selected dose, if the safety profile is tolerable.

A total of 7 confirmed responses (CR or PR) or more out of 29 patients (14 patients from Part 1 and 15 patients from Part 2a at the selected dose) will be necessary to reject the null hypothesis at the end of the Part 2a expansion cohort.

Sample size determination for the mild CA6 expression cohort (Part 2b).

At the selected dose, the relationship between CA6 expression level and efficacy outcome will be assessed by adding a cohort of 20 patients with mild CA6 membrane staining (at least 1% positive tumor cells at intensity ≥1+ and <30% of tumor cells at intensity 2+, 3+).

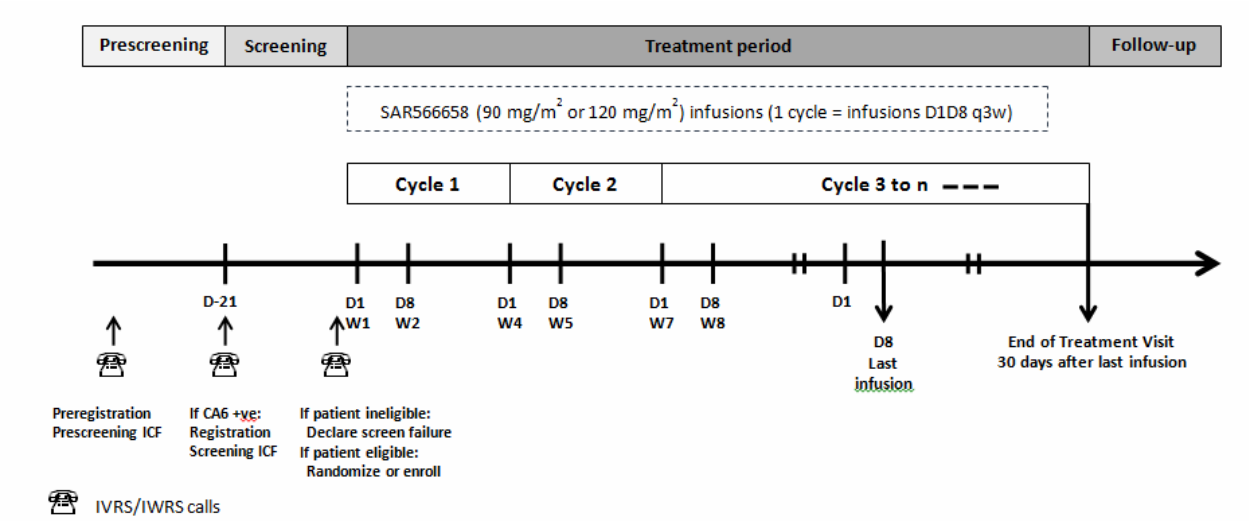
With a sample size of 20 patients, a true response rate above 15% can be rejected with 83% confidence if we observe 0 or 1 response. Moreover, if we observe 7 responses over the 29 patients in the biomarker positive group (≥30%, 2+,3+) and ≤1 response in the mild CA6 biomarker positive group (CA6 IHC = <30% 2+,3+ and ≥1%, ≥1+), the 1-sided p-value (according to Fisher's exact test) will be around 7%, suggesting an association between biomarker status and ORR.

1.4 STUDY PLAN

Patient pre-screening, screening and enrollment will be centralized by Interactive Voice Response System (IVRS).

Each pre-screened patient will have to sign a pre-screening informed consent which allows the collection of an archived available formalin-fixed paraffin-embedded (FFPE) tumor sample in order to assess centrally for CA6 expression by IHC. A first call to IVRS center will be done prior to sending the slides to central laboratory in order to obtain a pre-screening patient number. For each patient of whom CA6 expression is documented according to the protocol definition, and who have further signed the screening informed consent, the study consists of:

- up to 21-day screening phase prior start of study treatment infusion: the IVRS is called again (2^{nd} call) and the protocol study procedures could start,
- 21-day treatment cycle(s): once the patient has completed the necessary baseline study procedures and is deemed eligible, a third call is given to IVRS in order to enroll the patient. Patients from the Part 1 will be randomized in two treatment arms: 90 mg/m^2 D1D8Q3w or 120 mg/m^2 D1D8Q3w stratified by ECOG PS (0 versus 1). Initiation of study treatment should start within 3 calendar days after enrollment is provided,
- an end of treatment visit (30 days after the last study treatment infusion),
- and a follow-up phase every 6-weeks until radiological progression disease, death or withdrawal of consent.



Abbreviations: +ve: positive; D: day; ICF: informed consent form; IVRS/IWRS: Interactive Voice/Web Response System; q3w: every 3 weeks; W: week.

1.5 MODIFICATIONS TO THE STATISTICAL SECTION OF THE PROTOCOL

This section summarizes major changes to the protocol statistical section with emphasis on changes after study start (after the first patient was enrolled). The first patient was enrolled on 2017-03-23.

Table 1 - Protocol amendment statistical changes

Amendment Number	Date Approved	Rationale	Description of statistical changes
1		To update the sample size in order to use the same methods (exact methods) in the calculation of sample size as the one planned in statistical analyses section for the calculation of the confidence interval of the primary and secondary endpoints. To decrease the power from 85% to 80%, which is acceptable for a phase II study	Sample size: The sample size was increased to 29 patients at the selected dose in the overexpression cohort (instead of 28) Number of responses : Number of responses expected at the end of the POC at the selected dose in the overexpression cohort was increased to 7 (instead of 6) Number of responses expected at the end of the interim analysis for overwhelming efficacy was increased to 6 (instead of 5) Power: Power was decreased from 85% to 80%.

1.6 STATISTICAL MODIFICATIONS MADE IN THE STATISTICAL ANALYSIS PLAN

Not applicable.

2 STATISTICAL AND ANALYTICAL PROCEDURES

As a general convention, investigational medicinal product (IMP) administration or study treatment infusion are indifferently used through the document. Both mean SAR566658 infusion.

2.1 ANALYSIS ENDPOINTS

2.1.1 Demographic and baseline characteristics

The baseline value is defined as the last value or measurement available before the first study treatment infusion in the study. This definition applies for all variables unless otherwise specified. For patients randomized and not treated, the baseline value is defined as the last available value obtained up to the date and time of randomization.

All baseline safety and efficacy parameters (apart from those listed below) are presented along with the on-treatment summary statistics in the safety and efficacy sections (Section 2.4.4 and Section 2.4.5).

Demographic characteristics

Demographic variables include gender (Male, Female), race (White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Other Pacific Islander, Not reported, Unknown), age in years (quantitative and qualitative variable : <65, [65 - 75[and ≥75 years), ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not reported, Unknown), weight (kg) and BSA, and ECOG PS at baseline.

Medical or surgical history not related to the study disease

Medical and surgical history will include significant prior and concurrent illnesses, using the same definition as the one used in Section 2.4.5.1 for these AEs

This information will be coded using the version of Medical Dictionary for Regulatory Activities (MedDRA) currently in effect at Sanofi at the time of database lock.

Disease characteristics at baseline

Specific disease history will include:

- **Cancer at initial diagnosis:**
  - Time from initial diagnosis to first study treatment administration (in years),
  - Histopathology type,
  - Staging and TNM classification,
  - Scarff Bloom and Richardson histoprognostic factor.

- **Status of the disease at study entry:**
  - measurability of disease (yes / no),
  - number and location of organs with metastases (<3 vs. ≥3) according to the Investigator
- **CA6 expression** (see Section 2.1.7.1)
  - CA6 expression centrally assessed on an archival biopsy
  - Circulating CA6 (centrally assessed)
- **Molecular expression**: Basal-Like status, Luminal A status, Luminal B status, BRCA1 mutation status, BRCA2 mutation status, PD-1 and PD-L1 expression

Prior anticancer therapies

Prior anticancer therapies include previous surgery for cancer, radiation therapy, and systemic anti-cancer treatment (including chemotherapy, immunotherapy, targeted therapy, biologics…).

Systemic anti-cancer treatments and surgery will be respectively coded using the World Health Organization-Drug Dictionary (WHO-DD) and MedDRA using the versions currently in effect at Sanofi at the time of database lock.

Prior anticancer treatment will include:

- Patients with neoadjuvant therapy
- Patients with adjuvant therapy
- Patients with neoadjuvant/adjuvant therapy
- Patients with advanced chemotherapy

For all patients, results to be provided will be:

- Number of patients with previous anthracycline regimen alone,
- Number of patients with previous taxane regimen alone,
- Number of patients with previous anthracycline/taxane regimen.

Among patients with neoadjuvant/adjuvant followed by advanced chemotherapy, patients who relapse within 6 months (<7 months) of end of adjuvant therapy will be identified

For patients with advanced chemotherapy, results to be provided will be:

- number of prior lines (continuous and 1,2,3, >3):
  - Adjuvant/neo-adjuvant therapy will be counted as prior chemotherapy regimen if the patient had a progression/recurrence within 6 months after completion of the treatment,
  - Prior hormonal, biologic (eg, bevacizumab) or immunotherapy, without a cytotoxic agent, are allowed and are not counted as line of therapy

- A chemotherapy line in advanced/metastatic disease is an anticancer regimen that contains at least 1 cytotoxic chemotherapy agent and was discontinued due to progression. If a cytotoxic chemotherapy regimen was discontinued for a reason other than disease progression then this regimen does not count as a "prior line of chemotherapy" unless this regimen was discontinued after treatment response (CR, PR or Stable Disease SD) and disease progression occurring before another line of chemotherapy.

- best overall response of the last prior advanced chemotherapy
- time from completion of the last prior advanced chemotherapy to first study treatment administration (months)
- reason for discontinuation of the last prior advanced chemotherapy
- duration of last line of therapy
- prior treatment with eribulin (Yes vs No and if yes detail ; number of prior regimen with eribulin and best response across all the regimen, reason of discontinuation, time from completion to first study treatment administration (months))
- Disease free survival (DFS) defined as the interval from the first adjuvant chemotherapy date to the last date of relapse, classified into two groups: ≤24 months and >24 months

For prior surgery, number (%) of patients with any prior surgery related to study disease, type of surgery and time from last surgery to first study treatment administration (months) will be provided.

For prior radiotherapy, number (%) of patients with any prior radiotherapy related to study disease, intent and time from last radiotherapy to first study treatment administration (months) will be provided. Adjuvant radiotherapy (yes vs no) will be counted.

Vital signs

See Section 2.1.4.5.

Laboratory exams

See Section 2.1.4.4.

Any technical details related to computation, dates, and imputations for missing dates are described in Section 2.5.4.

ECG

See Section 2.1.4.6.

2.1.2 Prior, concomitant or post treatment medications (other than anticancer therapies)

All medications taken within 21 days before the first study treatment administration and at any time during the study treatment period are to be reported in the case report form pages.

All medications will be coded using the World Health Organization-Drug Dictionary (WHO-DD) using the version currently in effect at Sanofi at the time of database locks.

- Prior medications are those the patient used prior (within 21 days) to first IMP intake. Prior medications can be discontinued before first administration or can be ongoing during treatment phase.
- Concomitant medications are any treatments received by the patient concomitantly to the IMP(s), from first administration to the last treatment infusion +30 days. A given medication can be classified both as a prior medication and as a concomitant medication. Concomitant medications do not include medications started during the post treatment period (as defined in the observation period in Section 2.1.4).
- Post treatment medications are those the patient took in the period running from 30 days after the last treatment infusion up to the end of the study.

Premedications

As defined in Section 8.2.2 of the study protocol, patients will routinely receive premedications prior to SAR566685 infusion to reduce the risk and severity of infusion-related hypersensitivity reactions and corneal toxicities. Premedications are defined in the protocol as non-investigational medicinal product(s). The recommended premedication agents are:

- For hypersensitivity reactions:

Premedication with histamine H1 antagonist (diphenhydramine 50 mg given orally approximately 1 hour before SAR566658 administration) is required for all patients

- For corneal toxicity:

Topical artificial tears and /or hyaluronic ophthalmic gel, vasoconstrictor such as neosynephrine (or equivalent), corticosteroid ocular gel such as Sterdex or dexamethasone 0.16%.

The use of cold eye mask/pad (unless patient cannot tolerate it) from start until end of infusion

Any technical details related to computation, dates, imputation for missing dates are described in Section 2.5.

2.1.3 Efficacy endpoints

2.1.3.1 Primary efficacy endpoint(s)

The primary efficacy endpoint will be the Overall Response Rate (ORR) defined as the proportion of patients with confirmed complete response (CR) or confirmed partial response (PR) as Best

Overall Response (BOR), assessed by the Investigator (based on RECIST 1.1), relative to the total number of patients in the efficacy population (see Section 2.3.1). Confirmation of objective responses will be performed by repeat tumor imaging (Computed Tomography [CT] scan, MRI) at least 4 weeks after the first radiological documentation of response.

Tumor assessments will be performed every 6 weeks (ie, every 2 cycles) during the treatment period and every 6 weeks during the follow-up period if a patient discontinues treatment prior to radiologic disease progression.

Tumor assessments after the initiation of post-treatment anti-cancer therapy will be excluded from the determination of Best Overall Response in a patient. See Appendix B (Determination of tumor response) for details.

Best Overall Response

BOR is the best tumor response observed from first study treatment infusion until disease progression, death or initiation of post-treatment anti-cancer therapy, whichever occurs first.

Primary ORR analysis can be performed when all patients treated in Part 2a and Part 2b have had at least 4 postbaseline tumor assessments, experienced confirmed objective response, or have discontinued the study for any reason. However, if the primary criteria of 7 responders is either met or not met with certainty at the end of the Part 2a in the selected dose when all patients have had 2 postbaseline tumor assessments, experienced confirmed objective response, or discontinued the study for any reason and when all the patients of the Part 2b have had 2 postbaseline tumor assessments, experienced confirmed objective response, or discontinued the study for any reason, analysis may be done at that time instead.

Best tumor shrinkage

For patients with measurable disease, the best tumor shrinkage will be defined by the best relative change from baseline as follows:

$$Y = 100 \times \frac{Y_{smallest} - Y_{baseline}}{Y_{baseline}}$$

Where Y is the sum of longest diameter of the target lesions.

2.1.3.2 Secondary efficacy endpoint(s)

The secondary efficacy endpoints include duration of response (DOR), Disease Control rate (DCR), Time to progression (TTP) and Progression Free Survival (PFS).

2.1.3.2.1 Duration of response

The duration of response will be defined as the time from the date of first initial occurrence of a subsequently confirmed CR or PR to the date of first radiological documentation of disease

progression (using RECIST 1.1) before the initiation of any post-treatment anti-cancer therapy or death (due to any cause), whichever occurs first.

Duration of response is determined only for patients who have achieved a response of confirmed CR or PR.

If a patient receives post-treatment anti-cancer therapy before radiologic progression or death, the duration of response will be censored at the date of the last valid response assessment without evidence of progression performed before the initiation of the post-treatment anti-cancer therapy.

If a patient receives post-treatment anti-cancer therapy without radiologic progression or death, the duration of response will be censored at the date of the post-treatment anti-cancer therapy.

In the absence of radiologic progression, death or post–treatment anti-cancer therapy, the duration of response will be censored at the date of the last valid response assessment without evidence of progression.

A valid response assessment is defined as a response assessment with a response different of NE.

2.1.3.2.2 Disease control rate

Disease control rate is defined as the proportion of patients with confirmed CR or PR or stable disease (SD) lasting at least 3 months as assessed by the Investigator (based on RECIST 1.1) relative to the total number of patients in the relevant efficacy population (see Section 2.3.1).The calculation of the duration of stable disease is critical for the selection of patients with stable disease lasting at least 3 months. The duration of stable disease will be calculated according to the following rules:

- The duration of stable disease will be calculated as the time from the date of first study treatment infusion to the date of first documentation of progression, date of death or date of last valid tumour assessment before the initiation of any post-treatment anti-cancer therapy.
- If a patient receives post-treatment anti-cancer therapy before radiologic progression or death, the duration of response will be calculated until the date of the last valid response assessment without evidence of progression performed before the initiation of the post-treatment anti-cancer therapy.
- If a patient receives post-treatment anti-cancer therapy without radiologic progression or death, the duration of stable disease will be calculated until the date of initiation of the post-treatment anti-cancer therapy.

In the absence of disease progression, death or post-treatment anti-cancer therapy, the duration of stable disease will be calculated to the date of the last valid response assessment without evidence of progression.

A valid response assessment is defined as a response assessment with a response different of NE.

2.1.3.2.3 Time to progression

The time to progression is defined as the time interval from the date of first study treatment administration to the date of first radiologically documented tumor progression or death from disease progression before the initiation of any post-treatment anti-cancer therapy, whichever occurs first.

If a patient receives post-treatment anti-cancer therapy before radiologic progression or death, TTP will be censored at the date of the last valid response assessment without evidence of progression performed before the initiation of the post-treatment anti-cancer therapy.

If death different from disease progression occurs before receiving post-treatment anti-cancer, TTP will be censored at the date of the last evaluable response assessment without progression.

If a patient receives post-treatment anti-cancer therapy without radiologic progression or death, TTP will be censored at the date of the post-treatment anti-cancer therapy.

In the absence of tumor progression or death due to disease progression or post-treatment anti-cancer therapy, TTP will be censored at the date of the last valid response assessment without evidence of progression. An evaluable response assessment is defined as a response assessment with a response different of NE.

2.1.3.2.4 Progression-free survival

Progression-Free Survival (PFS) is defined as the time interval from the date of the first treatment infusion to the first documented date of disease progression or death due to any cause, whichever occurs first.

Rules for sensitivity analysis

PFS endpoint will be determined considering also clinical/non radiological progression as an event. If clinical/non radiological progression is observed in absence of documented radiological progression, the date of progression will be the date of clinical/non radiological progression. If the first progression is based on documented radiological progression and on clinical/non radiological progression, the date of progression will be the earliest of date of documented progression and date of clinical/non radiological progression.

Censoring rules for all PFS analyses:

For these 2 analyses, in the absence of disease progression or death (or clinical/non radiological progression for sensitivity analysis), the PFS will be censored at the date of last evaluable tumor assessment without documented progression. In addition, the patient without any valid post-baseline tumor assessment will be censored for PFS on the day of the first study treatment infusion (Day 1).

The summarized description of these analyses is given in Appendix F.

2.1.4 Safety endpoints

The safety analysis will be based on the reported adverse events (AE) and other safety information, such as clinical laboratory data, vital signs, ECG, weight and Eastern Cooperative Oncology Group (ECOG) performance status (PS).

Observation period

The observation period will be divided into 3 periods:

- The **pre-treatment** period is defined as the time from the signed informed consent date up to the first study treatment administration.
- The **treatment** period is defined as the time from the first study treatment administration to the last study treatment administration + 30 days.
- The **post-treatment** period is defined as the period of time starting the day after the end of the treatment period up to the end of the study.

2.1.4.1 Adverse events variables

Adverse event observation period

- Pre-treatment adverse events are defined as any adverse event reported during the pre-treatment period.
- Treatment-emergent adverse events (TEAEs) are adverse events that developed or worsened or became serious during the treatment period.
- Post-treatment adverse events are adverse events that developed or worsened or became serious during the post-treatment period.

All AEs (including SAEs and adverse events of special interest (AESI)) will be graded according to National cancer institute common terminology for adverse events (NCI-CTCAE) version 4.03 and coded to a lower-level term (LLT), preferred term (PT), high-level term (HLT), high-level group term (HLGT) and associated primary system organ class (SOC) using the version of Medical Dictionary for Regulatory Activities (MedDRA) currently in effect at Sanofi at the time of database lock.

AESI include the following terms:

- Pregnancy of female patient entered in a study as well as pregnancy occurring in a female partner of a male subject entered in a study with IMP/NIMP.
- Symptomatic overdose (serious or non-serious) with IMP/NIMP.

Primary safety endpoint for the Part 1 (predefined safety criteria):

Incidence of IMP-related predefined safety criteria during Cycle 1 and Cycle 2.

Predefined safety criteria are defined at the occurrence of any following related TEAE using NCI CTCAE v4.03:

- Grade ≥3 TEAE from the System Organ Class (SOC) Eyes disorders,
- Grade ≥3 peripheral neuropathy (Preferred Term [PT]),
- Grade ≥4 TEAE.

Adverse Events leading to specific analyses:

Specific analyses will be performed for the following AEs:

- Pre-specified safety criteria (cycle 1 and cycle 2).
- Corneal events (reported as AEs) as defined in the grouping GLB_CORNEAL_DISORDERS.

The last event will be identified from a grouping of adverse events based on MedDRA preferred terms.

2.1.4.2 Ocular examination

Specific ocular evaluations are planned at baseline and will be repeated at the time of the occurrence in ocular toxicity. Moreover tonometry will be performed at screening, end of cycle 6 and end of treatment.

2.1.4.3 Deaths

The deaths observation periods are per the observation periods defined above.

- Death on-treatment: deaths occurring during the TEAE period.
- Death post-treatment: deaths occurring during the post-treatment period.

2.1.4.4 Laboratory safety variables

Laboratory data consists of blood analysis, including hematology and biochemistry.

Blood samples for clinical laboratories will be taken at Day 8, Day 15 and Day 21 during the first two cycles (except for INR only at Day 21) and then at Day 8 and Day 21 for hematology and Day 21 for clinical chemistry and INR at subsequent cycles and then at end of treatment. The laboratory parameters will be classified as follows:

- Hematology:
  - **Red blood cells (RBC) and platelets and coagulation**: hemoglobin, platelet count, international normalized ratio (INR).
  - **White blood cells (WBC):** WBC with differential (neutrophils, lymphocytes, monocytes, basophils, eosinophils), absolute neutrophil count (ANC).

- Clinical chemistry:
  - **Metabolism:** blood glucose, lactate Dehydrogenase (LDH), total protein, albumin,
  - **Electrolytes**: sodium, potassium, chloride, calcium, corrected serum calcium (calculated according to Corrected calcium formula, see Section 2.5.1), phosphate,
  - **Renal function**: serum creatinine, creatinine clearance (calculated according to both Cockroft-Gault formula, see Section 2.5.1), urea or blood urea nitrogen (BUN),
  - **Liver function**: alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase, total and direct bilirubin.

Moreover, serum pregnancy test will be performed at baseline, during on-treatment period and end of treatment period for women of child bearing potential.

Laboratory values will be analyzed after conversion into standard international units by data management. These converted values will be graded according to the NCI-CTCAE version 4.03 whenever applicable, using laboratory normal ranges (liver function tests, creatinine, creatinine clearance, LDH) provided by the laboratory analyzing the sample whenever possible or using generic normal ranges for other parameters (see Appendix G). When the NCI-CTCAE is not applicable, laboratory values "out of normal ranges" will be summarized.

For each laboratory parameter, each assessment performed from the 2nd day of cycle N to the 1st day of cycle N+1 will be assigned to cycle N. The 1st day of any cycle will be the date of the 1st study treatment infusion for this cycle.

Parameters measured on the day of the first study treatment infusion (Cycle 1 Day 1) will be considered as baseline measurements. If missing, the baseline measurement will be the last available value before first study treatment infusion. For a given parameter, a patient will be considered as evaluable if at least one measure of this parameter is available during the on-treatment period.

Technical formulas are described in Section 2.5.1.

2.1.4.5 Vital signs variables

Vital signs include blood pressure (systolic and diastolic), BSA and body weight. Moreover ECOG performance status will be also provided. All of these parameters will be performed at screening, Days 8, 15 and 21 in Cycles 1 and 2 and Days 8 and 21 in subsequent cycles and then at end of treatment visit. ECOG PS and body weight will be also assessed in follow-up.

Parameters measured on the day of the first study treatment infusion (Cycle 1 Day 1) will be considered as baseline measurements. If missing, the baseline measurement will be the last available value before first study treatment infusion. For a given parameter, a patient will be considered as evaluable if at least one measure of this parameter is available during the on-treatment period.

2.1.4.6 Electrocardiogram variables

ECGs will be recorded automatically by the device at the Investigator site at screening and end of treatment visit.

Electrocardiogram assessments will be described as normal or abnormal.

For ECG, parameters measured during the screening period will be considered as baseline measurements.

2.1.5 Pharmacokinetic variables

Pharmacokinetic samples of SAR566658, unconjugated maytansinoids (DM4 and Me-DM4) and naked hu-DS6 will be collected at specific time points of each cycle, and at the end of treatment and follow-up visits as specified in the PK/immunogenicity flow chart of the protocol.

Main PK parameters that will be provided should include C_{max}, AUC, total clearance, distribution volume, and half-life.

2.1.6 Immunogenicity endpoints

Blood samples must be taken on Day 1 before the first study treatment infusion of each treatment cycle, then at the end of study treatment (Day 30) and at follow up (Day 60). For each time point, results are defined as negative, positive or inconclusive according to circulating levels of SAR566658.

Periods of observation:

- Antitherapeutic antibodies (ATA) pre-treatment period: The ATA pre-treatment period is defined as the time from signed informed consent to the first SAR566658 infusion administration.
- ATA on-study observation period: the ATA on-study observation period is defined as the time from first SAR566658 infusion administration until the end of the study.

ATA attributes:

- **Pre-existing ATAs** are defined as ATAs that were present in samples drawn during the pretreatment period.
- **Treatment boosted ATAs** are defined as pre-existing ATAs with an increase in titer value between pre-treatment and post-treatment samples of at least two titer steps. Assuming a 2-fold serial dilution schema is used for the study, this means that the post-treatment sample titer value is at least (≥) 4 fold of pretreatment titer value.
- **Treatment induced ATAs** are defined as ATAs that developed at any time during the ATA on-study observation period in patients without pre-existing ATAs, including

patients without pretreatment samples. If baseline sample is missing, baseline is considered by default as ATA negative, and thus, ATA is considered as treatment-induced.

- **Transient ATA response** is defined by:
  - Treatment-induced ATA detected only at one sampling time point during the ATA on-study observation period (excluding the last sampling time point), OR
  - Treatment-induced ATA detected at two or more sampling time points during the treatment (including follow-up period if any), where the first and last ATA -positive samples (irrespective of any negative samples in between) are separated by a period less than 16 weeks and the subject's last sampling time point is ATA -negative.
- Persistent ATA response is defined by:
  - Treatment-induced ATA detected at two or more sampling time points during the ATA on-study observation period, where the first and last ATA -positive on-study samples are separated by a period of 16 weeks or longer (irrespective of any negative samples in between), OR
  - Treatment-induced ATA detected in the last two sampling time points (both positive), irrespective of the time period in between.
- **Indeterminate ATA** is defined by treatment induced ATA that is not persistent or transient (only the last sampling time is positive).

Among evaluable population for immunogenicity (see Section 2.3.5 for definition), following patients will be defined:

- **ATA positive patients**, which are defined as patients with at least one treatment-induced or treatment-boosted ATA positive sample at any time during the on-study observation period;
- **ATA negative patients**, which are defined as patients without treatment-induced nor treatment-boosted ATA positive sample during the on-study observation period;
- **ATA inconclusive patients**, which are defined as patients which cannot irrefutably be classified as ATA negative.

Two main categories can be reported for the epidemiology of an ATA immune response: ATA prevalence and ATA incidence:

- ATA prevalence defines the proportion of all patients tested positive for ATAs (including preexisting antibodies, treatment boosted ATA and treatment induced ATA) at any point in time.
- In contrast, the term ATA incidence only defines the proportion of patients found to either have seroconverted (treatment induced ATAs) or boosted their pre-existing ATA response during the study.

The incidence will be defined as follows:

$$ATA\,incidence = 100 \times \frac{treatment\ boosted + treatment\ induced\ \ ATA\ \ positive\ \ patients}{number\ of\ \ evaluable\ \ patients}$$

$$Treatment\ induced\ \ ATA\,incidence = 100 \times \frac{treatment\ induced\ \ ATA\ \ positive\ \ patients}{number\ of\ \ evaluable\ \ patients\ \ ATA\ negative\ at\ baseline}$$

$$Treatment\ boosted\ \ ATA\,incidence = 100 \times \frac{treatment\ boosted\ \ ATA\ \ positive\ \ patients}{number\ of\ \ evaluable\ \ patients\ \ ATA\ \ positive\ at\ baseline}$$

2.1.7 Biomarkers

2.1.7.1 Main biomarkers

CA6 expression level in tumor tissue samples by IHC

IHC will be the reference assay to assess CA6 expression level and will be performed at a central laboratory. At prescreening, determination of CA6 positivity in tumors will be performed by IHC on FFPE slides collected from the most recent available tumor samples. A total of 6 4-micron sections cut from the most recent available FFPE tumor block (ie, archive tumor tissue at diagnosis, archive tumor tissue at surgery, or most recent tumor sample) will be collected and sent to a centralized laboratory for CA6 staining by muDS6-IHC and for appropriate control. Staining will be analyzed by a dedicated pathologist. Stained and unstained slides will be kept for up to 5 years from the completion of the clinical study.

Both the membrane and cytoplasmic staining will be recorded for the percent of tumor cells positive at all intensities (1+, 2+, and 3+).

The following variables will be derived based on the percentage of positive cells at all intensities (1+, 2+, and 3+):

- Percentage of positive cells at intensity ≥ 2+ on the sum of membrane, quantitatively and by categories.
- H-score on the sum of membrane (either whole or polarized) staining. H-score will be defined by ([% positive cells at intensity 0] x 0 + [% positive cells at intensity 1+] x 1 + [% positive cells at intensity 2+] x 2 + [% positive cells at intensity 3+] x 3).
- Main analyses will be performed on membrane staining but descriptive analyses on cytoplasmic staining could eventually be explored.

Circulating CA6 from blood samples

A blood sample will be collected at baseline prior to the first treatment administration and plasma will be extracted on site. Plasma will be stored at -70°C or below and transferred on a planned schedule to a centralized laboratory until their authorized transfer to a testing laboratory, where circulating CA6 levels will be determined. Circulating CA6 will be determined by a semiquantitative sandwich enzyme-linked immunosorbent assay (ELISA).

2.1.7.2 Exploratory pharmacodynamic biomarker

[redacted]

2.1.7.3 Other exploratory biomarkers

[redacted]

[redacted]

2.1.8 Post-treatment therapy after discontinuation of investigational medicinal product administration during the study

Post treatment anti-cancer therapies include post treatment systemic anti-cancer treatment (including chemotherapy, hormonal therapy, immunotherapy…), post treatment radiotherapy (other than given in palliative intent) and post treatment surgery related to cancer.

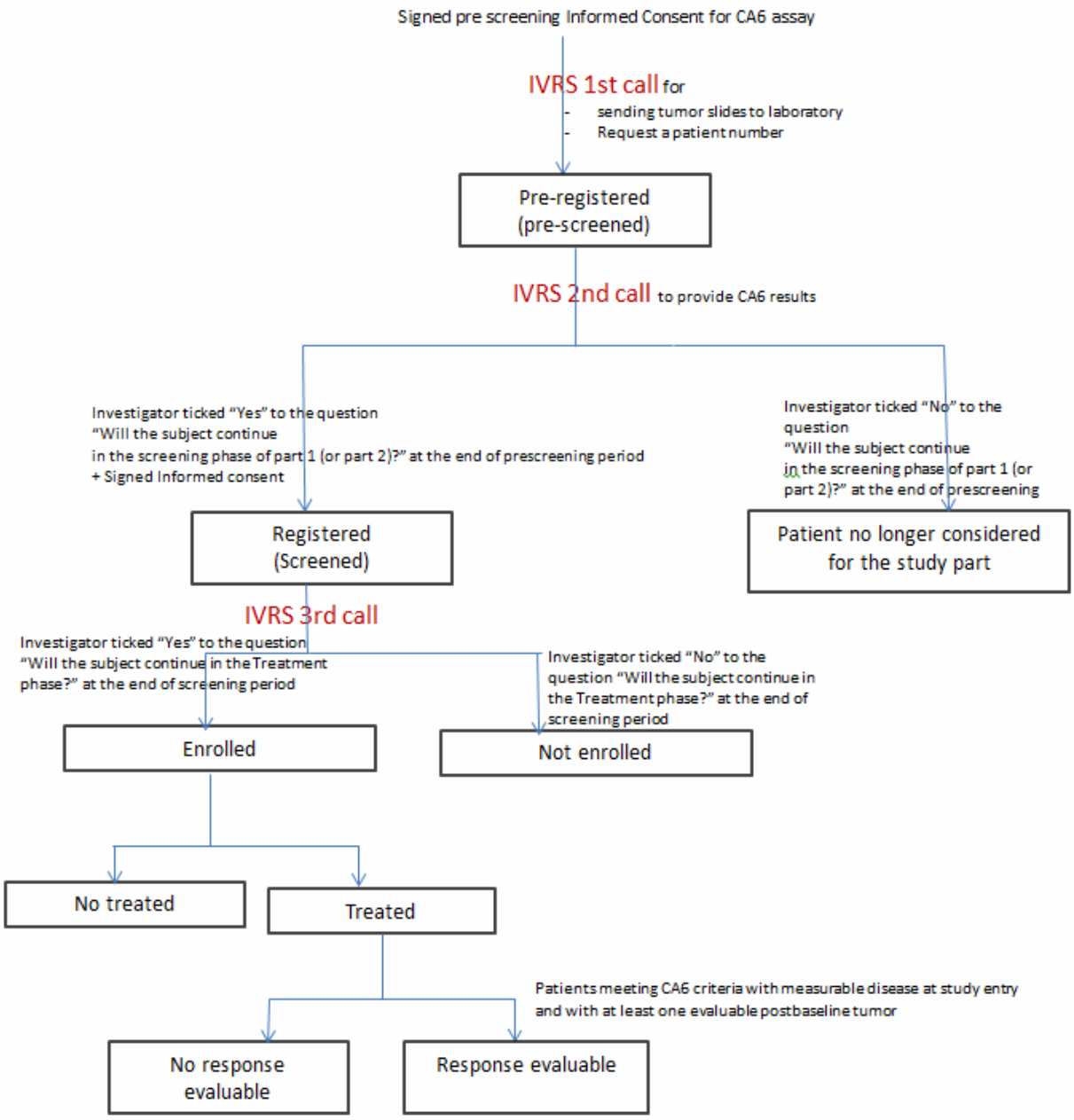
Palliative radiotherapy may be given for control of pain for palliative intents during treatment period. This will not be considered as post treatment anti-cancer therapy and thus will be documented separately.

2.2 DISPOSITION OF PATIENTS

This section describes patient disposition for both patient study status and the patient analysis populations.

An overview of the analysis populations is presented in the Figure 1.

Figure 1 : Analysis population



Pre-registered (pre-screened) patients are defined as patients who signed the prescreening informed consent for CA6 assay assessment of their archival biopsy

Registered (screened) patients are defined as patients who signed the informed consent for study participation and are planned to enter in the screening period, i.e. for whom the investigator ticked "Yes" to the question "Will the subject continue in the screening phase of part 1 (or part 2)?" at the end of prescreening period.

Enrolled patients includes registered patients who are planned to enter in the study period, ie, for whom the investigator ticked "Yes" to the question "Will the subject continue in the Treatment

phase?" at the end of screening period. In part 1, patients also need a confirmation of a successful allocation of a randomization number through the IVRS/IWRS in order to be randomized.

The total number of patients in each of the following analyses population will be presented in the clinical study report using a flowchart diagram or a summary table:

- Pre-registered population
- Registered population
- Not enrolled but treated population
- Enrolled population
- Enrolled but not treated population
- All-treated population
- Evaluable for predefined safety criteria population
- Response evaluable population

For patient study status, the total number of patients in each of the following categories will be presented in the clinical study report using a flowchart diagram or a summary table:

- Enrolled patients
- Enrolled and not treated patients with reason for non-treatment
- All-treated population with number of patients who discontinued study treatment by main reason for treatment discontinuation
- Status at last study contact: alive, dead

For all categories of patients (except for the pre-registered, registered and not enrolled patients) percentages will be calculated using the number of enrolled patients as the denominator.

All critical or major deviations will be summarized in tables giving numbers and percentages of deviations.

2.2.1 Randomization irregularities

Randomization irregularities occur whenever a randomization is not in accordance with the protocol-defined randomization method, such as a) an ineligible patient is randomized, b) a patient is randomized based on an incorrect stratum, c) a patient is randomized twice.

All randomization irregularities will be documented in the clinical study report. If the number of irregularities is large enough to make a tabular summary useful, the irregularities will be categorized and summarized among the treated patients (number and percentages). Nonrandomized/non enrolled but treated patients will be described separately. Listings with additional, relevant details will be provided in an appendix, eg, excluded second randomization/enrollment study data for patient randomized/enrolled twice.

Randomization irregularities to be prospectively identified include but are not limited to:

Randomization irregularities
Randomization by error
Patient randomized twice
Stratification error
CA6 level expression error
Patient switched to another site

2.2.2 Important deviations potentially impacting efficacy

The summary of important protocol deviations will be based on the all treated population. Other deviations will not be summarized in the Clinical Study Report.

Important protocol deviations are based on inclusion/exclusion criteria and are defined as follows:

- No metastatic TNBC (*I01*),
- Patient with no measurable disease as per RECIST 1.1 (I02),
- No histologically documented TNBC (I03),
- Not the corresponding CA6 membrane expression level according the threshold defined in the study part *(I04),*
- No prior chemotherapy regimen or more or equal to 4 for advanced/metastatic disease *(I05),*
- No prior anticancer therapy with anthracycline and a taxane in an adjuvant/neoadjuvant or metastatic setting *(I06),*
- Prior treatment with eribulin as last prior therapy or prior maytansinoid treatments (*E13*).

2.3 ANALYSIS POPULATIONS

Patients treated without being randomized (missing the third call IVRS in the Part 1) will not be considered randomized and will not be included in any efficacy or all treated population.

For any patient randomized or enrolled more than once, only the data associated with the first randomization will be used in any analysis population. The safety experience associated with any later randomization will be assessed separately.

2.3.1 Efficacy populations

The primary efficacy analysis will be performed on the all-treated population.

The response evaluable population will be used for sensitivity efficacy analyses of the BOR and the DCR.

2.3.1.1 All treated population

The all-treated population (AT) consists of enrolled patients who will actually receive at least one dose or any partial of a dose of study treatment.

2.3.1.2 Response evaluable population

The response evaluable population will be defined as the patients of the all treated population with measurable disease at study entry and with CA6 positivity according to the threshold defined in the relevant study part (Part 1 and Part 2a or Part 2b) and who had an evaluable response assessment during treatment (until 30 days after last SAR566658 administration). Treated patients who could not have a postbaseline tumor assessment because of early disease progression or early death due to progressive disease are also included in the response evaluable population.

An evaluable response assessment is defined as a response assessment with a response different of NE.

2.3.2 Safety population

The population used for safety analyses consists of the all-treated population. Any safety analysis will be performed on this population

2.3.3 Evaluable for predefined safety criteria population (Part 1)

Patients evaluable for predefined selection criteria will be defined as patients treated who have completed 2 cycles or who experienced predefined safety criteria. To complete 2 cycles, patients have to receive 4 infusions and more than 80% of the cumulative planned dose. This population is only defined for Part 1 for the purpose of the Bayesian design and the analysis of the predefined safety criteria.

2.3.4 Pharmacokinetic population

The PK population will include all patients who received at least 1 dose of SAR566658, even if incomplete, and have at least 1 valid PK parameter available.

2.3.5 Evaluable population for immunogenicity

The evaluable population for immunogenicity will include all patients with at least 1 sample drawn postbaseline after study drug infusion (on-treatment period or follow-up) that is appropriate for ATA testing with a reportable result (patients with missing ATA at baseline will be considered evaluable).

2.4 STATISTICAL METHODS

Continuous data will be summarized using the number of available data, mean, standard deviation, median, minimum, Q1, Q3 and maximum. Categorical and ordinal data will be summarized using the number and percentage of patients.

All the tables will be presented by the intended dose level in the Part1 and by actual expression cohort at the selected dose in the Part 2 (mild expression or overexpression cohorts). Separate tables will be provided for Part 1 and Part 2. Summary table including both dose levels and expression cohorts will be provided for some analyses as demographics or analysis population topics.

2.4.1 Demographics and baseline characteristics

The parameters described in Section 2.1.1 will be summarized on the all-treated population analyzed in the planned dose level using descriptive statistics.

The medical and surgical history will be summarized according to the SOC and PT. Separate tables will summarize ocular history, peripheral neuropathy history.

For the safety parameters, the baseline values will be described along with each safety analysis if relevant.

2.4.2 Prior, concomitant and post treatment medications (other than anticancer therapies)

The prior, concomitant and post-treatment medications will be presented for the all-treated population.

Medications will be summarized according to the WHO-DD dictionary, considering the first digit of the anatomic category (ATC) class (anatomic category) and the first 3 digits of the ATC class (therapeutic category). All ATC codes corresponding to a medication will be summarized, and patients will be counted once in each ATC category (anatomic or therapeutic) linked to the medication. Therefore patients may be counted several times for the same medication.

The table for prior medications will be sorted by decreasing frequency of ATC followed by all other therapeutic classes. In case of equal frequency regarding ATCs alphabetical order will be used. In part 1, sorting will be based on results for the dose level 120 mg/m^2 D1D8 Q3w and in part 2, sorting will be based on the All column including overexpression and mild expression cohorts at the selected dose.

The tables for concomitant and posttreatment medications will be sorted in the same way.

Premedications

Number (%) of patients compliant with premedications for ocular toxicities as defined in Section 2.1.2 will be provided.

Compliance will be calculated for each administered infusion. It will be a binary variable, patient will be considered as compliant or as not compliant for the whole infusion. Total compliance will be calculated as the mean compliance over all administered infusions.

2.4.3 Extent of investigational medicinal product exposure and compliance

The extent of study treatment exposure will be assessed and summarized on the all-treated population.

2.4.3.1 Extent of investigational medicinal product exposure

The overall extent of exposure will be assessed for each patient as:

- Actual number of weeks on-treatment (duration of treatment) defined as (date of last cycle first infusion – date of first cycle first infusion + 21)/7;
- Total number of cycles started (K);
- Total number of infusions started

Overall study treatment exposure will be summarized by the total number of infusions started, the total number of cycles started, the total number of infusions received by patient, the total number of cycles received by patient and the treatment duration

Following information will also be summarized:

- Cumulative dose (mg/m^2): the cumulative dose at cycle k is the sum of all doses from cycle 1 to and including cycle k, where k is based on investigator's report.

$$\text{Cumulative Dose} = \sum_{k=1}^{K} \text{Actual Dose in Cycle}[k]$$

- Actual dose intensity (mg/m^2/week):

$$\text{ADI} = \frac{\text{Cumulative dose } (mg/m^2)}{\text{Actual no. of weeks on treatment}}$$

- Planned dose intensity (mg/m^2/week):

$$\text{PDI} = \frac{\text{Total Planned Dose } (mg/m^2)}{\text{Planned no. of weeks on treatment}}$$

The total planned dose will be calculated by the sum of the theoretical planned dose (i.e., the planned dose as per protocol) at each cycle:

$$\text{Total Planned Dose} = \sum_{k=1}^{K} \text{Planned Dose Cycle}[k]$$

Where K=total number of treatment cycles received by the patient

The planned number of weeks on treatment is defined as 3*K.

As per protocol, the planned dose at each infusion at Day 1 and Day 8 is 90 mg/m2 or 120 mg/m2 in each cycle in the Part 1 and the selected dose in the Part 2.

- The relative dose intensity (RDI) is defined as

$$\text{Relative Dose Intensity} = \frac{\text{Actual Dose Intensity}}{\text{Planned Dose Intensity}}$$

Summary statistics will be provided for cumulative dose, actual dose intensity and relative dose intensity.

As per protocol, for patients with a BSA >2.0 m^2, a BSA equals to 2.0 m^2 will be used for the calculation of the dose in mg/m^2.

Dose modifications:

Dose reduction: dose reduction will be derived using the definition provided in Table 2 compared to the previous dose. For the second and subsequent infusions, a dose is deemed to have been reduced if the dose level a patient received differs from the previous actual dose level. If a dose is omitted at Day 8, then the dose reduction will be derived from the dose received at Day 1 from the previous cycle. The first infusion will not be counted for a dose reduction.

Table 2 – SAR566658 actual dose level definition (at each infusion)

Actual dose level (mg/m²)	Actual dose administered (mg/m²)
Upper dose	135 ≤ Dose
120	105 ≤ Dose < 135
90	83 ≤ Dose < 105
75	68 ≤ Dose <83
60	52 ≤ Dose < 68
Lower dose	0 < Dose <52
No dose administered	0

Dose omission: at a given infusion, dose omission will be considered if no drug is administered (dose=0 mg/m²). Number of dose omissions within a cycle will be equal to the difference between expected number of infusions and number infusions actually received.

Infusion delayed: a dose at Day 22 (Day 1 of next cycle) is deemed to have been delayed if the study treatment is ≥4 days beyond the theoretical day of treatment. The first dose will not be counted for dose delay.

Dose interruption: dose interruption will be considered if the study treatment infusion is temporarily stopped during the infusion.

Number of patients (%) and number of cycles (%) with at least one SAR566658 dose modification (including dose reduction, dose omission, cycle delayed) and dose interruption, overall and by type, will be provided.

Number of SAR566658 infusions (%) with a dose modification, overall and by type, will be also provided.

2.4.4 Analyses of efficacy endpoints

2.4.4.1 Analysis of primary efficacy endpoint(s)

Response rate will be analyzed based on all-treated population as defined in the Section 2.3.1.1.

Overall Response rates will be summarized using counts (n) and percentages (%) and presented with its 2-sided 80% confidence interval using the Clopper and Pearson exact method.

At the end of the Part 2a in the overexpression cohort at the selected dose, an exact binomial test of the null hypothesis that the true response rate is 12% will be performed using a one-sided 0.1 alpha level.

Best Overall Response (BOR) (CR, PR, SD, PD, and NE) will be summarized using counts (n) and percentages by treatment group. At the end of the part1, a summary of BOR by intended dose levels and ECOG stratification factor will also be provided.

For responder patients, a listing with number of prior advanced regimen, prior treatment with anthracycline and/or taxanes (yes/no), prior treatment with eribulin (yes/no), best response to last advanced previous anti therapy, baseline CA6 IHC, baseline circulating CA6, best overall response to SAR566658, duration of response and reason for treatment discontinuation will be provided.

Exploratory analyses

At the end of the Part 2a in the overexpression cohort, ORR will be estimated (with its 80% CI) in several subgroups of patients defined in the table below. A forest plot with estimation of ORR in each subgroup will be provided, if enough patients in each stratum. Depending upon the study results, additional subgroups may be examined, and subgroups with small sample sizes may be pooled to create a larger meaningful subgroup.

Table 3 – Subgroup analyses

Prognostic factor	Description
Age	[18-65[y, [65-75[y, ≥75y
ECOG PS at baseline	0, 1
Any prior eribulin agent	Yes vs No
Mutation BRCA1	Yes vs No
Mutation BRCA2	Yes vs No
Number of organs involved	1,2 ≥3
Number of prior regimen	1,2,3
Bone metastasis only vs visceral metastasis	Yes vs No
DFS	≤24 mo vs >24mo
PD1	Classes to be defined
PDL1	Classes to be defined

Sensitivity analyses of the primary efficacy endpoint

The ORR will also be estimated on the response-evaluable population as defined in Section 2.3.1.2. Depending on the results, exploratory analyses as defined in the Table **3** may also be performed on the response-evaluable population.

2.4.4.2 Analyses of secondary efficacy endpoints

2.4.4.2.1 Disease Control rate

Disease control rate will be summarized using counts (n) and percentages by treatment group. Confidence intervals at 80% will be computed using the Clopper and Pearson exact method.

The DCR will also be estimated on the response-evaluable population as defined in Section 2.3.1.2.

2.4.4.2.2 Duration of response

Duration of response will be analyzed in responders using the Kaplan-Meier method. Kaplan-Meier estimates and associated 80% CI will be provided.

2.4.4.2.3 Tumor shrinkage

A graphical representation of the best tumor shrinkage observed per patient (waterfall plot) will also be provided on the all-treated population.

2.4.4.2.4 Time to progression

The median TTP and its 80% confidence interval will be estimated on the all-treated population.

Kaplan-Meier estimates of the 25th, 50th, and 75th percentiles and their associated 80% CIs will be provided.

Number of patients at risk as well as the probabilities of surviving without disease progression at least 3, 6, 9, and 12 months with 80% confidence intervals will be estimated using the Kaplan-Meier method.

The number (%) of censored patients and the timing of their censoring (ie, censored at randomization, censored at last disease assessment) will be summarized. Kaplan-Meier curves will be plotted. These plots will include the number of patients at risk at key time points.

2.4.4.2.5 Progression Free Survival

PFS data will be analyzed using the Kaplan-Meier method in the all treated population:

- Kaplan-Meier estimates of the 25th, 50th, and 75th percentiles and their associated 80% CIs will be provided. The 80% confidence intervals will be constructed using a log-log transformation of the survival function and the methods of Brookmeyer and Crowley.
- Number of patients at risk as well as the probabilities of surviving without disease progression at least 3, 6, 9, and 12 months with 80% confidence intervals will be estimated using the Kaplan-Meier method.
- Kaplan-Meier curves will be plotted. These plots will include the number of patients at risk at key time points.

For patients with events, the type of event (objective disease progression or death for primary analysis and also clinical/non radiological progression for sensitivity analysis) will be summarized using counts and percentages. The type of disease progression will be presented for sensitivity analysis (objective disease progression or clinical/non radiological progression).

For patients who died without evidence of disease progression, the time from the last disease assessment to death will be summarized using number, mean, standard deviation, median, and range.

The number (%) of censored patients and the timing of their censoring (ie, censored at randomization, censored at last disease assessment) will be summarized.

2.4.4.3 Multiplicity issues

No adjustment for multiplicity linked to the selection of dose will be done.

2.4.5 Analyses of safety data

The summary of safety results will be presented by dose levels in the Part 1 and by CA6 expression level cohort in the Part 2 (mild expression and overexpression).

General common rules

All safety analyses will be performed on the all-treated population as defined in Section 2.3.1.1, unless otherwise specified, using the following common rules:

- Safety data in patients who do not belong to the all-treated population (eg, treated but not randomized) will be listed separately.
- The baseline value is defined as the last available value before first study treatment infusion.
- The analysis of the safety variables will be essentially descriptive.

2.4.5.1 Analyses of adverse events

Generalities

The primary focus of adverse event reporting will be on treatment-emergent adverse events as defined in Section 2.1.4. Pretreatment and post-treatment adverse events will be described separately.

If an adverse event date of onset (occurrence or becoming serious) is incomplete, an imputation algorithm will be used to classify the adverse event as pretreatment, treatment-emergent, or post-treatment. The algorithm for imputing date of onset will be conservative and will classify an adverse event as treatment emergent unless there is definitive information to determine it is pretreatment or post-treatment. Details on classification of adverse events with missing or partial onset dates are provided in Section 2.5.4.

The grade will be taken into account in the summary. For patients with multiple occurrences of the same event, the maximum (worst) NCI grade is used within the same period of observation (pretreatment, treatment, post-treatment). Summaries will be provided for all grades and for grade ≥3 (including grade 5). Missing grades, if any, will be included in the "all grades" category. Sorting within tables ensures the same presentation for the set of all adverse events within the observation period (pretreatment, treatment-emergent and post-treatment). For that purpose, the table of all treatment-emergent adverse events presented by SOC and PT sorted by the internationally agreed SOC order and decreasing frequency of PTs within SOCs will define the presentation order for all other tables unless otherwise specified. In part 1, sorting will be based on results for the dose level 120mg/m2 D1D8 Q3w and in part 2, sorting will be based on the All column including overexpression and mild expression cohorts at the selected dose.

Analysis of all treatment-emergent adverse events

The following treatment-emergent adverse event summaries will be generated for the all-treated population. Overview of treatment-emergent adverse events, summarizing number (%) of patients with any

- TEAE
- Grade ≥3 TEAE

- Grade 3-4 TEAE
- Grade 5 TEAE (any TEAE with a fatal outcome during the treatment period)
- Serious TEAE
- TEAE leading to permanent treatment discontinuation

In addition, an overview of Grade 5 AEs will be provided.

The following frequency distributions of AEs (incidence tables) will also be provided for the all-treated population, for all grades combined and grade ≥3.

- All treatment-emergent adverse events by primary SOC, HLGT, HLT, and PT, showing number (%) of patients with at least 1 TEAE sorted by the SOC internationally agreed order. The other levels (HLGT, HLT, PT) will be presented in alphabetical order.
- All treatment-emergent adverse events by primary SOC and PT, showing the number (%) of patients with at least 1 treatment-emergent adverse event, sorted by the internationally agreed SOC order and by decreasing incidence of PTs within each SOC as previously defined. This sorting order will be applied to all other tables, unless otherwise specified.
- All related treatment-emergent adverse events by primary SOC, HLGT, HLT, and PT, showing number (%) of patients with at least 1 related TEAE sorted by the internationally agreed SOC order. The other levels (HLGT, HLT, PT) will be presented in alphabetical order.
- All related treatment-emergent adverse events by primary SOC and PT, showing the number (%) of patients with at least 1 related TEAE, sorted by the sorting order defined above.

Analysis of all treatment emergent serious adverse event(s)

All serious TEAEs will be presented by primary SOC and PT, showing the number (%) of patients with at least 1 serious TEAE, sorted by the internationally agreed SOC order and by decreasing incidence of PTs within each SOC.

All treatment-emergent serious adverse events will be presented by primary SOC, HLGT, HLT, and PT, showing the number (%) of patients with at least 1 serious treatment-emergent adverse event, sorted by the internationally agreed SOC order. The other levels (HLGT, HLT, PT) will be presented in alphabetical order.

All treatment-emergent serious adverse events regardless of relationship and related to IMP will be presented by primary SOC, HLGT, HLT, and PT, showing the number (%) of patients with at least 1 serious treatment-emergent adverse event, sorted by the internationally agreed SOC order. The other levels (HLGT, HLT, PT) will be presented in alphabetical order.

Analysis of all treatment-emergent adverse event(s) leading to treatment discontinuation

All treatment-emergent adverse events leading to permanent treatment discontinuation will be presented by primary SOC, HLGT, HLT, and PT, showing the number (%) of patients sorted by the internationally agreed SOC order. The other levels (HLGT, HLT, PT) will be presented in alphabetical order.

All TEAE leading to treatment permanent discontinuation will be presented by primary SOC and PT, showing the number (%) of patients with at least 1 TEAE leading to treatment discontinuation, sorted by the internationally agreed SOC order.

Analysis of all treatment-emergent adverse event(s) leading to dose modification

The following summary tables are based on the investigator's intent reported in the AE page ("action taken"):

- All treatment-emergent adverse events leading to dose modification (delay and/or reduction/omission), by primary SOC and PT, showing the number (%) of patients, sorted by the sorting order defined above.
- All treatment-emergent adverse events leading to dose reduction/omission, by primary SOC and PT, showing the number (%) of patients, sorted by the sorting order defined above.
- All treatment-emergent adverse events leading to dose interruption, by primary SOC and PT, showing the number (%) of patients, sorted by the sorting order defined above.
- All treatment-emergent adverse events leading to dose delay, by primary SOC and PT, showing the number (%) of patients, sorted by the sorting order defined above.

Analysis all treatment-emergent adverse events(s) with fatal outcome

All TEAEs with fatal outcome will be presented by primary SOC, HLGT, HLT and PT, showing the number (%) of patients sorted by the internationally agreed SOC order. The other levels (HLGT, HLT, PT) will be presented in alphabetical order

Analysis of adverse events of special interest

Number (%) of patients with at least one AESI cited in Section 2.1.4 will be presented.

Analysis of ocular toxicities and peripheral neuropathy

For each TEAE, the number (%) of patients by worst grade, the cycle of first occurrence regardless of the grade, the cycle of first occurrence of the worst grade, the relationship to study treatment (in case of several occurrences with different relationships, if any event is related, then the relationship will be 'Related'), the action taken with the study treatment (in case multiple occurrences with different actions, the most severe action taken will be tabulated and selected according to the following order of severity: drug withdrawn, dose reduced and delayed, dose reduced/omitted, dose delayed, drug interruption, dose not changed, not applicable), the outcome

(in case of multiple occurrences with different outcomes, the most severe outcome will be tabulated and selected according to the following order of severity: fatal, not recovered or resolved, recovering or resolving, recovered or resolved with sequelae, recovered or resolved, unknown) and the number of recurrences.

Analysis of pre-treatment and post-treatment adverse events

Overview of pre-treatment adverse events, summarizing number (%) of patients with any

- Pre-treatment adverse events
- Grade ≥3 pre-treatment treatment adverse events
- Grade 3-4 pre-treatment adverse events
- Serious pre-treatment adverse events
- Grade 5 pre-treatment adverse events

Overview of post-treatment adverse events, summarizing number (%) of patients with any

- Post-treatment adverse events
- Grade ≥3 post-treatment treatment adverse events
- Grade 3-4 post-treatment treatment adverse events
- Grade 5 post-treatment adverse events (including TEAE worsening to Grade 5 during the post-treatment period)
- Serious post-treatment adverse events

All pre-treatment AEs by primary SOC and PT, showing the number (%) of patients with at least 1 pre-treatment AE, sorted by the internationally agreed SOC order and decreasing incidence of PTs within each SOC.

All post-treatment AEs by primary SOC and PT, showing the number (%) of patients with at least 1 post-treatment AE, sorted by the internationally agreed SOC order and decreasing incidence of PTs within each SOC

Some listings might also be reported if judged more relevant.

2.4.5.2 Deaths

The following summaries of deaths will be generated for the all-treated population:

- Number (%) of patients who died by study period (on-treatment, post-treatment as defined in Section 2.1.4.3) and reasons for death by treatment received.
- Deaths in nonrandomized patients or randomized but not treated patients. A listing may be provided.

- Summary of AEs leading to death, by Primary SOC and PT
  - in context of disease progression (death within 30 days from last study treatment administration and for whom cause of death is disease progression),
  - in context other than PD (death within 30 days from last study treatment administration and for whom cause of death is different from disease progression, or death more than 30 days from last study treatment administration and for whom cause of death equals to adverse event)

Some listings might also be reported if judged more relevant.

2.4.5.3 Analyses of ocular examinations

Number (%) of visual symptoms will be provided.

During the treatment, change from baseline will be summarized.

For slit lamp examination, patients with abnormal or normal status will be provided. For patients with abnormal status, number of patients with corneal mycrocysts, corneal deposits, corneal erosion, corneal ulcer, and corneal inflammation at one of the eyes will be described. Number of patients with clinically significant findings will be reported with corresponding adverse event.

For Schirmer's test, descriptive statistics on the result will be provided. Patients will be classified into four classes with anesthetics: normal ≥15 mm, mild 14-9 mm; moderate 8-4 mm; severe <4 mm and ocular adverse events for the classification severe at baseline will be reported.

For tonometry test, number (%) of patients with abnormal/normal status with and without dilation at baseline will be provided. During the treatment, descriptive statistics of the change from baseline without dilation for each eye will be provided.

2.4.5.4 Analyses of laboratory variables

Hematological and clinical biochemistry toxicities will be assessed from laboratory test parameters defined in Section 2.1.4.4.

The frequency of patients in each grade of laboratory test during treatment will be summarized. For patients with multiple occurrences of the same laboratory variable during the treatment, the maximum grade (worst) per patient will be used.

The denominator used for percentage calculation is the number of patients with at least 1 evaluation of the laboratory test during the considered observation period.

The summary table will present the frequency of patients with any grade of abnormal laboratory tests and with Grade 3-4 abnormal laboratory tests.

For the non-NCI gradable parameters (monocytes, eosinophils, basophils, chloride, total protein, BUN, LDH, urea, direct bilirubin), frequency of evaluable patients outside normal ranges will be provided.

Regarding calculated creatinine clearance, a summary of the lowest value recorded by patient will be provided using summary statistics. Frequencies by class (≥90 mL/min, ≥60 mL/min and <90 mL/min, ≥30 mL/min and <60 mL/min, ≥15 mL/min and <30 mL/min, <15 mL/min) will also be provided.

Moreover, a listing of possible Hy's law cases identified (eg, patients with any elevated AST or ALT of >3 ULN **and** elevated Total bilirubin >2 ULN, 1 day apart), will be provided, displaying ALT, AST, Total bilirubin and ALP values.

A summary reporting the number of patients with at least one positive serum pregnancy test during the study will be provided.

2.4.5.5 Analyses of vital sign variables

Vital signs parameters are described in Section 2.1.4.5.

- A summary of baseline results will be provided for all parameters.
- For ECOG performance status, a shift table will be provided for the last and worst evaluations respectively relative to baseline.
- For blood pressure (SBP, DBP), a table will be provided with the last and worst evaluations (minimum and maximum value). In addition, for blood pressure, a graph describing mean changes from baseline and associated +/- SEM will also be done throughout the on-treatment period.
- Weight and BSA will be summarized at baseline only

For blood pressure parameters and weight, potentially clinically significant abnormality (PCSA) values are defined as abnormal values considered medically important by the Sponsor according to predefined criteria/thresholds based on literature review (Appendix E).

Moreover, the incidence of PCSAs at any time during the treatment-emergent adverse event period will be summarized irrespective of the baseline level and/or according to the following baseline status categories:

- Normal/missing
- Abnormal according to PCSA criterion or criteria

2.4.5.6 Analyses of electrocardiogram variables

The incidence of patients with at least 1 abnormal ECG at end of treatment will be summarized or according to the following baseline status categories:

- Normal/missing
- Abnormal

2.4.6 Analyses of biomarkers

Parameters described in Section 2.1.7 will be summarized on the all treated population using descriptive statistics.

2.4.6.1 Descriptive analyses

CA6 expression

At the end of part 1, baseline CA6 expression in tumor tissue will be summarized as the percent of positive cells at intensity 3+ and ≥2+ on the sum of membrane, as well as with the H-score on the sum of membrane. At the end of part 2, percent of positive cells at intensity 1+ will also be described. Descriptive analyses of cytoplasmic data may also be considered.

Baseline circulating CA6 will also be summarized:

- The relation between circulating CA6 at baseline and the percentage of positive cells of CA6 expression at intensity ≥ 2 in tumor tissue sample at baseline by IHC will be explored on the overall membrane by a scatter plot. Responders and non-responders will be identified by different symbols or colors.
- A monotonic correlation evaluated with Spearman's rank-order correlation coefficient could be done between the circulating CA6 at baseline and the percent of positive cells at intensity ≥2+ of CA6 expression by IHC.

[redacted]

[redacted]

Genetic alterations

For each genetic alteration listed in Section 2.1.7.3, a shift table will be performed between baseline and end-of-treatment (EOT) status to detect mutation changes.

2.4.6.2 Analyses between biomarkers expression and efficacy endpoints

CA6 expression

The analyses between CA6 IHC expression and efficacy endpoints will include the two expression cohorts (Part 2a and Part 2b).

To assess the potential link between CA6 IHC expression and efficacy endpoints, CA6 IHC expression will be described according to patients' status (responder/non responder).

Following graphical representations will be provided:

- Boxplots of percent of positive cells of CA6 at intensity ≥2+ according to patients' status (responder/non-responder)
- Potential link between percent of positive cells CA6 expression at intensity ≥2+ and the best relative tumor shrinkage will be assessed graphically with a Waterfall chart.
- Scatter plot of percent positive cells of CA6 in intensity ≥2+ by the best relative tumor shrinkage

To detect an association between biomarker status and ORR, a 1-sided Fisher's exact test will be calculated considering the number of responses in each cohort (overexpression cohort and mild expression cohort) and the sample size of the cohorts.

[redacted]

Genetic alterations

To assess the potential link between genetic alterations and efficacy data, the following analyses will be summarized for each genetic alteration:

- Number and percentage of patients with each genetic alteration status (baseline and EOT) by response status (responder/non-responder).
- For each genetic alteration, a waterfall plot of the best relative tumor shrinkage according to genetic alteration status will be displayed.

These analyses will be performed on the all-treated population.

2.4.7 Analyses of pharmacokinetic variables

PK parameters of SAR566658, naked huDS6 (SAR404461) and unconjugated maytansinoids (DM4 and Me-DM4) will be summarized with descriptive statistics (arithmetic and geometric means, standard deviation, coefficient of variation, minimum, median, and maximum).

The relationship between PK parameters and safety/efficacy endpoints may be investigated using logistic regression models with covariates selection using stepwise procedure.

Those estimates will then be investigated as prognostic factors for clinical outcome including safety and efficacy endpoints, if possible.

Previous PK data collected in other clinical studies such as TED10499, in which intensive blood sampling protocols were used, may be considered to help in model building.

2.4.8 Analyses of immunogenicity variables

Number (%) of patients with pre-existing ATAs, number (%) of ATAs negative patients, number (%) of inconclusive patients, number (%) of ATAs positive patients, number (%) of patients with treatment boosted ATAs, number (%) of patients with treatment induced ATAs (either transient or persistent) will be carried out.

Prevalence and incidence will be presented as well as descriptive statistics of titers and a listing of ATAs time point assessment.

The impact of positive immune response may be evaluated on efficacy, PK and safety endpoints.

2.4.9 Further therapy after discontinuation of investigational medicinal product administration during the study

Number (%) of patients with at least one post-treatment anti-cancer therapy (post-treatment systemic anti-cancer treatment, post-treatment radiotherapy or post-treatment surgery combined) will be summarized.

Moreover, a summary table will be provided for post-treatment anti-cancer therapy based on WHO-DD coding. Descriptive statistics will be also provided for post-treatment radiotherapy and post-treatment surgery.

2.5 DATA HANDLING CONVENTIONS

2.5.1 General conventions

The following formulas will be used for computation of parameters.

Demographic formulas

Body surface area value will be derived using the variation of DuBois and DuBois formula:

$$\text{BSA } (m^2) = \text{weight in kg}^{0.425} \times \text{height in cm}^{0.725} \times 0.007184$$

Renal function formulas

Creatinine clearance value will be from the equation of Cockroft and Gault:

$$\text{Creatinine clearance} = \frac{(140 - \text{age [years]} \times \text{body mass [kg]})}{\text{Plasma creatinine (mg/dL)} \times 72} \times \begin{array}{c}\text{Gender correction factor}\\ \text{(male: 1.00; female: 0.85)}\end{array}$$

Creatinine clearance in ml/min at a cycle will be calculated using last weight assessed before or at the date of serum creatinine measurement.

Corrected calcium formula

Corrected Calcium = Serum Calcium (in mmol/L) + 0.8 (4 - serum albumin [in G/dL])

2.5.2 Baseline values

In general, the baseline value of an analysis variable is defined as the last observation before the first dose of the treatment.

2.5.3 Data handling conventions for secondary efficacy variables

Valid/Evaluable tumor assessment:

A valid post-baseline tumor assessment is one for which the derived "overall objective response" is not "NE" (not evaluable) as per RECIST 1.1 criteria. If the date of assessment is missing, the tumor assessment is not valid.

In case of missing scans, i.e. if one lesion is not evaluated/missing for a given tumor assessment, the overall response should generally be NE unless there is a clear evidence of progression (independently of the missing lesion).

Date of last valid tumor assessment (different from PD): if several scans are performed at different dates for a given tumor evaluation reported in the e-CRF, the date of the last valid tumor assessment is the date of latest scan.

Date of disease progression determination:

- PD by New lesion: for progression based on a new lesion, the date of progression is the date of the first observation that the new lesion was detected,
- PD by target lesion: if multiple assessments based on the sum of target lesion measurements are done at different times, the date of progression is the date of the first radiological measurement,
- PD by non-target lesion: if the first documented progression is based on a non-target lesion alone, the date of PD is the initial date at which unequivocal progression is associated with non-target disease,

- PD by a combination of factors: if the first documented progression is based on a combination of target and new and/or non-target disease, the date of PD shall be the earliest date associated with any of the contributing factors.

The rules to determine date of progression are consistent with FDA recommendation described in "Guidance for Industry Clinical"

- Efficacy response variable: when a proportion is calculated for a binary response variable (eg, response rate), the denominator is based on the total number of patients in the analysis population used for the summary. There can be 3 observations: Yes, No and Nonevaluable (or missing). For the patients with non-evaluable outcomes, the default rule is that the patients will be treated as "no events".
- Time to event data: missing outcomes due to different reasons will be handled using adequate censoring rules. The censoring rules are specified as part of the definition of the analysis variables in corresponding section.
- Handling of missing/partial death dates:
  - If the day of the death date is missing, it will be imputed as the first day of the month, except if the date of the patient's last contact is in the same month as the death date. In this case, the death date is imputed as the date of last contact + 1 day.
  - If the day and month of the death date is missing, the date of death will be imputed to the first of January of the year, except if the date of the patient's last contact is in the same year as the death date. In this case, the death date will be imputed as the date of last contact + 1 day.
  - If the death date is missing, no imputation will be done and the patient will be censored at the last contact date.
- Incomplete date of first post-treatment anti-cancer therapy: if the day of first post-treatment anti-cancer therapy is missing, the date will be imputed to the first day of the month.

2.5.4 Missing data

In general, no imputation is planned for missing data. The following approaches are default methods for missing data handling.

For categorical variables, patients with missing data are not included in calculations of percentages unless otherwise specified. When relevant, the number of patients with missing data is presented.

Continuous data: the analyses and summaries for variables with continuous scales will be based on observed data only. However, the number of patients with missing observations will be provided.

Incomplete date of initial cancer diagnosis

If the day of initial cancer diagnosis is missing, the date will be imputed to the first day of the month.

Incomplete start/end date of prior anti-cancer therapy

If the day of start date of prior anti-cancer therapy is missing, the date will be imputed to the first day of the month.

If the day of end date of prior anti-cancer therapy is missing, the date will be imputed to the last day of the month.

Handling of medication missing/partial dates

No imputation of medications start/end dates will be performed (other than prior and further anti-cancer therapy). If a medication date is missing, so it cannot be determined whether it was taken prior or concomitantly, it will be considered as a prior, concomitant, and a post-treatment medication.

Handling of adverse events with missing or partial date/time of onset

Missing or partial adverse event onset dates and times will be imputed so that if the partial adverse event onset date/time information or visit number does not indicate that the adverse event started prior to treatment or after the treatment-emergent adverse event period, the adverse event will be classified as treatment-emergent. No imputation of adverse event end dates/times will be performed. These data imputations are for categorization purpose only and will not be used in listings. No imputation is planned for date/time of adverse event resolution

Handling of missing assessment of relationship of adverse events to investigational medicinal product

If the assessment of the relationship to IMP is missing, then the relationship to IMP has to be assumed and the adverse event considered as such in the frequency tables of possibly related adverse events, but no imputation should be done at the data level.

Handling of missing grades of adverse events

If the grade is missing for one of the treatment emergent occurrences of an AE, the maximal severity on the remaining occurrences will be considered. If the severity is missing for all the occurrences, no imputation will be done and the AE will be reported in all grades.

Handling of potentially clinically significant abnormalities

If a patient has a missing baseline he will be grouped in the category “normal/missing at baseline.”

For PCSAs with 2 conditions, one based on a change from baseline value or a normal range and the other on a threshold value, with the first condition being missing, the PCSA will be based only on the second condition.

2.5.5 Windows for time points

Summaries by cycles focusing on exposure and tumor assessments will be tabulated based on cycles as recorded in the e-CRF. For specific measurements such as vital signs and laboratory parameters, window defined around the date of treatment may be considered.

2.5.6 Unscheduled visits

Unscheduled visit measurements of laboratory data, vital signs, and ECG will be used for computation of baseline and worst values and/or grades.

2.5.7 Pooling of centers for statistical analyses

Data from all sites will be pooled together for analyses.

2.5.8 Statistical technical issues

Not applicable

3 INTERIM ANALYSIS

Interim analyses of efficacy, safety, and other data will be performed after 14 patients are treated at each dose level at the end of Part 1. These interim analyses are planned to occur when all patients have completed at least 4 postbaseline tumor assessments, experienced confirmed objective response, or discontinued the study for any reason. However, if the futility criteria of 2 responders is either met or not met with certainty for both arms when all patients have completed 2 postbaseline tumor assessments, experienced confirmed objective response, or discontinued the study for any reason, the analyses may be conducted at that time instead. Enrollment will be interrupted after completion of enrollment in Part 1, until completion of the Part 1 interim analysis.

An interim analysis based on efficacy will be performed to assess futility (to reject early the hypothesis that the response rate is 30%). If there is 0 or 1 response (CR or PR) in the first 14 patients treated at each dose, the 7.1% critical boundary is not met and the alternative hypothesis of at least 30% will be rejected. If the critical boundary of 7.1% is reached (2 responses or more), the study can continue in Part 2 at the selected dose.

At the same interim analysis, overwhelming efficacy will be assessed (to reject early that the response rate is less than or equal to 12%). If there are 6 responses or more in the first 14 patients treated at 1 dose, the critical boundary of 42.9% (1-sided nominal significance level of 2%) is reached and the null hypothesis of 12% will be rejected.

Overall, this procedure has 80% statistical power (1-sided alpha level of 10%) to reject the null hypothesis, using a gamma (-2) beta spending function for futility analysis and a Lan DeMets (OF) alpha spending function in a 1 sample test for a binomial proportion (East version 6.3 using exact computations, Cytel Software, Cambridge, MA)

4 DATABASE LOCK

The database is planned to be locked when all patients have completed at least 4 postbaseline tumor assessments, experienced confirmed objective response, or discontinued the study for any reason.

In case of early discontinuation of the study (for futility for instance), the database lock can be performed when all patients have completed at least 2 postbaseline tumor assessments, experienced confirmed objective response, or discontinued the study for any reason.

5 SOFTWARE DOCUMENTATION

All summaries and statistical analyses will be generated using SAS version 9.4 or higher.

6 REFERENCES

1. Eisenhauer EA, Therasse P, Boagerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-47.

7 LIST OF APPENDICES

Appendix A: Specification of the Bayesian design in Part 1

Appendix B: Determination of tumor response

Appendix C Modified Response Evaluation Criteria in Solid Tumors (RECIST 1.1)

Appendix D SOC internationally agreed order

Appendix E Potentially clinically significant abnormalities criteria for vital signs

Appendix F Description of primary and sensitivity analyses of PFS

Appendix G Generic ranges for hematological and biochemistry parameters

Appendix A Specification of the Bayesian design in Part 1

Primary objective of the study:

Objectives of the study are to select the dose based on overall response rate and safety of 2 dose levels 90 mg/m2 D1D8 Q3w and 120 mg/m2 D1D8 Q3w.

During the part 1 of the study, 14 patients will be included in both dose levels.

Safety will be monitored and the risks for predefined safety criteria (PSC) during the first 2 cycles will be followed for patients in part 1of the study. We will use a Bayesian binomial model as a monitoring tool.

Following predefined safety criteria have been selected for the study:

- Grade ≥3 TEAE from SOC eye disorders
- Grade ≥3 TEAE peripheral neuropathy (Preferred Term)
- Grade ≥4 TEAE

PSC evaluation will be performed on patients treated in the study who have completed 2 cycles, or who experienced predefined safety criteria (whatever the dose received).

The design will recommend a GO decision if it verifies overdosing risk constraints in the population:

- No more than 0.25 risk to recommend an overdosing dose (true PSC rate above 0.40).
- No more than 0.05 risk to recommend an unacceptable toxicity dose (true PSC rate above 0.60).

And so we will recommend to stop a dose level when overdosing dose or unacceptable toxicity are not controlled at the specific levels.

The Bayesian model is planned to be run after 4 evaluable patients at dose level 120 mg/m2 D1D8Q3W.

The calculation will be done independently for the two tested doses 90 mg/m2 D1D8Q3w and 120 mg/m2 D1D8Q3W.

Assumptions for prior elicitation

Prior distribution is elicited using the data from the first in human study (TED10499).

As dose level 90 mg/m^2 D1D8 Q3w has been already given in the phase I study, these data will be taken into account for the prior distribution of this dose level.

Dose level 120 mg/m2 D1D8 Q3w has never been tested before. As preliminary PK/PD analyses show that the dose intensity (mg/m2/week) is predictive of the occurrence of ocular toxicity (one of the predefined safety criteria), the data from dose level 240 mg/m2 Q3w with same dose intensity will be used for the elicitation of the prior distribution for the dose level 120 mg/m^2 D1D8 Q3w.

Table 4 provides the final results for both dose level selected from the phase I as prior distribution of the Bayesian model:

Table 4 - Results from the phase I TED10499 study

Dose Levels and schedules	**90 mg/m^2 D1D8 Q3w**	**240 mg/m^2 Q3w**
Number of patients	17	8
Number of PSC	1	2

Data from phase I are less informative than from phase II. Indeed patients selected for phase II are only CA6 positive with metastatic Triple Negative Breast Cancer and also they will all receive primary prophylaxis during their treatment whereas only some patients received primary prophylaxis during the phase I.

That is why a discount parameter k and k' is added for the corresponding dose level 90 mg/m^2 D1D8 Q3w and 120 mg/m^2 D1D8 Q3w such as:

$$0 < k' < k \leq 1$$

and so 1 patient from the phase II will correspond to k or k' patients from the phase I. k' will also be lower than k, since prior information for dose level 120 mg/m2 D1D8 Q3w are not coming from the exact same dose level.

Table 5 provides the information from phase I including discount parameter:

Table 5 - Results from the phase I TED10499 study including discount parameter

Dose Levels and schedules	**90 mg/m^2 D1D8 Q3w**	**240 mg/m^2 Q3w**
Number of patients	17k	8k'
Number of PSC	1k	2k'

There is no information before the phase I study, so the minimally informative beta distribution will be used. It will be the Jeffrey's prior beta distribution Beta (0.5, 0.5).

Resulting beta distributions after phase I:

- Beta(0.5+1k, 0.5+17k-1k) for dose level 90 mg/m^2 D1D8Q3w
- Beta(0.5+2k',0.5+8k'-2k') for dose level 120 mg/m^2 D1D8Q3w

Posterior distribution

The posterior distributions after n (or n') additional patients and x (or x') PSC respectively for dose level 90 mg/ m2 D1D8Q3w (or 120 mg/m2 D1D8Q3w) become

- Beta(0.5+1k+x, 0.5+17k-1k+n-x) for dose level 90 mg/m^2 D1D8Q3w
- Beta(0.5+2k'+x',0.5+8k'-2k'+n'-x') for dose level 120 mg/m^2 D1D8Q3w

Risk of overdosing and unacceptable toxic dose

If we note the previous distributions Beta(a,b), p-values for overdosing and unacceptable toxicity are derived for each dose following the formulas as below:

$$p_{overdosing} = 1 - Fbeta(threshold_{overdosing}, a, b)$$
$$p_{unaccaptable\ toxicity} = 1 - Fbeta(threshold_{unacceptable\ toxicity}, a, b)$$

with Fbeta the distribution function of Beta(a,b).

Choice for k and k'

Based on rules already defined above for k and k', some simulations have been run to find a value for k and k'. The target was to find a k and k' which will minimize the probability to take a wrong decision in any cases whatever the number of PSC found when we will run the Bayesian model at 4 or 8 patients (which are numbers of patients where we planned to use the Bayesian model). And in the same time, we would like k and k' which allow us to give some weight to the phase I and phase II in order for both results to be a part of the Go/No Go decisions.

For simulations, we will:

- Run the Bayesian model for 4 or 8 patients and see what will be the decision when 0, 1, 2, 3 or 4 PSC(s) occur(s).
- Using phase I result and for each case, a probability of PSC have been generated to simulate the rest of the patient until to get 14 patients for the phase II.
- 1000 simulations have been generated in order to calculate the probability of wrong decision where we compare the first decision at 4 or 8 patients with the simulated decision at 14 patients (do the decision will be the same at 4 or 8 patients compare to the decision taken with simulated data at 14 patients).

So simulations will be performed for each k and k' from 0.1 to 0.9 and following tables and figures show the results for probability of wrong decision for k and k' when we run the Bayesian model at 4 and 8 patients:

Figure 2 - Graph of probability of wrong decision by k values –90 mg/m^2 D1D8 Q3w at 4 patients

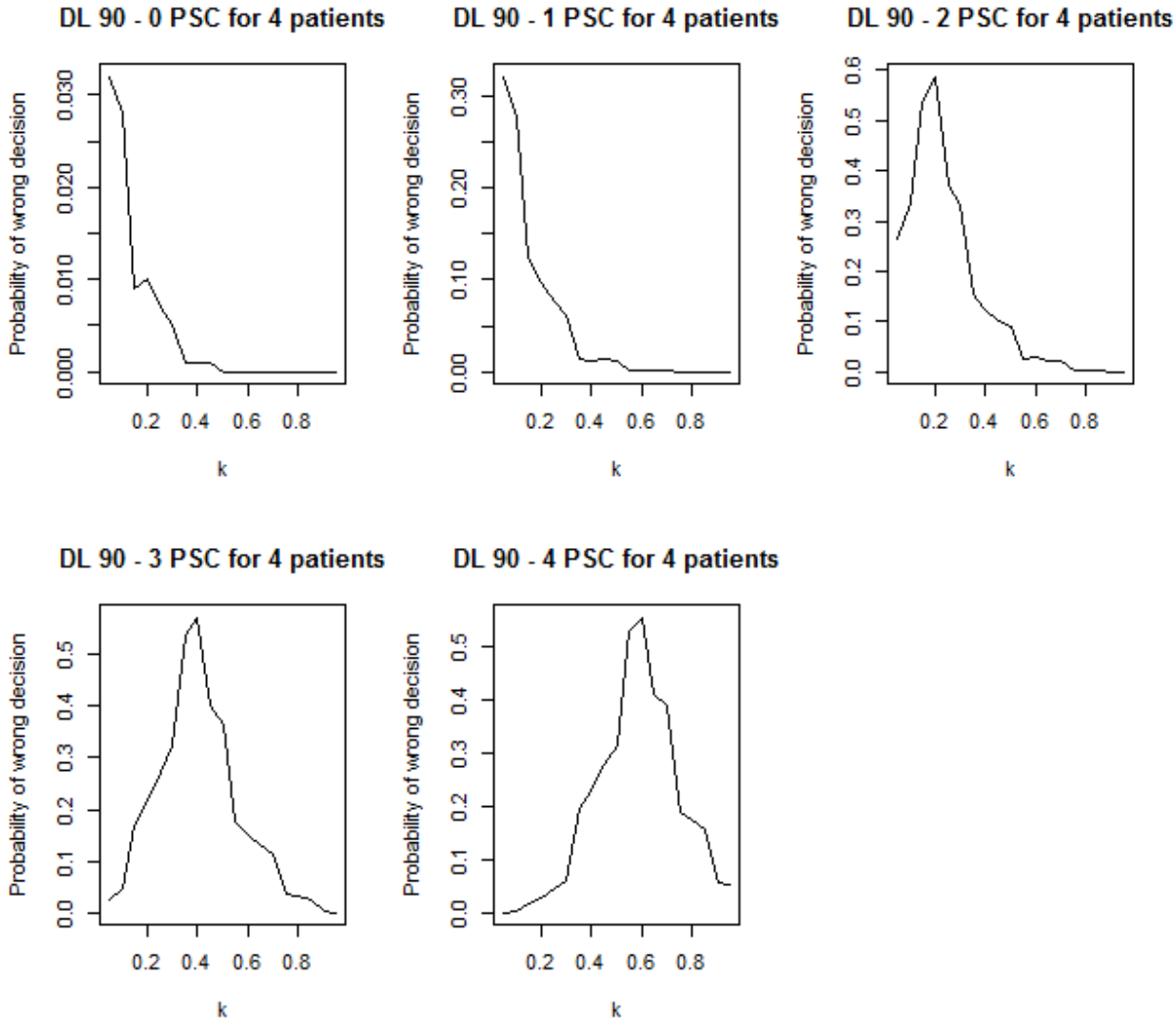


Figure 3 - Graph of probability of wrong decision by k values – 90 mg/m2 D1D8 Q3w at 8 patients

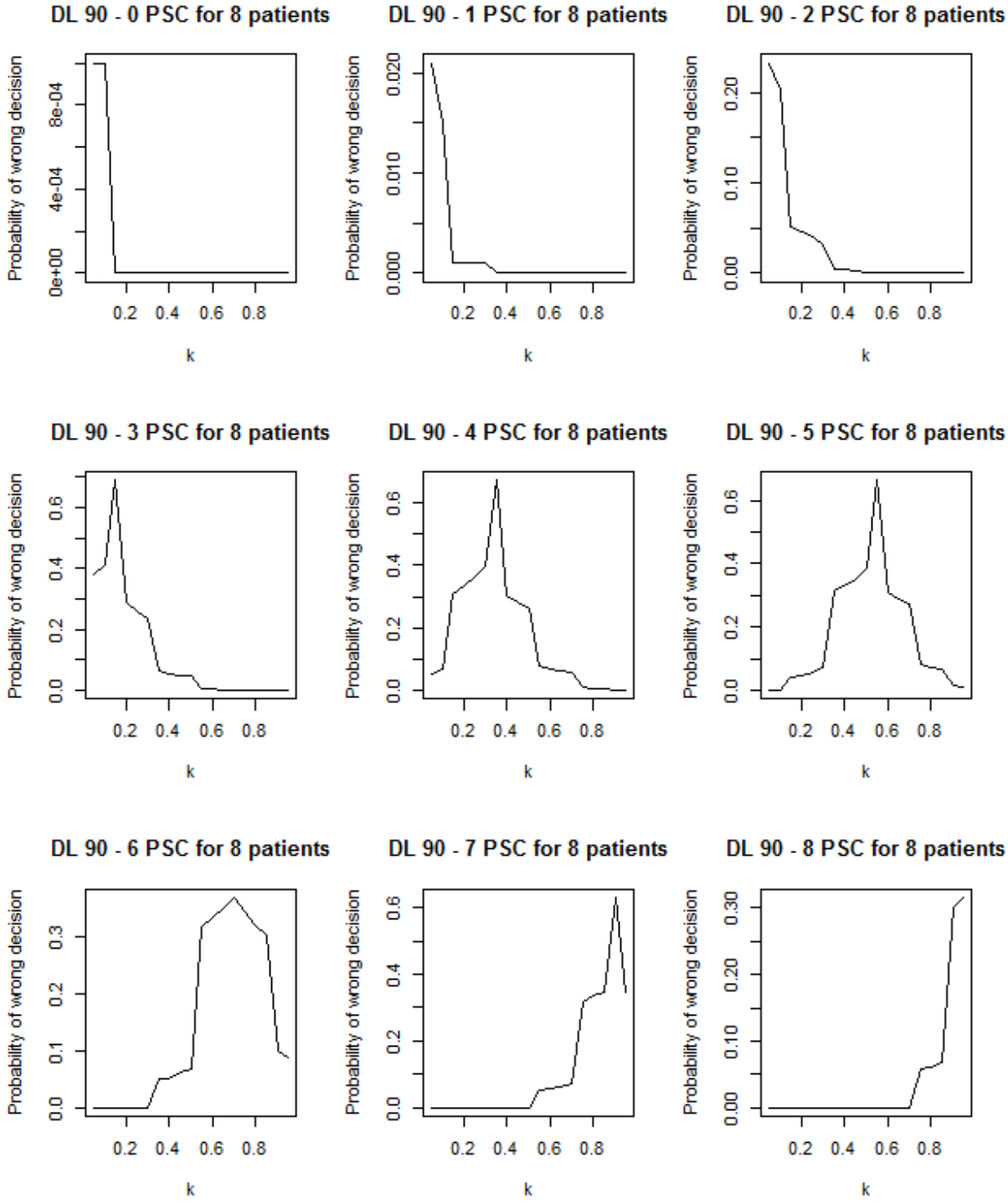


Table 6 - Probability of wrong decision by k values – 90 mg/m² D1D8 Q3w at 4 and 8 patients

Probability of wrong decisions by k values - Dose level 90 mg/m² D1D8 Q3w at 4 patients									
	k								
PSC	**0,1**	**0,2**	**0,3**	**0,4**	**0,5**	**0,6**	**0,7**	**0,8**	**0,9**
0	0,028	0,01	0,005	0,001	0	0	0	0	0
1	0,276	0,096	0,061	0,012	0,012	0,001	0,001	0	0
2	0,335	0,588	0,331	0,125	0,091	0,028	0,021	0,002	0
3	0,05	0,216	0,325	0,57	0,363	0,152	0,115	0,034	0,006
4	0,004	0,028	0,06	0,228	0,314	0,554	0,388	0,174	0,057
Sum1	0,693	0,938	0,782	0,936	0,78	0,735	0,525	0,21	0,063
Probability of wrong decisions by k values - Dose level 90 mg/m² D1D8 Q3w at 8 patients									
	k								
PSC	**0,1**	**0,2**	**0,3**	**0,4**	**0,5**	**0,6**	**0,7**	**0,8**	**0,9**
0	0,001	0	0	0	0	0	0	0	0
1	0,015	0,001	0,001	0	0	0	0	0	0
2	0,204	0,046	0,031	0,004	0,001	0	0	0	0
3	0,411	0,292	0,236	0,055	0,047	0,004	0,001	0	0
4	0,07	0,334	0,395	0,304	0,26	0,066	0,057	0,006	0
5	0	0,048	0,07	0,331	0,386	0,312	0,273	0,075	0,016
6	0	0	0	0,052	0,068	0,334	0,368	0,318	0,098
7	0	0	0	0	0	0,057	0,073	0,334	0,63
8	0	0	0	0	0	0	0	0,06	0,302
Sum2	0,701	0,721	0,733	0,746	0,762	0,773	0,772	0,793	1,046
Sum1+Sum2	1,394	1,659	1,515	1,682	1,542	1,508	1,297	1,003	1,109

As a higher k will give a weight too important to the phase I study and not enough weight to the phase II, we choose k=0.3 for dose level 90 mg/m² D1D8 Q3w since this is the best intermediate value which minimizes the probability of wrong decision whatever the number of PSC that appear at 4 or 8 patients.

Same work has been done for the dose level 120 mg/m² D1D8 Q3w.

Figure 4 - Graph of probability of wrong decision by k' values – 120 mg/m2 D1D8 Q3w at 4 patients

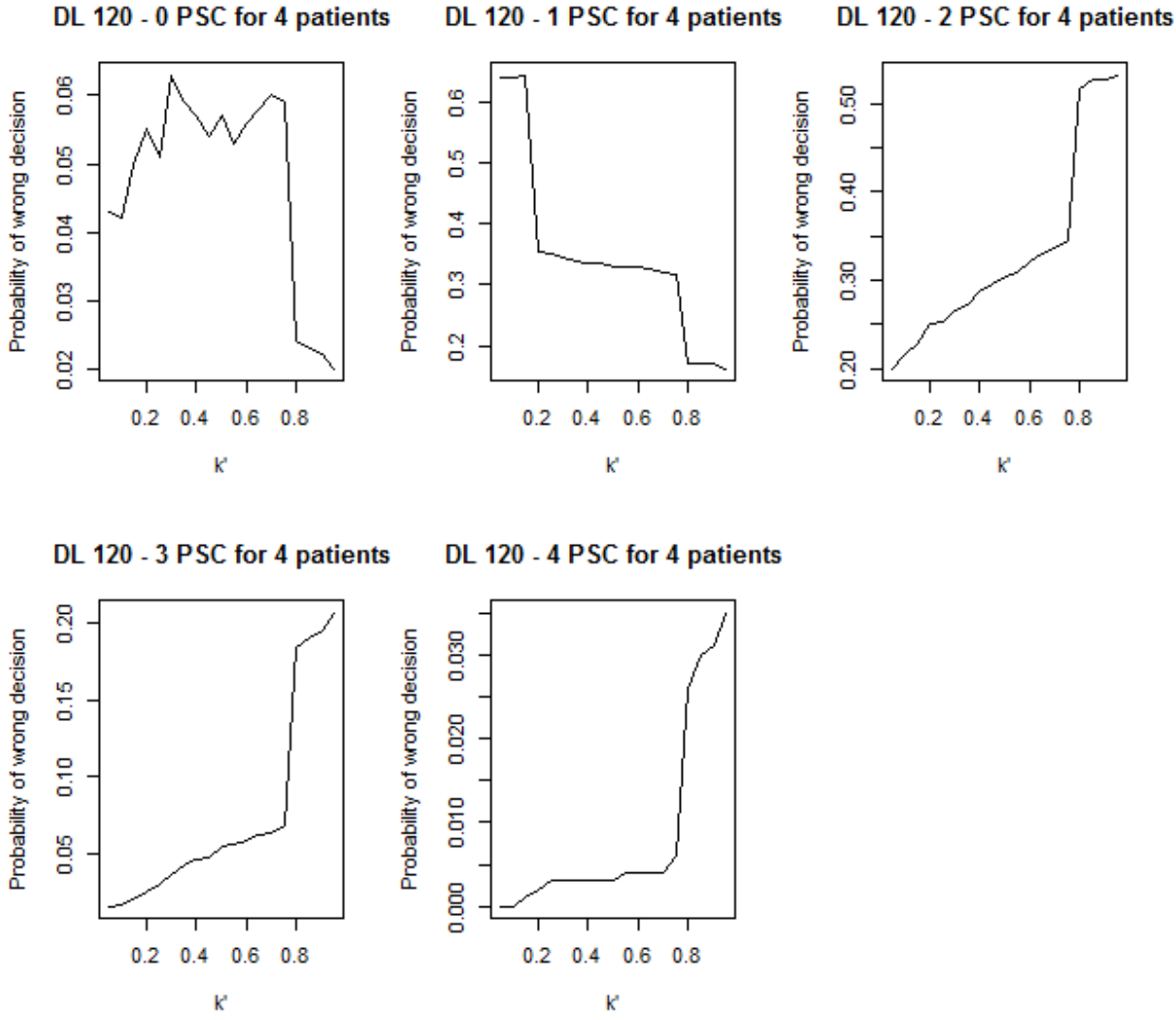


Figure 5 - Graph of probability of wrong decision by k' values – 120 mg/m2 D1D8 Q3w at 8 patients

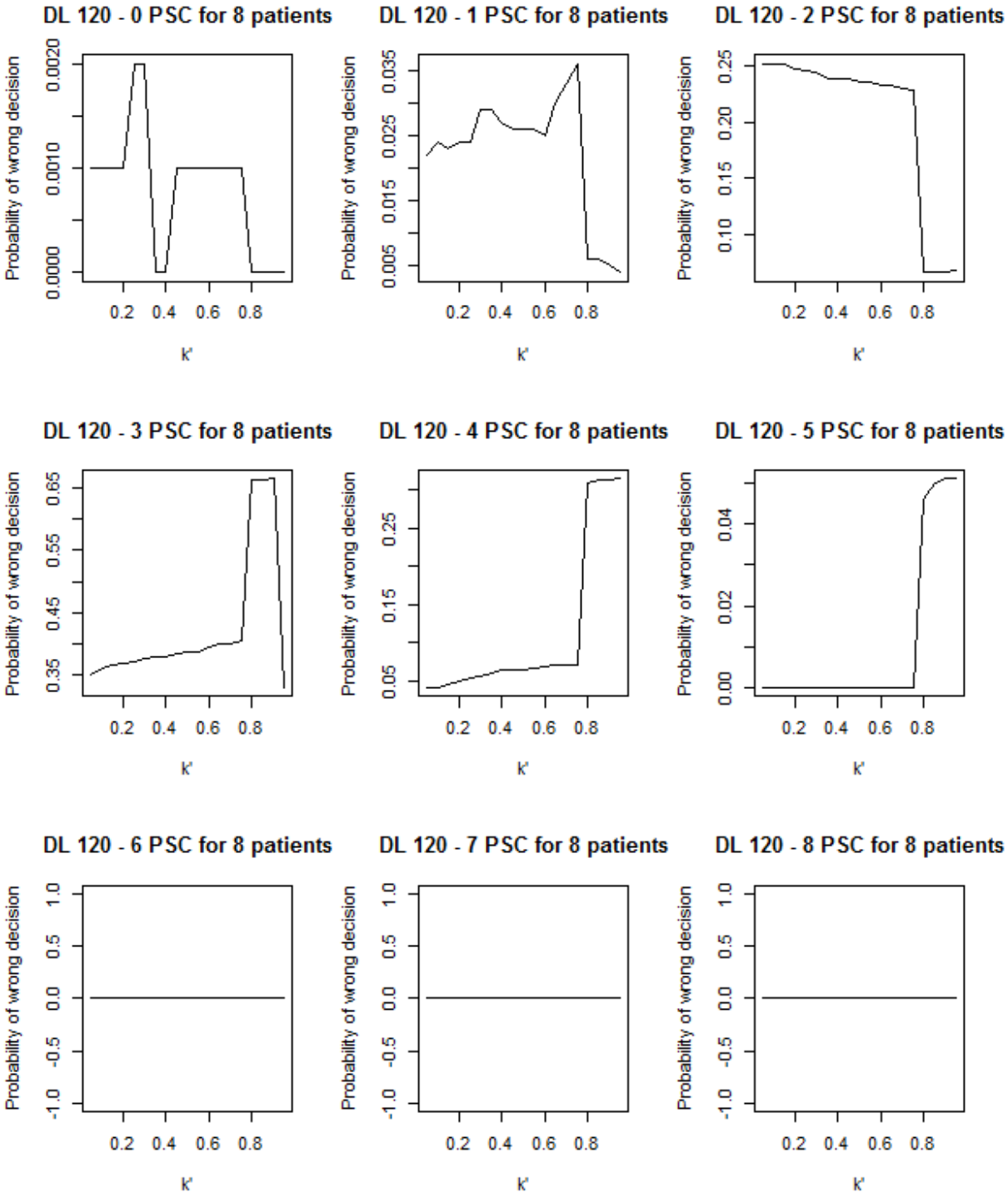


Table 7 - Probability of wrong decision by k' values –120 mg/m² D1D8 Q3w at 4 and 8 patients

Probability of wrong decisions by k values - Dose level 120 mg/m² D1D8 Q3w at 4 patients									
	k'								
PSC	**0,1**	**0,2**	**0,3**	**0,4**	**0,5**	**0,6**	**0,7**	**0,8**	**0,9**
0	0,042	0,055	0,063	0,057	0,057	0,056	0,06	0,024	0,022
1	0,643	0,356	0,345	0,336	0,329	0,33	0,321	0,171	0,169
2	0,216	0,25	0,265	0,287	0,304	0,321	0,336	0,516	0,527
3	0,017	0,026	0,037	0,046	0,055	0,059	0,064	0,185	0,196
4	0	0,002	0,003	0,003	0,003	0,004	0,004	0,026	0,031
Sum1	0,918	0,689	0,713	0,729	0,748	0,77	0,785	0,922	0,945
Probability of wrong decisions by k values - Dose level 120 mg/m² D1D8 Q3w at 8 patients									
	k'								
PSC	**0,1**	**0,2**	**0,3**	**0,4**	**0,5**	**0,6**	**0,7**	**0,8**	**0,9**
0	0,001	0,001	0,002	0	0,001	0,001	0,001	0	0
1	0,024	0,024	0,029	0,027	0,026	0,025	0,033	0,006	0,005
2	0,252	0,247	0,244	0,239	0,236	0,232	0,23	0,067	0,067
3	0,359	0,368	0,375	0,379	0,387	0,393	0,4	0,663	0,666
4	0,042	0,049	0,056	0,064	0,065	0,068	0,07	0,309	0,313
5	0	0	0	0	0	0	0	0,046	0,051
6	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0
Sum2	0,678	0,689	0,706	0,709	0,715	0,719	0,734	1,091	1,102
Sum1+Sum2	1,596	1,378	1,419	1,438	1,463	1,489	1,519	2,013	2,047

For dose level 120 mg/m² D1D8 Q3w, k'=0.2 give the lowest probability of wrong decision, so it will be choose for this dose level.

Results of the Bayesian model:

As we decide to use k=0.3 for dose level 90 mg/m² D1D8 Q3w and k'=0.2 for dose level 120 mg/m² D1D8 Q3w, we can run the Bayesian model for 3, 4, 6 and 8 patients and see what will be the decision when 0, 1, 2, 3 or 4 toxicity occurs.

As we did previously, we will do simulations in order to calculate two probabilities of wrong decision:

Compare to simulated decision at 14 patients: do the decision will be the same at 3, 4, 6 or 8 patients compare to the decision taken with simulated data at 14 patients.

Compare to the probability of toxicity generated using phase I results and data at 3, 4, 6 or 8 patients: do the decision taken after 3, 4, 6 and 8 patients is in agreement with the probability of toxicity.

Table 8 - Bayesian decision – 90 mg/m^2 D1D8 Q3w at 3 patients

Analysis at 3 patients at 90mg/m2 Q3W D1D8					
Number of PSC	**Decision**	**Overdosing probability**	**Unacceptable probability**	**Probability of wrong decision**	
				Compare to simulated decision at 14 patients (1)	**Compare to the probability of toxicity (2)**
0	GO	0,94%	0,03%	1,4%	0,7%
1	GO	7,61%	0,51%	11,3%	7,5%
2	No GO	25,64%	3,42%	56,7%	74,9%
3	No GO	52,69%	13,23%	20,6%	48,4%

Table 9 - Bayesian decision – 90 mg/m^2 D1D8 Q3w at 4 patients

Analysis at 4 patients at 90mg/m2 Q3W D1D8					
Number of PSC	**Decision**	**Overdosing probability**	**Unacceptable probability**	**Probability of wrong decision**	
				Compare to simulated decision at 14 patients (1)	**Compare to the probability of toxicity (2)**
0	GO	0,55%	0,01%	0,5%	0,2%
1	GO	4,94%	0,22%	6%	4,7%
2	GO	18,43%	1,67%	33%	17,2%
3	No GO	41,87%	7,34%	32,7%	58,8%
4	No GO	67,79%	21,45%	6,1%	32,5%

Table 10 - Bayesian decision – 90 mg/m^2 D1D8 Q3w at 6 patients

Analysis at 6 patients at 90mg/m2 Q3W D1D8					
Number of PSC	**Decision**	**Overdosing probability**	**Unacceptable probability**	**Probability of wrong decision**	
				Compare to simulated decision at 14 patients (1)	**Compare to the probability of toxicity (2)**
0	GO	0,19%	0%	0,1%	0,1%
1	GO	2,04%	0,04%	1%	2,1%
2	GO	9,09%	0,38%	13,3%	7,7%
3	GO	24,71%	2,06%	46,1%	24,4%
4	No GO	47,45%	7,56%	18,7%	52,5%

Table 11 - Bayesian decision – 90 mg/m^2 D1D8 Q3w at 8 patients

Analysis at 8 patients at 90mg/m2 Q3W D1D8					
Number of PSC	**Decision**	**Overdosing probability**	**Unacceptable probability**	**Probability of wrong decision**	
				Compare to simulated decision at 14 patients (1)	**Compare to the probability of toxicity (2)**
0	GO	0,07%	0%	0%	0,12%
1	GO	0,83%	0,01%	0,24%	0,74%
2	GO	4,28%	0,08%	3,82%	4,36%
3	GO	13,59%	0,53%	23,22%	13,04%
4	No GO	30,4%	2,33%	36,98%	69,74%

Table 12 - Bayesian decision – 120 mg/m^2 D1D8 Q3w at 3 patients

Analysis at 3 patients at 120mg/m2 Q3W D1D8					
Number of PSC	**Decision**	**Overdosing probability**	**Unacceptable probability**	**Probability of wrong decision**	
				Compare to simulated decision at 14 patients (1)	**Compare to the probability of toxicity (2)**
0	**GO**	7,72%	1,11%	9,6%	6,8%
1	**No GO**	35,18%	9,94%	50,9%	65,6%
2	**No GO**	70,84%	35,72%	12,3%	29,3%
3	**No GO**	92,97%	71,72%	1%	5,3%

Table 13 - Bayesian decision – 120 mg/m^2 D1D8 Q3w at 4 patients

Analysis at 4 patients at 120mg/m2 Q3W D1D8					
Number of PSC	**Decision**	**Overdosing probability**	**Unacceptable probability**	**Probability of wrong decision**	
				Compare to simulated decision at 14 patients (1)	**Compare to the probability of toxicity (2)**
0	**GO**	4,56%	0,44%	5,4%	4,2%
1	**GO**	24,2%	4,64%	35,3%	24,1%
2	**No GO**	56,58%	20,25%	25,3%	43,9%
3	**No GO**	84,12%	50,12%	2,7%	14,4%
4	**No GO**	96,83%	81,14%	0,2%	3,1%

Table 14 - Bayesian decision – 120 mg/m^2 D1D8 Q3w at 6 patients

Analysis at 6 patients at 120mg/m2 Q3W D1D8					
Number of PSC	**Decision**	**Overdosing probability**	**Unacceptable probability**	**Probability of wrong decision**	
				Compare to simulated decision at 14 patients (1)	**Compare to the probability of toxicity (2)**
0	**GO**	1,6%	0,07%	1,1%	1,2%
1	**GO**	10,89%	0,95%	13,7%	10,1%
2	**No GO**	32,71%	5,63%	50,6%	68,7%
3	**No GO**	61,31%	19,41%	13,9%	38,9%
4	**No GO**	84,29%	44,33%	1%	14,5%

Table 15 - Bayesian decision – 120 mg/m^2 D1D8 Q3w at 8 patients

Analysis at 8 patients at 120mg/m2 Q3W D1D8					
Number of PSC	**Decision**	**Overdosing probability**	**Unacceptable probability**	**Probability of wrong decision**	
				Compare to simulated decision at 14 patients (1)	**Compare to the probability of toxicity (2)**
0	**GO**	0,57%	0,01%	0,22%	0,48%
1	**GO**	4,69%	0,19%	3,48%	4,36%
2	**GO**	17,26%	1,38%	24,06%	16,78%
3	**No GO**	39,52%	6,15%	33,9%	60,7%
4	**No GO**	65,01%	18,44%	5,04%	34,92%

Appendix B Determination of tumor response

The overall response at each time point and the best overall response will be determined by the Investigator derived from the tumor measurements data according to RECIST 1.1 criteria (Appendix C) using the following algorithm.

- Overall Response:

Overall response, or integrated response, will be determined by the investigator using assessments of target, non-target and new lesions according to RECIST 1.1. See Appendix C for details.

- Confirmation:

According to RECIST 1.1 a confirmed CR or PR will be based on two tumor assessments at least four weeks apart as follows: if a CR or PR outcome is reported, then a second tumor assessment with CR or PR at least four weeks later is needed for confirmation, regardless of how many "unknown/not applicable/missing/non evaluable" tumor assessments are in between the two outcomes.

- Best Overall Response (BOR) :

In general, a patient's best tumor response will depend on the achievement of both: documented tumor measurement assessments and confirmation of the tumor response by RECIST1.1 criteria. For each patient the BOR will be derived based on the following algorithm:

− a) PR/CR will be determined based on confirmed PR and confirmed CR described in the algorithm above.

− b) SD and PD will be determined from the overall tumor response at each timepoint. If SD is the overall response and it occurs less than 35 days from first study administration, then overall response will not be SD. For patients who stop the treatment due to a symptomatic deterioration or death, without any other tumor assessment before the symptomatic deterioration or death, then the BOR will be considered as NE. These cases will be described separately as patients in early progression or early death.

If PD is the overall response and it occurs less than 42 days from the first study administration, then the best overall response is PD.

− c) Convention for PD after several NE cycles: in the case where the assessment at study entry is followed by several visits (or cycles) with no assessment (or only NE assessments or SD <35 days from first study administration) for more than twice the periodicity of tumor assessments (ie, for more than 42 days), and the next assessment is PD, the best overall response will be NE. If the PD occurs within this period of time, the best overall response will be PD.

Using the outcome from a), b) and c) above, the BOR will be determined based on the following order of Overall Responses: CR, PR, SD, PD and NE.

In the BOR determination, tumor assessments performed after a patient started post-treatment anti-tumor therapy will not be considered. The date of post-treatment anti-tumor therapy for a patient will be the earliest date among the dates for “Post Treatment Anti-Cancer Drug Therapy”, “Post Treatment Radiation Therapy” and “Post-Treatment Anti-Cancer Surgery” for target and non-target lesions identified at baseline. Assessments performed after the earliest post-treatment anti-tumor therapy will be excluded from the determination of BOR. If, in the same tumor evaluation period, several assessments are performed before the first post-treatment anticancer therapy and several assessments are performed afterwards, the disease assessment will be included in the determination of BOR.

Appendix C Modified Response Evaluation Criteria in Solid Tumors (RECIST 1.1)

Definitions

At baseline, tumor lesions/lymph nodes will be categorized measurable or non-measurable as follows:

Measurable:

- *Tumour lesions*: Must be accurately measured in at least one dimension (longest diameter in the plane of measurement is to be recorded) with a minimum size of:
  - 10 mm by CT scan (CT scan slice thickness no greater than 5 mm).
  - 10 mm caliper measurement by clinical exam (lesions which cannot be accurately measured with calipers should be recorded as non-measurable).
  - 20 mm by chest X-ray.
- *Malignant lymph nodes*: To be considered pathologically enlarged and measurable, a lymph node must be ≥15 mm in short axis when assessed by CT scan (CT scan slice thickness recommended to be no greater than 5 mm). At baseline and in follow-up, only the short axis will be measured and followed (see Special Issue 15). See also notes below on "Baseline documentation of target and non-target lesions" for information on lymph node measurement.

Non-measurable

- All other lesions, including small lesions (longest diameter <10 mm or pathological lymph nodes with 10 to <15 mm short axis) as well as truly non-measurable lesions. Lesions considered truly non-measurable include: leptomeningeal disease, ascites, pleural or pericardial effusion, inflammatory breast disease, lymphangitic involvement of skin or lung, abdominal masses/abdominal organomegaly identified by physical exam that is not measurable by reproducible imaging techniques.

Special considerations regarding lesion measurability

Bone lesions, cystic lesions, and lesions previously treated with local therapy require particular comment:

Bone lesions:

- Bone scan, PET scan or plain films are not considered adequate imaging techniques to measure bone lesions. However, these techniques can be used to confirm the presence or disappearance of bone lesions.
- Lytic bone lesions or mixed lytic-blastic lesions, with identifiable soft tissue components, that can be evaluated by cross sectional imaging techniques such as CT or MRI can be

considered as measurable lesions if the soft tissue component meets the definition of measurability described above.

- Blastic bone lesions are non-measurable.

Cystic lesions:

- Lesions that meet the criteria for radiographically defined simple cysts should not be considered as malignant lesions (neither measurable nor non-measurable) since they are, by definition, simple cysts.
- "Cystic lesions" thought to represent cystic metastases can be considered as measurable lesions, if they meet the definition of measurability described above. However, if non-cystic lesions are present in the same patient, these are preferred for selection as target lesions.

Lesions with prior local treatment:

- Tumour lesions situated in a previously irradiated area, or in an area subjected to other loco-regional therapy, are usually not considered measurable unless there has been demonstrated progression in the lesion. Study protocols should detail the conditions under which such lesions would be considered measurable.

Methods of measurement

Measurement of lesions: All measurements should be recorded in metric notation, using calipers if clinically assessed. All baseline evaluations should be performed as close as possible to the treatment start and never more than 4 weeks before the beginning of the treatment.

Method of assessment: The same method of assessment and the same technique should be used to characterize each identified and reported lesion at baseline and during follow-up. Imaging based evaluation should always be done rather than clinical examination unless the lesion(s) being followed cannot be imaged but are assessable by clinical exam. Clinical lesions: Clinical lesions will only be considered measurable when they are superficial and 10 mm diameter as assessed using calipers (e.g. skin nodules). For the case of skin lesions, documentation by color photography including a ruler to estimate the size of the lesion is suggested. As noted above, when lesions can be evaluated by both clinical exam and imaging, imaging evaluation should be undertaken since it is more objective and may also be reviewed at the end of the study.

- *Chest X-ray*: Chest CT is preferred over chest X-ray, particularly when progression is an important endpoint, since CT is more sensitive than X-ray, particularly in identifying new lesions. However, lesions on chest X-ray may be considered measurable if they are clearly defined and surrounded by aerated lung.
- *CT, MRI*: CT is the best currently available and reproducible method to measure lesions selected for response assessment. This guideline has defined measurability of lesions on CT scan based on the assumption that CT slice thickness is 5 mm or less. When CT scans have slice thickness greater than 5 mm, the minimum size for a measurable lesion should be twice the slice thickness. MRI is also acceptable in certain situations (eg, for body scans).

- *Ultrasound*: Ultrasound is not useful in assessment of lesion size and should not be used as a method of measurement.
- *Endoscopy, laparoscopy*: The utilisation of these techniques for objective tumour evaluation is not advised. However, they can be useful to confirm complete pathological response when biopsies are obtained or to determine relapse in trials where recurrence following complete response or surgical resection is an endpoint.
- *Tumour markers*: Tumour markers alone cannot be used to assess objective tumour response. If markers are initially above the upper normal limit, however, they must normalise for a patient to be considered in complete response. Specific guidelines for both CA-125 response (in recurrent ovarian cancer) and PSA response (in recurrent prostate cancer), have been published. In addition, the Gynecologic Cancer Intergroup has developed CA125 progression criteria which are to be integrated with objective tumour assessment for use in first-line trials in ovarian cancer.
- *Cytology, histology*: These techniques can be used to differentiate between PR and CR in rare cases if required by protocol (for example, residual lesions in tumour types such as germ cell tumours, where known residual benign tumours can remain). When effusions are known to be a potential adverse effect of treatment (eg, with certain taxane compounds or angiogenesis inhibitors), the cytological confirmation of the neoplastic origin of any effusion that appears or worsens during treatment can be considered if the measurable tumour has met criteria for response or stable disease in order to differentiate between response (or stable disease) and progressive disease.

Tumour response evaluation

To assess objective response or future progression, it is necessary to estimate the overall tumour burden at baseline and use this as a comparator for subsequent measurements. Only patients with measurable disease at baseline should be included in protocols where objective tumour response is the primary endpoint. Measurable disease is defined by the presence of at least one measurable lesion. In studies where the primary endpoint is tumour progression (either time to progression or proportion with progression at a fixed date), the protocol must specify if entry is restricted to those with measurable disease or whether patients having non-measurable disease only are also eligible.

Response Criteria

Table 16 - Evaluation of target lesions

Evaluation of target lesions	
Complete Response (CR)	Disappearance of all target lesions. Any pathological lymph nodes (whether target or non-target) must have reduction in short axis to <10 mm.s
Partial Response (PR):	At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters.

Evaluation of target lesions	
Progressive Disease (PD):	At least a 20% increase in the sum of diameters of target lesions, taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm. (Note: the appearance of one or more new lesions is also considered progression).
Stable Disease (SD):	Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameters while on study

Table 17 - Evaluation of non-target lesions

Evaluation of non-target lesions	
Complete Response (CR):	Disappearance of all non-target lesions and normalisation of tumour marker level. All lymph nodes must be non-pathological in size (<10 mm short axis).
Incomplete Response/Stable Disease (SD):	Persistence of one or more non-target lesion(s) and/or maintenance of tumour marker level above the normal limits.
Progressive Disease (PD):	Unequivocal progression (see comments below) of existing non-target lesions. (Note: the appearance of one or more new lesions is also considered progression).

Although a clear progression of "non target" lesions only is exceptional, in such circumstances, the opinion of the treating physician. Should prevail and the progression status should be confirmed later on by the review panel (or study chair).

Evaluation of Best overall response

It is assumed that at each protocol specified time point, a response assessment occurs. The following table provides a summary of the overall response status calculation at each time point for patients who have measurable disease at baseline.

Table 18 - Evaluation of overall response

Target lesions	Non-target lesions	New lesions	Overall response
CR	CR	No	CR
CR	Non-CR/non-PD	No	PR
CR	Not evaluated	No	PR
PR	Non-PD or not all evaluated	No	PR
SD	Non-PD or not all evaluated	No	SD
Not all evaluated	Non PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

Table 19 - Evaluation of best overall response

Overall response	Overall response	BEST overall response
First time point	Subsequent time point	
CR	CR	CR
CR	PR	SD, PD or PR[a]
CR	SD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	NE	SD provided minimum criteria for SD duration met, otherwise NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
PR	NE	SD provided minimum criteria for SD duration met, otherwise NE
NE	NE	NE

[a] If a CR is truly met at first time point, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). Best response would depend on whether minimum duration for SD was met. However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the patient had PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.

The best overall response is determined once all the data for the patient is known.

Best response determination in trials where confirmation of complete or partial response IS NOT required: Best response in these trials is defined as the best response across all time points (for example, a patient who has SD at first assessment, PR at second assessment, and PD on last assessment has a best overall response of PR). When SD is believed to be best response, it must also meet the protocol specified minimum time from baseline. If the minimum time is not met when SD is otherwise the best time point response, the patient's best response depends on the subsequent assessments. For example, a patient who has SD at first assessment, PD at second and does not meet minimum duration for SD, will have a best response of PD. The same patient lost to follow-up after the first SD assessment would be considered invaluable.

Best response determination in trials where confirmation of complete or partial response IS required: Complete or partial responses may be claimed only if the criteria for each are met at a subsequent time point as specified in the protocol (generally 4 weeks later). In this circumstance, the best overall response can be interpreted as in Table 19.

Appendix D SOC internationally agreed order

The internationally agreed order (Guideline on summary of product characteristics, December 1999, European commission) for SOC:

1. Infections and infestations
2. Neoplasms benign and malignant (including cysts and polyps)
3. Blood and the lymphatic system disorders
4. Immune system disorders
5. Endocrine disorders
6. Metabolism and nutrition disorders
7. Psychiatric disorders
8. Nervous system disorders
9. Eye disorders
10. Ear and labyrinth disorders
11. Cardiac disorders
12. Vascular disorders
13. Respiratory, thoracic and mediastinal disorders
14. Gastrointestinal disorders
15. Hepato-biliary disorders
16. Skin and subcutaneous tissue disorders
17. Musculoskeletal, connective tissue and bone disorders
18. Renal and urinary disorders
19. Pregnancy, puerperium and perinatal conditions
20. Reproductive system and breast disorders
21. Congenital and familial/genetic disorders
22. General disorders and administration site conditions
23. Investigations
24. Injury and poisoning
25. Surgical and medical procedures
26. Social circumstances

The other terms are sorted by dictionary code order

Appendix E Potentially clinically significant abnormalities criteria for vital signs

Parameter	PCSA	Comments
Vital signs		
SBP	≤95 mmHg and decrease from baseline ≥20mmHg ≥160 mmHg and increase from baseline ≥20 mmHg	To be applied for all positions (including missing) except STANDING.
DBP	≤45 mmHg and decrease from baseline ≥10 mmHg ≥110 mmHg and increase from baseline ≥10 mmHg	To be applied for all positions (including missing) except STANDING.
Weight	≥5% increase from baseline ≥5% decrease from baseline	FDA Feb 2007.

Appendix F Description of primary and sensitivity analyses of PFS

Table 20 - PFS Primary analysis

Situation	Date of progression or censoring	Outcome
No baseline tumor assessments	Date of first treatment administration	Censored
Documented progression according to investigator	Earlier of: - Date of radiological assessment showing new lesion or progression of non-target lesion - Date of last radiological assessment of target lesions (if progression is based on increase in sum of target lesions)	Event
Death without documented progression according to investigator	Date of death	Event
Both documented progression according to investigator and death	Date of progression	Event
No documented progression according to investigator and no death	Date of last valid tumor assessment without evidence of progression	Censored
Symptomatic deterioration reported and no documented progression according to investigator and no death	Ignored	Ignored
Initiation of new anti-cancer treatment	Ignored	Ignored

Table 21 - PFS sensitivity analysis #1 (progression based on investigator's assessment of lesions, and including symptomatic deterioration)

Situation	Date of progression or censoring	Outcome
No baseline tumor assessments	Date of first treatment administration	Censored
Documented Progression (without symptomatic deterioration) according to the investigator	Earlier of: - Date of radiological assessment showing new lesion or progression of non-target lesion - Date of last radiological assessment of target lesions (if progression is based on increase in sum of target lesions)	Event
Death without documented progression and without symptomatic deterioration	Date of death	Event
Symptomatic deterioration reported with no documented progression according to investigator and no death	Date of symptomatic deterioration	Event
Symptomatic deterioration and documented progression according to the investigator regardless of death occurrence	Earlier of the date of symptomatic deterioration and date of progression	Event
Symptomatic deterioration and death with no documented progression according to the investigator	Date of symptomatic deterioration	Event
Documented progression according to the investigator and death with no symptomatic deterioration	Date of progression	Event
No progression according to the investigator, no death, and no symptomatic deterioration	Earlier of the date of last valid tumor assessment without evidence of progression and cut-off date	Censored
New anti-cancer treatment started	Ignored	Ignored

Appendix G Generic ranges for hematological and biochemistry parameters

The current list of generic ranges for hematological parameters (for adults) is provided in the table below:

Table 22 – Generic ranges for hematological parameters

LBTESTCD	LBTEST	GENDER	LBSTRESU	LBGNNRLO
HGB	Hemoglobin	F	g/L	120
HGB	Hemoglobin	M	g/L	135
LYM	Lymphocytes		10^9/L	1
NEUT	Neutrophils		10^9/L	1,8
PLAT	Platelets		10^9/L	150
WBC	Leukocytes		10^9/L	4,5
EOS	Eosinophils		10^9/L	0
BASO	Basophils		10^9/L	0
MONO	Monocytes		10^9/L	0,18
HCT	Hematocrit	M	%	0,41
HCT	Hematocrit	F	%	0,36
RBC	Erythrocytes	F	10^12/L	4
RBC	Erythrocytes	M	10^12/L	4,5

Based on NEJM (N Engl J Med 2004;351:1548-63.): “Laboratory Reference Values”, Alexander Kratz, M.D., Ph.D., M.P.H., Maryjane Ferraro, Ph.D., M.P.H., Patrick M. Sluss, Ph.D., and Kent B. Lewandrowski, M.D.

The current list of generic ranges for biochemistry parameters (for adults) is provided in the table below:

Table 23 – Generic ranges for biochemistry parameters

LBTEST	LBSTRESU	LBGNNRLO - LBGNNRHI
Albumin	g/L	35 - 55
Blood Urea Nitrogen (BUN)	mmol/L	NA-17
Calcium	mmol/L	2,2 - 2,6
Chloride	mmol/L	80 - 115
Glucose	mmol/L	3,9 - 7
Potassium	mmol/L	3,5 - 5
Magnesium	mmol/L	0,8 - 1,2
Sodium	mmol/L	129 - 160
Phosphate	mmol/L	1 - 1,4
Protein	g/L	55 - 80
Urea	mmol/L	3,6 - 7,1

ACT14884 16.1.9 Statistical analysis plan

ELECTRONIC SIGNATURES

Signed by	Meaning of Signature	Server Date (dd-MMM-yyyy HH:mm)
[redacted]	Clinical Approval	12-Jun-2018 14:43 GMT+0200
[redacted]	Clinical Approval	12-Jun-2018 19:28 GMT+0200