etop
Information | Research

PROMISE-meso

ETOP 9-15

A multicentre randomized phase III trial comparing pembrolizumab versus standard chemotherapy for advanced pre-treated malignant pleural mesothelioma

Statistical Analysis Plan (SAP)
Final efficacy analysis

A clinical trial of ETOP

SAP VERSION	1.1
SAP DATE	15.01.2019

Table of Contents

INTRODUCTION

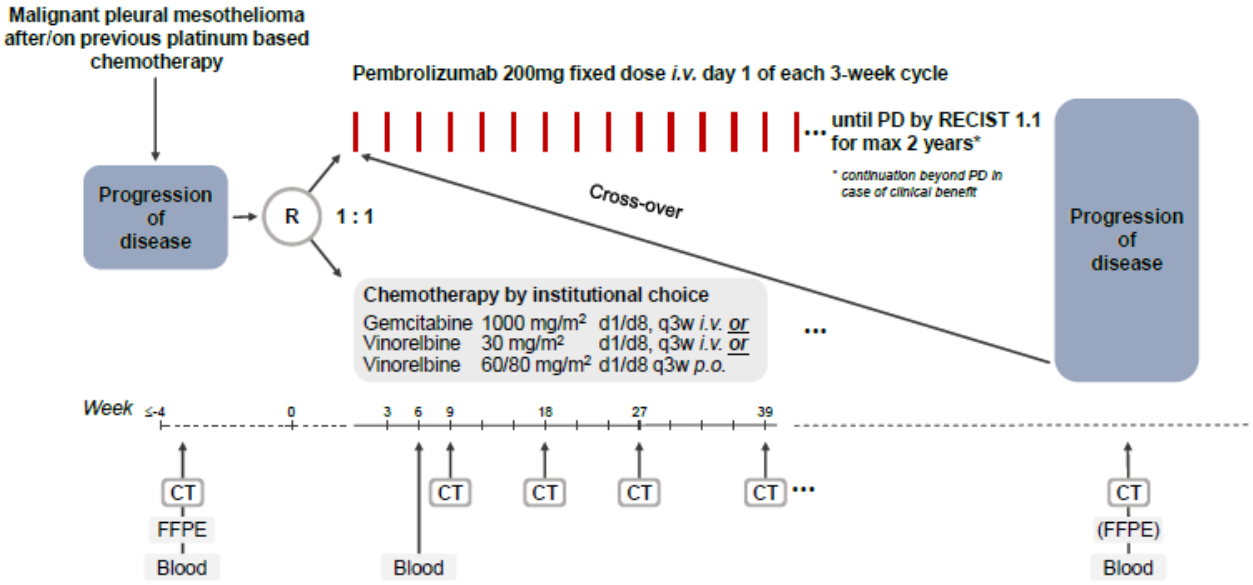
This is a technical document describing thoroughly and in detail an analysis framework for the ETOP 9-15 PROMISE-meso trial.

A short description of the contents of this statistical analysis planned is provided below:

1. **PROMISE-meso trial outline/oversight**: trial's schema, eligibility criteria, objectives, trial duration, sample size & power, treatment
2. **General considerations**: Analysis timing, data retrieval
3. **Statistical considerations**: definition of primary and secondary endpoints, (serious) adverse events definition, analysis populations
4. **Study Analysis:** Baseline characteristics, treatment administration, efficacy analysis, safety analysis and exploratory analysis, handling of missing data
5. **Technical issues:** Software and testing

1 Trial oversight

This is a randomized phase III multicentre clinical trial aiming to demonstrate superiority of pembrolizumab versus standard, institutional-choice chemotherapy (gemcitabine or vinorelbine) in patients with advanced pre-treated malignant mesothelioma. Patients randomized to chemotherapy arm will be allowed to cross-over and receive pembrolizumab at progression.



SCHEMA 1. Trial design

1.1 Objectives

Primary objective

The **primary objective** of the study is to investigate whether treatment with pembrolizumab improves progression-free survival (PFS), as assessed by independent radiological review, compared to standard, institutional choice chemotherapy (gemcitabine/vinorelbine).

Secondary objectives

The **secondary objectives** of the study include:

- To evaluate secondary measures of clinical efficacy including objective response (OR), investigator assessed PFS (ia-PFS), overall survival (OS), and time to treatment failure (TTF)
- To asses the safety and tolerability of the treatment

1.2 Endpoints

Primary endpoint:

• PFS according to RECIST 1.1 criteria based on independent radiological review

Secondary endpoints:

- Objective response determined by RECIST 1.1
- Overall survival
- Time to treatment failure
- Investigator assessed PFS determined according to RECIST 1.1
- Tolerability assessed by adverse events graded according to CTCAE v4.0

Correlative endpoints:

- Responses according to PD-L1 expression levels, measured by IHC
- TILs analysis
- Mutation load

1.3 Most important eligibility criteria

Inclusion criteria at enrolment:

- Histologically confirmed malignant pleural mesothelioma (all subtypes are eligible)
- Progressing after or on previous platinum-based chemotherapy.
- Availability of tumour tissue for translational research
- Male and female patients
- Age 18 years ECOG performance status 0-1
- Life expectancy of at least 3 months
- Measurable or evaluable disease according to RECIST 1.1 criteria
- Adequate hematological, renal, and liver function

Exclusion criteria at enrolment:

- Prior therapy with an anti-PD-1, anti-PD-L1/L2, anti-CD137, or anti-CTLA-4 antibody
- Prior therapy with gemcitabine or vinorelbine
- Known active central nervous system metastases and/or carcino matous meningitis.

- Active autoimmune disease that has required systemic treatment in past 2 years (i.e. with use of disease modifying agents, corticosteroids or immunosuppressive drugs).

1.4 Treatment

Experimental arm

Pembrolizumab, 200 mg fixed dose *i.v*. on day 1 of every 3-week (+/-3 days) cycle until progression of disease determined according to RECIST 1.1 criteria or lack of tolerability or until further protocol treatment is declined by the patient, for a maximum of 2 years.

In case of clinical benefit, with physician and patient agreement, pembrolizumab treatment can continue beyond documented disease progression according to RECIST 1.1 criteria until a maximum of 2 years on pembrolizumab treatment is reached. Patients need to meet the following criteria: ECOG performance status 0-1, absence of rapid progression of disease,

absence of progressive tumour at critical anatomical

Control arm

For the control chemotherapy options, the choice of vinorelbine (*p.o.*) or vinorelbine (*i.v.*) or gemcitabine chemotherapy will be made on a per-patient basis prior to randomisation.

- Gemcitabine 1000 mg/m2 *i.v*., day 1 and day 8 of every 3-week (+/-3 days) cycle
- Vinorelbine *i.v.* 30 mg/m2 *i.v.*, day 1 and day 8 of every 3-week (+/-3 days) cycle
- Vinorelbine 60/80 mg/m2 *p.o.*, day 1 and day 8 of every 3-week (+/-3 days) cycle

At documented disease progression according to RECIST 1.1 criteria, patients in the control arm are allowed to receive pembrolizumab, if they meet the cross-over criteria:

ECOG performance status 0-1, absence of progressive tumours at critical anatomical sites (e.g., cord compression) requiring urgent alternative medical intervention.

1.5 Sample size & power

This is a 1:1 randomized phase III trial (stratified by predominately epitheloid vs. non-epitheloid histological subtype), designed to detect with **80% power**, **using a one-sided test at significance level of 0.025**, an **increase** of the **median PFS from 3.5 months on the chemotherapy arm to 6 months on pembrolizumab**. This corresponds to a **6-month PFS of 30% vs 50% for the chemotherapy and the pembrolizumab group respectively**

[hazard ratio (HR)=0.58]. A cumulative drop-out of 5% is assumed. A total of **110** events need to be observed to achieve the trial goal.

The expected accrual rate is 2 patients per month for the first 6 months, increasing to 10 patients per month thereafter. A total sample size of ***142*** patients accrued over a period of 19 months would be required to observe the total number of 110 events with a maximum total trial follow-up of 24 months.

The primary analysis is expected to be available 36 months after the inclusion of the first patient. The trial treatment phase will continue for a maximum of 2 years from the inclusion of the last patient and the trial is expected to end at all sites approximately 43 months after the inclusion of the first patient.

1.6 Total trial duration

The trial treatment phase will continue for a maximum of 2 years from the inclusion of the last patient and the trial is expected to end at all sites approximately 43 months after the inclusion of the first patient.

2 General considerations

2.1 Analysis timing

According to the statistical design, the primary analysis is scheduled to be performed when 110 PFS events have been observed, for a total of 142 patients accrued over a period of 19 months. Under the protocol assumptions, this is expected to occur 24 months after the inclusion of the first patient.

2.2 Data Retrieval Information

The final analysis will be based on the database download that will take place, as soon as the total number of 110 PFS events (based on independent radiological review) required according to the statistical design of the trial are observed. Using this database extraction, a set of queries will be produced and forwarded to trial's data manager with the expectation to be answered in a pre-specified time period (approximately four weeks). Corrections and responses based on these queries, will be used for correcting the previously downloaded database, in order to create the final clean dataset to be used for the analysis.

3 Statistical considerations

3.1 Study's endpoints

Primary endpoint

The primary endpoint of the trial is PFS, defined as time from randomization until documented progression (based on independent radiological review. according to RECIST 1.1 criteria) or death, whichever occurs first.

Secondary endpoints

Secondary endpoints include OR (tumor response based on independent radiological review), OS, TTF, investigator assessed PFS (ia-PFS) and adverse events graded according to CTCAE V4.0.

More specifically OR is defined as the best overall response (complete or partial) across all assessment time-points according to RECIST Criteria 1.1, during the period from randomization to termination of trial treatment. OS is defined as the time from the date of randomization until death from any cause. In addition, TTF is defined as time from the date of randomization to treatment failure for any reason, including progression of disease, treatment toxicity, refusal/withdrawal and death (even after treatment completion). Ia-PFS, is defined as the time from randomization to documented progression (investigator assessed) or death from any cause.

More details on the exact definition and calculation of efficacy endpoints is provided in Table 1 that follows.

Finally, a detailed description of AEs and SAEs is provided in the section 3.2.

Correlative endpoints

Correlative endpoints of the study include responses according to PD-L1 expression levels, measured by IHC (cut-offs considered will be 1% and 50%) as well as TILs analysis and mutation load, *depending on data availability*.

Exploratory endpoints

Time to treatment discontinuation (TTD), defined as the time from randomization to treatment discontinuation for any reason and Duration of response (DoR), defined as the time from documentation of tumor response (either partial or complete based on independent radiological review) to disease progression or death, will be explored in

patients with respect to their initial treatment assignment and additionally for patients who crossed-over to receive pembrolizumab, after progression under standard institutional choice chemotherapy.

Furthermore, PFS2 defined as time from randomisation to objective tumor progression (based either on independent radiological review or investigator assessed) on further line of treatment or death from any cause, whichever occurs first, will be also presented.

3.2 (Serious) Adverse Events

Adverse events

The main criterion for treatment tolerability is the occurrence of toxicities and adverse events. The severity and causality will be classified according to the NCI CTCAE Version 4. The CTCAE is available for downloading (http://evs.nci.nih.gov/ftp1/CTCAE/About.html).

An adverse event is defined as any untoward medical occurrence that occurs from the first dose of study medication until 30 days after the final dose, regardless of whether it is considered related to a medication. In addition, any known untoward event that occurs subsequent to the adverse event reporting period that the investigator assesses as possibly related to the protocol treatment should be considered an adverse event.

Serious Adverse Events (SAE)

An SAE is defined in general as any undesirable medical occurrence/adverse drug experience that occurs during or within 30 days after stopping study treatment that, at any dose, results in any of the following:

- is fatal (any cause)
- life-threatening,
- requires or prolongs inpatient hospitalization,
- results in persistent or significant disability/incapacity
- is a congenital anomaly or birth defect
- is a secondary malignancy
- requires significant medical intervention

Other significant/important medical events which may jeopardize the patient are also considered serious adverse events. Serious also includes any other event that the investigator or the ETOP Safety Office judges to be serious or which is defined as serious by the regulatory agency in the country in which the event occurred.

Severity Grade

The adverse event severity grade provides a qualitative assessment of the extent or intensity of an adverse event, as determined by the investigator or as reported by the subject. *The severity grade does not reflect the clinical seriousness of the event*, only the degree or extent of the affliction or occurrence (e.g. severe nausea, mild seizure), and does not reflect the relationship to study drug.

Severity grade for other adverse events not covered in the toxicity grading scale:

1=Grade 1	Mild
2=Grade 2	Moderate
3=Grade 3	Severe
4=Grade 4	Life-threatening
5=Grade 5	Fatal

Prognostic score

Patients will be categorized in "good" or "poor" prognostic score groups according to the EORTC mesothelioma prognostic score (EPS). The formula for the calculation of the EPS is as per Curran et al (1998). Here, EPS(0.55)a + (0.60)b + (0.52)c + (0.67)d + (0.60)e

where a= if the total white blood cell count is more than 8.3 109 /L, b=if the Eastern Cooperative Oncology Group (ECOG) performance status is 1 or 2, c=if the histology is probable, d= if the histology is sarcomatous, and e=if the sex is male.

Patients are categorized into "good prognosis" if total score ≤1.27 (corresponding to having zero, one, or two poor prognostic factors) and "poor prognosis" if total score >1.27 (corresponding to having three, four, or five poor prognostic factors). The impact of the EPS "good" and "poor" categories will be explored in time-to-event analyses.

3.3 Analysis populations

Intent-to-Treat Population (ITT): The ITT population will include all patients randomized into the trial based on the initial treatment assignment and not on the treatment eventually

received. The ITT population will be used for the assessment of patient baseline characteristics and all efficacy endpoints.

As-treated population (AT): The AT population will include all patients that received at least one dose of trial treatment, with treatment assignments designated according to actual study treatment received. The AT population will be the primary population for evaluating treatment administration/compliance and safety.

Cross-over cohort (COC): Patients originally randomized in the control arm, that progressed and decided to switch to experimental treatment, will encompass the **cross-over cohort.**

4 Study Analysis

4.1 Patient accrual, balance of stratification factors and baseline characteristics

- Patient accrual by center will be presented in tabular format.
- In addition, observed vs. expected accrual will be graphically displayed.
- For patients deemed ineligible (patients registered in the online database but eventually not randomized) a table summarizing reason for non-randomization will be provided.
- Balance of treatment allocation by center and by stratification factor will be summarized as well.
- A consort diagram will be created to graphically depict the flow of patients and the phases of the trial.
- Patient baseline characteristics (categorical: gender, smoking history, ECOG performance status at diagnosis and continuous: age at randomization), will be presented overall and separately by treatment arm. Frequencies and corresponding percentages will be presented for categorical variables (if missing cases exist, a separate category named "*Missing*" will be created), while the following descriptive measures will be considered for the continuous ones: n (non-missing sample size), mean, 95% CI for the mean, median, maximum and minimum. Balance of baseline characteristics by treatment arm will be assessed via the fisher's exact test for categorical variables and the Mann-Whitney U test for continuous. Also available information on prior treatment will be summarised.
- Median follow-up (FU) of the patients (overall and by treatment arm) along with the respective interquartile range (IQR) and the number (%) of patients that are still on FU, will be summarized in a table. A Kaplan-Meier (K-M) will be also provided for a graphical representation of the respective information.

4.2 Treatment administration

Treatment information will be summarized overall and separately by treatment arm. More specifically the following information will be presented:

- Number of patients that started treatment, information on number of cycles (median, min-max). For those patients randomized in experimental arm, progressed but with physician's and their own agreement continued receiving treatment, information on treatment cycles as well as treatment failure after 1st progression will be additionally provided.
- Number of patients that did not receive any dose of trial treatment, along with reasons for not doing so
- Number of treatment failures, median TTF (95% CI), rate of TTF events and reasons for treatment failure/discontinuation will be presented overall and by treatment arm. The log-rank test will be used to compare TTF between the two treatment arms.
- A Kaplan Meier curve for TTF, overall and by treatment group will be created. In case a statistically significant difference between the two treatment arms is found the overall line will be displayed only for illustrative purposes and will be light grey and dashed. A similar approach will be followed for TTD analysis.
- The number of patients that completed treatment and whether or not they experience a PFS thereafter, will be also recorded.
- For those patients that progressed information on further lines of treatment will be also provided.

4.3 Efficacy analysis

Efficacy analysis will be performed on all randomized patients, based on their initial treatment assignment (ITT cohort).

4.3.1 Primary endpoint: PFS based on independent radiological review

Primary analysis on PFS

The study is designed to test the hypothesis that treatment of advanced malignant pleural mesothelioma with pembrolizumab will lead to an increase in median PFS (based on independent radiological review) to 6 months, from 3.5 months under standard institutional choice chemotherapy. This according to the study design, corresponds to a HR of 0.58. Using 80% power and a one-sided type I error of 2.5%, a total of 110 PFS events are needed to be observed in order to achieve the trial goal.

The total number of PFS events observed, overall and by treatment arm will be presented. In addition, 3-month, 6-month PFS estimates, median PFS and respective 95% CIs will be provided. To formally compare PFS between the two treatment arms, the stratified log-rank test will be used (with histological subtype being the stratification factor). Unstratified log-rank will be also calculated. This would be of particular value in case that the we have strata levels with very low number of patients.

To assess treatment effect on PFS, a stratified Cox proportional hazards model will be fitted, using histological sub-type as the stratification factor, adjusted for clinicopathological variables of interest (gender, age, smoking history, ECOG performance status at diagnosis, PD-L1 status, mutation load, TILs (if available, appropriately categorized). The backward elimination method, with a removal criterion at 10% will be implemented to conclude on the statistically significant variables of the model. The HRs and corresponding 95% CIs for all significant PFS predictors will be summarized in a forest plot (final model).

Graphical representation of PFS, by treatment arm will be performed via a Kaplan-Meir plot (in case of statistically significant differences between groups overall lines will be grey and dashed). The plot will be produced separately by histological subtype.

PFS will be analysed separately by subgroups defined by PD-L1 status (1% and 50% cut-off levels will be considered), mutation load and TILs. For these subgroups number of PFS events, 3-month/6-month PFS estimates, median PFS and corresponding 95% CI overall and by treatment arm will be presented. Observed differences in hazard will be assessed via the

log-rank test and will be graphically depicted by Kaplan-Meier plots. In addition, number of PFS events, median PFS and unstratified/unadjusted HRs (along with 95% CIs) will be summarized for the subgroups defined by treatment arm and the following variables: age (appropriately categorized), ECOG performance status, histological subtype, gender, EORTC score, PD-L1 status (cut-offs considered 1% & 50%), mutation load (using median value as a cut-off) and TILs. This information will be depicted in a tabular format in the report and a forest plot will be created for the presentation/manuscript, if investigators consider it interesting.

Finally, a waterfall plot for each treatment arm will be created to graphically depict the percent tumor change and a swimmer plot to illustrate efficacy information by patient (time-on-treatment, time to response and progression, follow-up), and spider plots generated split by treatment type (chemo vs pembrolizumab) to graphically overview the changes in tumour size over time with spider lines coloured by histological type (epithelioid vs non-epithelioid).

4.3.2 Secondary endpoints: Overall Survival

At the time of final analysis, the OS events will most probably not be sufficient to provide enough power for comparison between treatment. An additional OS analysis with longer follow-up will be performed at a later time point. The following will be presented at both analyses (final and longer follow-up), with the understanding that at the final analysis, statistically significant differences will not be expected.

The number of deaths will be presented overall and by treatment arm. In addition, 3-month/6-month/1-year OS estimates, median OS and respective 95% CI will be provided. OS comparison by treatment arm will be assessed using a stratified log-rank test (histological subtype will be the stratification factor, as for PFS the unstratified log-rank will be also calculated). OS, by treatment group, will be graphically displayed by Kaplan Meier plot (if a statistically significant difference between treatment groups is detected the overall line will be grey and dashed). Similar to PFS, the plot will be produced separately by histological subtype.

Similar to PFS, OS information (number of deaths, median OS, unstratified/unadjusted HRs with 95% CIs) will be summarized by subgroups defined by treatment arm and the following variables of interest: age (appropriately categorized), ECOG performance status,

histological subtype, gender, EORTC score , PD-L1 status (cut-off considered 1% & 50%), mutation load (cut-off: median value) and TILs. This information will be depicted in a tabular format in the report and a forest plot will be created for the presentation/manuscript, if investigators consider it interesting.

In addition, Restricted Mean Survival Time (RMST) will be compared between groups, to accommodate a possible delayed survival effect in the treatment group.

OS will be modelled via stratified Cox proportional hazards model (histological sub-type will be the stratification factor), adjusted for clinicopathological variables of interest: gender, age, EORTC score, ECOG performance status at diagnosis, PD-L1 status, mutation load, TILs. Same with PFS analysis, he backward elimination method, with a removal criterion at 10% will be implemented to conclude on the statistically significant variables of the model. The HRs and corresponding 95% CIs for all significant OS predictors will be summarized in a forest plot.

4.3.3 OS analysis taking into account crossover

To circumvent the selective crossover effect, a censored and an inverse probability weighted (IPW) analyses will be performed for OS (Robins et al, 2000). In the censored analysis, all observation patients who switch to receive experimental treatment at progression, will be censored the first day they received active treatment. The IPW approach is based on the idea of artificially censoring the follow-up of each patient at the time of crossover. Then, the real treatment effect is assessed by recreating the population that would have been observed without crossover through statistical modelling and weights assignment. In this way, the follow-up of patients whose experience receiving the control treatment could not be observed because they selectively switched to experimental treatment, will be replaced by the follow-up of patients with similar characteristics (both baseline and post-randomization factors) who remain in the control arm at the time of treatment switching. Fundamental to the IPW approach is the assumption of no unmeasured confounders, which implies that all common predictors are appropriately measured and accounted for in the analysis. Even though this assumption cannot be tested based on the observed data, the availability of a sufficient number of covariates and the use of both baseline and time-dependent covariates that are common predictors of the outcome of interest and artificial censoring limits the risk of important omissions. Apart from IPW, Landmark analysis will also be implemented to address cross-over.

4.3.4 Secondary & exploratory endpoints: Objective Response & Duration of response

Objective response rate (ORR) will be presented overall and separately for the two treatment arms, along with a 95% exact binomial CI. The stratified Miettinen and Nurminen's method (Miettinen O, Nurminen M., 1985) will be used for the comparison of ORR between treatment arms. The difference in ORR (with 95% CI), will be estimated using the stratification factor applied in randomization. Finally, if the number of responses allows, ORR will be presented by PD-L1, mutation load and TILs subgroups, and graphically as scatter plots, or waterfall plots.

Median DOR, along with 95% CI will be presented, for all patients and separately for the two treatment groups. Graphical representation of duration of response will be performed via swimmer plots.

Some trials of checkpoint inhibitors have suggested a treatment effect in those with grade 3 or higher immune-related adverse events. Contingent on numbers of grade 3 or more events, exploratory analyses of PFS, OS, TTF, and ORR will be performed in those with grade 3 or more immune-related adverse events or not.

4.3.5 Secondary endpoints: Investigator assessed PFS

A similar approach to the PFS as assessed by independent radiological review will be followed for the analysis of investigator assessed PFS.

4.3.6 Subgroup analysis

To determine whether the treatment effect is consistent across various subgroups the between-group treatment effect for all efficacy endpoints will be estimated within each category of the following pre-specified variables.

<u>Main subgroup analyses:</u>

- Histological subtype
- Gender
- PD-L1 (cut-offs considered: 1%)

<u>Other pre-planned subgroup analyses:</u>

- EORTC prognostic score
- PD-L1 (cut-offs considered: 50%)

- Patients with immune related grade ≥3 event vs. not
- Age group
- Mutation load
- TILs

Note: In case of imbalance in number of patients with the subgroups created by treatment and the variables of interest, this subgroup analysis will not count in the multiple comparison adjustment.

4.4 Safety analysis

The safety analysis will be performed in the AT population with the following information presented:

- Overview of the number of patients who experienced AE/SAE, as well as the number of patients in the safety cohort who did not experience an event, along with respective percentages will be shown. This information will be presented overall and by treatment arm. Also, number of patients that entered the study with baseline symptoms will be reported.
- Number of AEs/SAEs and rate of AE/SAE occurrence per month of FU (again overall and by treatment arm).
- Number of patients experiencing a specific number of AEs/SAEs, overall and by treatment arm.
- Distribution of (S)AEs by grade and CTACE category, overall and separately for the two treatment arms.
- Maximum severity of adverse events (AE/SAE) for patients, overall and by treatment arm.
- Number and corresponding percentages of treatment related (S)AEs by grade, leading either to treatment discontinuation or death will be summarized for the two trial arms. Treatment related (S)AEs (of any grade) occurring in more than 15% (or any other relevant %) will be presented for the two treatments.

- The risk difference, along with corresponding 95% CIs of specific adverse evets (most frequent i.e. ≥15%, of any grade and also focusing on grade ≥3), between the two treatment arms will be presented.
- For all fatal SAEs, cause of death will be provided.

Note: In cases where a patient may experience the same event (AE/SAE) more than once, the event will be counted only once for the calculation of the total number of events reported for the overall safety cohort, using the highest observed grade.

Finally, adverse events occurring after cross-over, will be analysed separately.

4.5 Exploratory analysis of the cross-over cohort

A similar analysis, whenever applicable will be performed (efficacy & safety) for the cross-over cohort.

4.6 Missing Data

- **Baseline characteristics**

For categorical baseline characteristics if missing cases exist, a separate category named "*Missing*" will be created. As far as continuous values, missing cases will not be replaced by any statistics calculated over non-missing data.

- **Dates:**

If the day of the month is missing for any date used in the analysis, the 15th of the month will be used to replace the missing date unless the calculation results in a negative time duration (eg, date of onset cannot be prior to day one date). In this case, the date resulting in one day of duration will be used. If the day of the month and the month are missing for any date used in a calculation, i.e., January 1 will be used to replace the missing date. Missing dates for adverse events will be imputed based on the similar principle.

- **Incomplete tumor assessment information**

In patients who have no on-study assessments:

- If death is recorded prior to the first planned tumor assessment, the death date will be considered as the date of the PFS event.

- If clinical progression is recorded prior to the first planned tumor assessment, the date of the reported clinical progression will be considered as the date of the PFS event.

4.7 Sensitivity analysis

In a sensitivity analysis framework, the efficacy analysis will be repeated using the AT population.

5 TECHNICAL DETAILS

Data will be primarily analysed using the SAS software package (version 9.4), while the R statistical software will be also used for specific analyses, such as the IPW analysis, and specific plots.

A second statistician, the reviewing statistician, will independently reproduce all analysis and summary statistics. The reviewing statistician will have an overview of the entire analysis and will explicitly check the code producing tables and figures, as well as any other pieces of code as desired.

References

Curran D, Sahmoud T, Therasse P, van Meerbeeck J, Postmus PE, Giaccone G., Prognostic factors in patients with pleural mesothelioma: the European Organization for Research and Treatment of Cancer experience. JCO 1998 Jan;16(1):145-52.

Robins JM, Finkelstein DM. Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests. Biometrics. 2000;56(3):779-788

Miettinen O, Nurminen M. Comparative analysis of two rates. Statistics in Medicine 1985; 4: 213–226.