Janssen Research & Development, LLC

Statistical Analysis Plan

A Randomized, Open-Label Study Comparing the Combination of YONDELIS® and DOXIL®/CAELYX® With DOXIL®/CAELYX® Monotherapy for the Treatment of Advanced-Relapsed Epithelial Ovarian, Primary Peritoneal or Fallopian Tube Cancer

Protocol ET743-OVC-3006: Phase 3

R279741 (trabectedin)

Original Issue Date: May 21, 2012
Final Issue Date: April 12, 2018

Document No.: EDMS-ERI-43318524: version 6.0

TABLE OF CONTENTS

1. INTRODUCTION

This statistical analysis plan (SAP) contains definitions of analysis populations, derived variables, and statistical methods for the analysis of efficacy and safety.

1.1. Trial Objectives

The primary objective of the study is to compare the overall survival (OS) after treatment with the trabectedin + DOXIL combination therapy to that observed after treatment with DOXIL monotherapy for subjects with advanced-relapsed epithelial ovarian, primary peritoneal, or fallopian tube cancer.

Secondary Objectives are:

- To evaluate PFS.
- To evaluate the objective response rate (ORR).
- To assess population pharmacokinetics (PK) of trabectedin using data from the trabectedin + DOXIL treatment group.
- To evaluate the safety of the trabectedin + DOXIL combination therapy and DOXIL monotherapy.

Exploratory Objectives are:

- To conduct pharmacogenomic evaluations of OS, PFS and other endpoints in subjects with and without mutations in BRCA1 and BRCA2.
- To evaluate patient-reported outcomes (PROs).

1.2. Trial Design

This is a randomized, open-label, active-controlled, multicenter study designed to assess the efficacy and safety of trabectedin + DOXIL as a third-line chemotherapy in subjects with platinum-sensitive advanced-relapsed epithelial ovarian, primary peritoneal, or fallopian tube cancer who received 2 previous lines of platinum-based chemotherapy. Approximately 670 subjects will be enrolled in this study, with 335 subjects per planned treatment group. At randomization, subjects will be stratified by the time from the last dose of first-line platinum therapy to disease progression (6 months to 12 months vs >12 months to 24 months vs >24 months), Eastern Cooperative Oncology Group (ECOG) performance status score (0 vs 1), BRCA 1/2 status (mutation vs no mutation), and prior pegylated liposomal doxorubicin therapy (no vs yes). Then they will be randomly assigned in a 1:1 ratio to the trabectedin + DOXIL combination therapy group (Arm A) or to the DOXIL monotherapy group (Arm B). Study treatment will be

administered by IV infusion on Day 1 of a 21-day cycle in Arm A and on Day 1 of a 28-day cycle in Arm B.

- **Treatment Arm A**: DOXIL 30 mg/m^2 90-minute IV infusion (per DOXIL package insert) followed by trabectedin 1.1 mg/m^2 3-h IV infusion, via central venous catheter, every 3 weeks. Central venous access is required for the administration of trabectedin.
- **Treatment Arm B:** DOXIL 50 mg/m^2 as a 90-minute IV infusion (per DOXIL package insert) every 4 weeks. If no infusion-related reaction is observed with DOXIL after the first infusion, DOXIL may be administered as a 30- to 90-minute infusion in subsequent cycles, as indicated in the product's Package insert.

Subjects assigned to the trabectedin + DOXIL group (Arm A) will be pretreated with 20 mg of dexamethasone (or an equivalent corticosteroid) 30 minutes prior to initiation of infusion of DOXIL IV, on Day 1 of each treatment cycle.

On 26 June 2017, the IDMC requested a futility analysis of OS to be performed at the time of the next meeting 6 months later. In view of the IDMC request, study data up to 20 September 2017 were evaluated for safety and OS (45% rather than 60% of the required death events at the planned interim analysis). Following the review of the study data by the IDMC on 15 December 2017, the HR for OS was 0.962, crossing the previously agreed upon threshold for futility of 0.93. In view of this result and the observed and expected higher toxicity in Arm A as compared with Arm B, the IDMC recommended discontinuing the study. Therefore, as of Amendment 6, no new subjects would be randomized to study treatment, and treatment with trabectedin should be immediately discontinued for subjects assigned to Arm A (trabectedin + DOXIL). All study subjects (Arm A or Arm B) currently on study treatment who, in the opinion of the investigator, were deriving clinical benefit may continue treatment with single-agent DOXIL as per the local standard of care. Treatment for these subjects may continue as long as the subjects comply with protocol-specified prohibitions and restrictions for treatment (Protocol Section 4.3) and as long as they experience clinical benefit in the opinion of the investigator. DOXIL would be provided by JR&D as needed until post-trial access to DOXIL is available or until disease progression, whichever occurs first.

Accordingly, the end of study data collection was defined as when all subjects on study treatment have completed treatment termination visit assessments as specified in the Time and Events Schedule for Amendment 6 or by 18 January 2018, whichever occurs first. For subjects continuing treatment with single-agent DOXIL, as per the local standard of care, only serious adverse events should be reported to JR&D.

1.3. Sample Size Justification

It is assumed that failure will follow an exponential distribution with a constant hazard rate. Assuming a median OS of 16 months for the active control group (DOXIL monotherapy), a planned sample size of approximately 670 subjects will provide 80% power to detect a HR of 0.78 (16 months versus 20.5 months; corresponding to a 28% improvement in median OS) at a 2-tailed significance level of 0.05 and a enrollment duration of approximately 52 months (13 subjects/month enrollment) over a total study duration of 64 months to obtained the required 514 events.

The OS endpoint will incorporate group sequential design by including 1 interim analysis and 1 final analysis using the O'Brien-Fleming boundaries as implemented by the Lan-DeMets alpha spending method. Subjects will be randomized in a 1:1 ratio to receive trabectedin + DOXIL combination therapy or DOXIL monotherapy.

2. ANALYSES PLANNED

2.1. General Analysis Definitions

2.1.1. Definition of Cycle, Treatment Phase, Baseline

Cycle: The duration of cycles may not consistently be as the planned 21 days in the trabectedin + DOXIL Arm or 28 days in the DOXIL Arm, it will depend on the earliest date of the DOXIL/trabectedin doses given in the next cycle. If a subject experiences toxicity at the planned start of a cycle that leads to a delay in DOXIL/trabectedin administration, this new cycle will be delayed, and the previous cycle will be prolonged (without additional DOXIL or trabectedin treatment).

The start of the cycle is the date of the first administration of DOXIL or trabectedin in a cycle. The end of a cycle is the day before the first infusion of the immediate subsequent cycle recorded on the case report form (CRF).

The last cycle ends 3 weeks after the start of this cycle for the trabectedin + DOXIL Arm, and 4 weeks for the DOXIL Arm.

The assessments performed during the 'End-of-treatment visit' will be included in the last treated cycle for analysis.

All lab samples taken after the last treatment and up to 30 days after the last DOXIL/trabectedin infusion will be captured in the last cycle.

End-of-treatment phase: The end of the treatment phase is the date the last cycle ends, i.e., 3 weeks after the first trabectedin or DOXIL dosing in the last cycle for Arm A, or 4 weeks after the dose of DOXIL in the last cycle for Arm B.

Follow-up phase: Following termination of DOXIL/trabectedin treatment, subjects are followed every 8 weeks for survival for the first 2 years after randomization and every 3 months thereafter until death or the clinical cut-off date defined at the time when 514 deaths are observed, whichever is sooner. If disease progression has not occurred at treatment termination, then disease assessment should continue every 8 weeks until there is evidence of disease progression or death, or until the clinical data cut-off date, or until the start of first subsequent therapy for ovarian cancer, whichever is earlier. Disease assessment will also be done if necessary to confirm response or document disease progression.

Relative day: Assessments will be presented chronologically by cycle day or study day, which are defined in the following:

Cycle day = assessment date - date of first DOXIL or trabectedin infusion for the cycle + 1

Study day = assessment date - date of first DOXIL or trabectedin infusion in cycle 1 + 1

Baseline definition:

Baseline of each parameter is defined as the latest value on or prior to the first DOXIL or trabectedin infusion in cycle 1.

In addition to the cycle day and study day defined above, the laboratory assessments completed on Day 1 of a cycle are considered to be baseline for that cycle as well as the last laboratory assessment for the previous cycle.

2.1.2. Pooling Strategy for Analysis Centers

Treatment-by-center interaction will not be formally tested. It can be assessed by graphics with regions/countries ordered according to their size.

2.1.3. Analysis Sets

"All Randomized Subjects" analysis set is defined as all subjects who are randomized to this study, independent of whether they received study drug or not.

"All Treated Subjects" analysis set is defined as all randomized subjects who receive at least 1 dose of DOXIL or trabectedin. Subjects who receive dexamethasone (or an equivalent corticosteroid) as pre-medication but who do not receive DOXIL or trabectedin will not be included in "All Treated Subject" analysis set.

The PK analysis set is defined as all subjects who received at least 1 dose of trabectedin and had at least 1 PK measurement.

2.1.3.1. Efficacy Analysis Sets

The "All Randomized Subjects" analysis set will be used.

2.1.3.2. Safety Analysis Set

The safety analysis is based on the "All Treated Subject" analysis set, where the grouping into the different arms is based on the true treatment received during the first cycle.

2.1.3.3. Definition of Subgroup

In addition to the main analysis, the efficacy analysis will be performed by the levels of stratification factors.

2.2. Methods of Analysis

Final efficacy and safety analysis for the CSR will be using clinical cut-off date of January 18, 2018.

2.2.1. Statistical Hypotheses for Trial Objectives

Overall Survival (OS) will be compared between treatment arms using an unstratified one-sided log-rank test. The trabectedin + DOXIL combination therapy will be declared better than DOXIL monotherapy if the OS is better with a p-value less than or equal to the significance level as specified by the alpha spending function. The overall 2-tailed significance level of 0.05 will be spread over 1 interim and 1 final OS analyses, when approximately 308 and 514 events on OS (death) are seen.

2.2.2. Interim Analysis

2.2.2.1. Futility Analysis

A non-binding futility analysis for OS will be implemented after observing 33% (170 events) of the total number of required 514 events. The study may be considered futile if the estimated hazard ratio (HR) from Cox proportional-hazard model is equal to or greater than 0.95. The probability of stopping under null hypothesis is approximately 62%, and the probability of stopping under alternative hypothesis is approximately 10%.

After the futility analysis at 33% of the total number of OS events, on 26 June 2017, the IDMC requested a futility analysis once again at 45% of the total number of OS events. So, a non-binding futility analysis for OS will also be implemented after observing 45% (232 events) of the total number of required 514 OS events. The study may be considered futile if the estimated hazard ratio (HR) from Cox proportional-hazard model is equal to or greater than 0.93. The probability of stopping under null hypothesis is approximately 71%, and the probability of stopping under alternative hypothesis is approximately 9%.

2.2.2.2. Interim Analysis

One interim analysis is planned for this study after observing 60% (308 events) of the total number of required (514) events. The alpha spend will be 0.008 for the interim analysis. The exact timing of the interim analysis will be determined according to the planned number of events. The O'Brien-Fleming boundaries, as implemented by Lan-DeMets alpha spending function will be used for the efficacy boundary. Operating characteristics for these boundaries are presented in the following table.

Stopping Boundaries for OS

	Analyses	
Variable	**Interim**	**Final**
Projected Observed OS Events	308	514
Anticipated Time to Analysis (months)	43	64
Anticipated Enrollment (n)	563	669
Efficacy Boundary (HR)	0.74	0.84
Boundary Crossing Prob. (H_0)	0.008	0.048
Cumulative α spent	0.008	0.050

HR=Hazard ratio; Ho = 0% improvement; H_1 = 28% improvement

An Independent Data Monitoring Committee consisting of external members including clinicians and statistician will be formed. The committee will conduct periodic reviews of the safety data and will review the efficacy data according to the prescribed statistical plan.

Following the review of the study data for the second utility analysis by the IDMC on 15 December 2017, the HR for OS was 0.962, crossing the previously agreed upon threshold for futility of 0.93. In view of this result and the observed and expected higher toxicity in Arm A as compared with Arm B, the IDMC recommended discontinuing the study. The planned interim analysis will not be performed.

2.2.3. Demographics and Baseline Characteristics

Demographics and baseline characteristics will be summarized for all randomized subjects.

Age, weight, height, and body surface area (BSA) recorded on the CRF will be summarized descriptively. Age is calculated based on the date of birth at the time date of randomization. Baseline body mass index (BMI) will be calculated, using the formula:

BMI = weight (kg) / length (m^2).

The BMI will be summarized, and the frequency counts of the different categories (<20, 20-<25, 25-<30, ≥30) will be displayed.

Age category (<65, ≥65), and race will be summarized with frequency counts.

For the ovarian cancer history, the histology and histological grade will be summarized. Time from initial diagnosis to the randomization, time from progression/recurrence during/after initial chemotherapy regimen to randomization, and time from end of first platinum-containing therapeutic line to disease progression will be calculated in months and summarized descriptively. Incomplete dates will be imputed by the methods specified in the attachment (Section 4.1).

Baseline ECOG score and BRCA1 or BRCA2 mutation status will be summarized using frequency counts. A frequency tabulation of the number of subjects with and the different types of previous cancer surgery, radiotherapy, or systemic therapy will be given. A listing of all pre-planned surgeries will be generated. A summary of medical and surgical history will be presented per system (none, history, currently active). Baseline ECG and LVEF (Left Ventricular Ejection Fraction) will be summarized.

2.2.4. Discontinuation Information

The number of subjects enrolled per center and subject disposition will be summarized by randomization arm. Subject disposition includes the number of subjects in each of the following categories:

- Randomized
    - Not Treated (subjects who are randomized but not treated with trabectedin or DOXIL
    - Treated (subjects who are treated with trabectedin or DOXIL)

An additional table will be created for all treated subjects, displaying for each treatment group the reason for treatment discontinuation by cycle.

A separate listing will be provided for subjects not treated.

Reasons for treatment termination will be collected on the CRF and will be summarized for all randomized subjects with the following categories:

- Overt progression of neoplastic disease
- Subject Choice
- Adverse Event

- Death
- Subject completed 2 cycles of treatment after confirmed CR
- Other
- Lost to follow-up

The discontinuation reason due to an adverse event is further classified into drug-related and non-drug-related adverse events.

2.2.5. Extent of Exposure

Cycle duration:

Cycle duration (weeks) is defined as:

[(The earliest dosing date of trabectedin or DOXIL in the next cycle) - (the earliest dosing date of trabectedin or DOXIL in the current cycle)]/7

For the final cycle, cycle duration is equal to the pre-planned duration, ie, 3 weeks for the trabectedin + DOXIL arm and 4 weeks for the DOXIL arm.

Treatment duration:

The treatment duration (weeks) is defined as:

[(The last cycle end date) – (the first cycle start date) + 1]/7

Cumulative dose:

Cumulative will be calculated for DOXIL and trabectedin separately.

The cumulative dose during the treatment is defined as:

Sum of dose infused across at each cycle (mg) / the Body Surface Area (m^2) at each cycle

Dose intensity:

Dose intensity will be calculated for DOXIL and trabectedin separately.

Dose intensity during the treatment is defined as:

Cumulative dose/treatment duration.

The dose intensity is expressed in mg/m²/per 3 weeks for the trabectedin + DOXIL arm, and in mg/m²/per 4 weeks for the DOXIL arm.

Relative Dose intensity:

Relative dose intensity will be calculated for DOXIL and trabectedin separately.

Relative dose intensity is the dose intensity divided by the planned dose intensity. The planned dose intensity is 1.1 mg/m²/per 3 weeks for trabectedin in the trabectedin + DOXIL arm, 30mg/m²/per 3 weeks for DOXIL in the trabectedin + DOXIL arm, and 50mg/m²/per 4 weeks for DOXIL in the DOXIL arm.

Number of Cycles:

For each subject, total number of cycles received will be calculated.

Descriptive statistics of treatment duration, cumulative dose, dose intensity, relative dose intensity, and the number of cycles received will be presented.

Cycle Delay, Dose Reduction:

The number of subjects with a cycle delay will be summarized by the number of delays. The number of subjects with a dose reduction will be summarized by the number of reductions. The reasons for cycle delay or dose reduction are further classified into drug-related adverse event (AE), non drug-related AE, or other reason. Only delays of at least 5 days will be used for the summary table. All delays will be provided on the listing.

A listing will be generated containing the data of those trabectedin and DOXIL infusions that have been interrupted along with the reason.

Dosing information of dexamethasone will be listed.

2.2.6. Protocol Deviations

Protocol Deviations will be summarized for all randomized subjects.

A summary table with the number of inclusion/exclusion violators will be presented per criterion. These subjects will be listed with the criteria that are not met.

Important protocol deviations will be documented during the trial execution based on pre-defined criteria. The detailed criteria will be documented in the data management plan and will be agreed to by the Janssen Research & Development, LLC (Janssen R&D) project physician. The classification of such protocol deviations will also be reviewed by the Janssen R&D project physician and finalized prior to the database lock. Once identified and recorded according to the rules, deviations will not be removed from the database.

Summary will be presented including the following categories:

- Selection criteria not met
- Subject not withdrawn as per protocol
- Treatment deviation
- Excluded concomitant medication
- Efficacy assessment deviation
- Other

2.2.7. Concomitant Medications

Concomitant therapies will be categorized per coded term. The number of subjects receiving each type of therapy during the treatment phase will be tabulated in 2 separate tables: a frequency tabulation of the different therapies that started pre-trial, and a frequency tabulation of the different therapies that started during the trial. The accompanying listing will contain all concomitant therapies.

2.2.8. Efficacy

2.2.8.1. Analysis Specifications

2.2.8.1.1 Level of Significance

For the primary endpoint OS, the overall (two-sided) significance level is 0.05. This 0.05 will be spread over 2 analyses by an O'Brien-Fleming boundaries as implemented by Lan-DeMets alpha spending function. The significance of efficacy will be claimed if the p-value is less than or equal to the significance level as calculated based on the above specified alpha spending function and the observed number of events.

Hierarchical Testing

This following preplanned hierarchical testing will not be performed due to lack of efficacy per the futility analysis.

Only when OS is significant, the Hochberg procedure will be used to test the secondary endpoints at the overall (two-sided) significance level 0.05: testing will begin with p value $P_{(m)}$ with corresponding hypothesis $H_{(m)}$, where $P_{(m)}$ is the largest ordered p values amongst 'm' secondary endpoints and $H_{(m)}$ is the corresponding hypothesis. If $P_{(m)} \leq \alpha$, then all hypotheses are rejected. If not, then $P_{(m-1)}$, the (m-1)th largest p-value, is compared with $\alpha/(m-1)$. If smaller, then all hypotheses from $H_{(m-1)}$ to $H_{(1)}$ are rejected. The testing procedure continues in this manner until no significant result is found

2.2.8.2. Primary Endpoint: Overall Survival

Overall survival (OS) is defined as the time between the randomization and death. Subjects who die, regardless of the cause of death, will be considered to have had an event. All subjects who are lost to the follow-up prior to the end of the trial or who are withdrawn from the trial will be censored at the time of last contact. Subjects who are still being treated will be censored at the last available date where the subject is known to be alive.

OS will be compared between treatment arms using an unstratified two-sided log-rank test as primary analysis. The Kaplan-Meier method will be used to estimate the distribution functions of OS for each treatment arm. The number of events, subjects censored, the estimate of medians and 95% confidence interval for the medians will be presented. The plots of OS using the Kaplan-Meier method will be presented.

Six-month and one-year survival rates will be calculated using the Kaplan-Meier method. The unstratified Cox proportional hazards model will be used to obtain the HR and it's 95% CI.

OS will also be compared between treatment arms by a two-sided log-rank test, stratified for the stratification factors: the time from the last dose of first-line platinum therapy to disease progression (6 to 12 months vs >12 to 24 months vs >24 months), Eastern Cooperative Oncology Group (ECOG) performance status score (0 vs 1), BRCA 1/2 status (mutation vs. no mutation) and prior pegylated liposomal doxorubicin therapy (no vs yes).

In case an imbalance in baseline prognostic factors is observed, especially PFI (platinum-free interval: the time from the last dose of first-line platinum therapy to disease progression), a Cox proportional hazards model will also be used to compare 2 treatment arms. The following baseline information includes as covariates: baseline ECOG ("0" vs. "1"), PFI (as continuous), BRCA1/2 status (mutation vs. no mutation), prior pegylated liposomal doxorubicin therapy (no vs yes), and any imbalanced factors. From the Cox proportional hazards regression, HR estimates and their 95% CIs will be estimated for treatment and for the prognostic factors.

Subgroup analyses

The following preplanned subgroup analysis will not be performed due to lack of efficacy per the futility analysis.

Subgroup analysis is planned for the primary endpoints OS to investigate whether treatment effects are consistent within subgroups. Each subgroup will be analyzed separately. The subgroups are as follows:

- Age (<65, ≥65)
- PFI (6 to 12 months vs >12 to 24 months vs >24 months),
- ECOG performance status score (0 vs 1),
- BRCA 1/2 status (mutation vs. no mutation)
- prior pegylated liposomal doxorubicin therapy (no vs yes)

2.2.8.3. Tumor Assessment

Tumor response will be assessed by investigators, according to the RECIST 1.1 response criteria.

Complete tumor assessments will be performed approximately every 8 weeks after randomization. Using RECIST 1.1 response criteria, tumor lesions are evaluated separately by target lesions and non-target lesions during each assessment. Then the combined results of target and non-target lesions will provide an overall tumor response for this assessment.

2.2.8.3.1 Response Criteria (RECIST Version 1.1)

Response evaluation for the target lesions and for the non-target lesions, and new lesions as assessed by the investigator, will be recorded on the CRF and included in the database.

For details on the RECIST Version 1.1 criteria, see Attachment 2 of the PED.

2.2.8.3.2 Evaluation of Overall Response for Each Assessment

Overall response of the subject at an assessment cycle will be determined based on investigator's assessments on target, non-target and new lesions according to the guidelines below:

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

CR = complete response; PR = partial response; SD = stable disease; PD = progressive disease; NE = Not evaluable.

For subjects with either target or non-target lesions only listed at baseline, the overall response will be evaluated by either the target or non-target lesions only (depending on which is missing), except in the case where a new lesion appears whereby the overall response would then become PD.

Subjects with a clinical deterioration of health status thought by the investigator to be related to disease progression and requiring discontinuation of treatment without objective evidence of disease progression at that time should be reported as "symptomatic deterioration". Every effort should be made to document the objective progression even after discontinuation of treatment.

Date of Progression

For the analysis based on the investigator's tumor evaluations, the day of disease progression is determined by using the earliest time point from the following dates: the date of PD as evaluation of response on the CRF; the date of a new lesion on the CRF tumor measurements page.

For the analysis based on all investigator's data, the day of disease progression is determined by using the earliest time point from the following dates: the date of PD as evaluation of response on the CRF; the date of a new lesion on the CRF tumor measurements page; the death date (if death is caused by PD);, the progression date as documented on the CRF treatment termination page, the progression date as documented on the CRF follow-up visit page.

Best Overall response

The best overall response is the best response recorded from the start of the treatment until disease progression or start of subsequent anticancer therapy. The best overall

response will be summarized per treatment group in a frequency table with categories: complete response (CR), partial response (PR), stable disease (SD), progressive disease (PD), not evaluable (NE).

2.2.8.4. Secondary Endpoint: Progression-Free survival

The analysis of PFS will be based on the data received from the investigator, combined with information from the database on the subject's death and on subsequent therapy for ovarian cancer.

Progression free survival (PFS) is defined as the time between the randomization and disease progression or death. Subjects who progressed or died will be considered to have had an event, except if this event occurs after the start of subsequent therapy for ovarian cancer, in which case the subject is censored at the time of last tumor assessment (prior to or on the first day of the first subsequent therapy for ovarian cancer). Subjects who did not progress nor died (lost to follow-up, or still being treated without documented disease progression, or started subsequent therapy for ovarian cancer and still alive) will be censored at the date of the last tumor assessment (prior to or on the first day of the first subsequent therapy for ovarian cancer).

2.2.8.4.1 Handling of Missing Assessments

Tumor scans are scheduled every 8 weeks for the first 4 assessments and then every 12 weeks. Despite this, assessments are sometimes missing (not performed, not all lesions measured, lost, technically inadequate). If there is a time interval between the last non-PD scan and the date of death or the date of progression, the following algorithm will be applied:

- if a subject has an interval of ≥18 weeks during the first 4 assessment or ≥26 weeks thereafter, this is an indication that more than one measurement is missing. Then the PFS is censored at the last disease assessment prior to the interval.
- if a subject has an interval of <18 weeks prior to the fourth assessment or <26 weeks after the first 4 assessments, this is an indication that there is only 1 measurement missing, and then the subsequent information (PD date, death date) is used as event.

2.2.8.4.2 Analysis Methods

PFS will be compared between treatment arms using an unstratified two-sided log-rank test. The Kaplan-Meier method will be used to estimate the distribution functions of PFS for each treatment arm. The number of events, subjects censored, the estimate of medians and 95% confidence interval for the medians will be presented. The plots of PFS using the Kaplan-Meier method will be presented. Six-month and one-year progression-free rates will be calculated using the Kaplan-Meier method. The

unstratified Cox proportional hazards model will be used to obtain the HR and it's 95% CI.

2.2.8.5. Other Secondary Efficacy Analyses

2.2.8.5.1 Definition of Other Secondary Efficacy Variables

Objective Response

Objective response is defined as having "CR" or "PR" as best overall response based on reconciled tumor assessment. Detail tumor response evaluation is described previously. The ***objective response rate*** is calculated as the number of objective responders divided by the number of subjects.

The response rate variable will be evaluated using the Chi-square statistic (unstratified).

2.2.8.5.2 Analysis Methods

The best overall response will be summarized per treatment arm in a frequency table with categories: complete response (CR), partial response (PR), stable disease (SD), progressive disease (PD), not evaluable (NE).

Subjects' overall response per cycle assessment will be determined and will be presented in a listing.

The response rate variable will be evaluated using the Chi-square statistic (unstratified).

2.2.8.6. Supportive Analyses

2.2.8.6.1 Symmetry of Tumor Assessment Schedules

Tumor assessments were to be performed every 8 weeks. Timing of assessments in both treatment groups will be presented side by side with boxplots.

2.2.8.7. ECOG

The ECOG value at the end of treatment visit will be tabulated (not cross-tabulated against the baseline value). The listing will contain both ECOG values.

2.2.8.8. Patient Reported Outcomes

The following preplanned patient reported outcome analysis will not be performed due to lack of efficacy per the futility analysis.

The Patient Reported Outcomes will be measured by the OV28 and the EQ-5D questionnaires. Patients will be requested to fill these out at baseline, at the end of every cycle, at the treatment termination visit, and every 8 weeks till start of subsequent treatment. The analyses of these data will be performed on summary scores as well as subscales, and individual symptoms.

Change from baseline at each time point and treatment termination will be analysed for selected scales of the OV28 items. In particular, Abdominal/Gastrointestinal Symptoms scales of OV28 will be tested based on the grouping criteria as specified in Greimel et al.[1] Mean score changes from baseline at each time point and treatment termination will be compared between the treatment arms using T–test. The changes will also be analyzed using model for longitudinal data, taking the potential missing data and dropout into account.

2.2.8.9. Subsequent treatment

Frequency tabulation will be given, indicating how many subjects in each arm started subsequent therapy for ovarian cancer.

The accompanying listing will contain details on the type of subsequent therapy for ovarian cancer as well as the start date.

2.2.9. Safety

2.2.9.1. Adverse Events

Treatment emergent adverse events will be summarized by disease group and overall. A treatment-emergent adverse event is defined as any AE occurring or worsening on or after the first treatment of DOXIL/Trabectedin, and within 30 days after the last dose. For imputation rules for incomplete dates, see Attachment 4.1.

Adverse events are documented on the CRF together with their severity, according to the NCI CTC version 4.0, also referred to as NCI toxicity grading. For the categorization of the adverse events, the MedDRA dictionary will be used. Adverse events are considered drug related when the relation to trial medication is considered to be possible, probable, or very likely, according to the investigator's opinion.

Summary of overall adverse events will be done by body system and preferred term, by severity (worst toxicity grade), by relationship to the study drugs (trabectedin or DOXIL), and by AE outcome. Tables will be sorted by body system/preferred term and by the highest incidence. In the AE database, "neoplasm malignant aggravated", and "condition aggravated" are created terms to match the Global Medical Safety (GMS) death reporting database (in case of death due to disease progression). These will not be included in the AE summaries.

A listing will be prepared on all adverse events. If data is available, a listing for delayed adverse events (defined as new onset of AE occurring 30 days after last dose as judged by investigator to be related to trial medication) will be presented.

First occurrence of grade 3 and 4 adverse events will be reported by cycle. The drug related grade 3 and 4 adverse events would be summarized.

A frequency table will be made for the AEs leading to cycle delay, dose reduction, skipped dose, or withdrawal of study medication. Adverse events leading to permanent stop and AE with outcome death will also be presented by drug relatedness.

For all serious adverse events (SAEs), the investigator has to send an SAE form to the Janssen R&D GMS. The AEs that are indicated on the CRF as serious are used for the summary tables. A summary table by any grade SAEs will be presented. A similar table will be presented for the drug-related SAEs.

2.2.9.2. Clinical Laboratory Tests

2.2.9.2.1 Hematology

Laboratory results will be classified according to the NCI CTC version 4.0 as described in the attachment of the protocol. For absolute neutrophil count (ANC), platelet count and hemoglobin, the worst grade per subject will be tabulated overall during treatment and per cycle.

Cross tabulation will be presented for the worst grade during treatment versus the baseline toxicity grading.

All data will be listed. In addition, a listing will present subjects' data by cycle and laboratory test for those whose laboratory values reach grade 4 anytime during the trial.

2.2.9.2.2 Serum chemistry

Similar to hematology analysis, the worst grade during treatment will be cross tabulated to the baseline grade for sodium, alanine aminotransferase [ALT], potassium, total bilirubin, direct bilirubin, blood urea nitrogen [BUN], alkaline phosphatase [ALP], creatinine, creatine phosphokinase [CPK], glucose, albumin, aspartate aminotransferase [AST], and total protein).

Cross tabulation will be presented for the worst grade during treatment versus the baseline toxicity grading.

All data will be listed. In addition, a listing will present subjects' data by cycle and laboratory test for those whose laboratory values reach grade 4 anytime during the trial.

2.2.9.2.3 Hy's Law[2] Cases Analysis

For subjects with any elevated AT (AST or ALT) of >3xULN, ALP <2xULN, and associated with an increase in bilirubin ≥2xULN, a listing for all subjects with all such records will be produced and a summary table of number of such subjects by treatment arm will also be generated.

2.2.9.3. Vital Signs and Physical Examination Findings

Vital signs (pulse and blood pressure, as well as temperature) and body weight will be measured during the screening phase (within 14 days of randomization). Tabulation will be made summarizing the vital signs and the body weight.

A physical examination will be done during the screening phase (within 14 days of randomization).

2.2.9.4. Deaths

Deaths during treatment or within 30 days from the last DOXIL/Trabectedin treatment administration will be tabulated, as well as their primary cause of death. A distinction will be made between drug related and non-drug related adverse events.

An additional table will specify the number of subjects who are still alive, who died within 30 days after the last treatment (together with the reason), and who died more than 30 days after the last treatment (together with the reason).

2.2.9.5. ANALYSIS OF CARDIAC SAFETY

For detailed cardiac safety analysis results please refer to post marketing request report for soft tissue sarcoma submission.

2.2.10. Safety Narratives

Safety narratives will be written for the following 3 groups of subjects:

1. Deaths for reasons other than disease progression that occurred within 30 days of the last dose of study medication;
2. Treatment-emergent serious adverse events;
3. Treatment-emergent adverse events that led to discontinuation of study medication

2.2.11. Clinical Pharmacology Analyses

Sparse blood samples were collected to measure trabectedin plasma concentrations at the specified times as shown in the schedule of events of the protocol. Population PK analysis will be based on PK-evaluable data set.

Population PK of trabectedin will be modelled from the data obtained in the phase III ET743-OVC-3006 study. Since PK characteristics of trabectedin have been extensively

studied in adult cancer population, the population PK model developed previously in adults (using data from phase I and phase II studies) will be utilized initially for describing trabectedin PK data from ET743-OVC-3006 study. A detailed analysis plan for population PK analysis will be developed separately, and a stand-alone report will be written to summarize the results of the population PK analysis.

3. REFERENCES

1. Greimel E, et al., An international field study of the reliability and validity of a disease-specific questionnaire module (the QLQ-OV28) in assessing the quality of life of patients with ovarian cancer. European Journal of Cancer, Volume 39, Issue 10 , Page 1402

2. Guidance for Industry, Drug-Induced Liver Injury: Premarketing Clinical Evaluation Oct 2007.

4. ATTACHMENTS

4.1. Imputation Rules for Incomplete Dates

The following method is only for general reporting purpose. In the case of primary analysis, the imputation method has to be clearly and carefully specified in the Statistical Analysis Plan.

The dates of certain historical or current clinical activities are key component for statistical analysis. Incomplete date appears when day, month or year is/are missing, and it could be imputed so that variables like time to and duration of certain event can be calculated. If none of day, month and year is available, then the date is missing, no imputation is necessary.

The following dates are usually used for baseline characteristics and safety summaries.

1. Date of initial diagnosis
2. Date of the first progression/recurrence
3. Date of the last progression/recurrence
4. Date of chemotherapy
5. Date of radiation
6. Date of surgery
7. AE start date
8. AE end date
9. Follow-up therapy start date
10. Date of death

Dates 1 to 6 will be mentioned as historic dates in the context. All above mentioned dates can be either start date or the end date in the calculation depending on the analysis, the imputation result could be different for the same date while used either as the start date or the end date. General rules for imputation will be specified for the start date and end date respectively.

Historic Dates

For historic dates such as 1-6, the imputation will not distinguish between the usage of the start date or the end date, rules are as following:

- If day is unknown, the imputed date = 15th of the same month
- If day and month are missing, and the first visit started before July of the *same* year, the imputed date = 15th of the month prior to the month where the first visit started

- If day and month are missing, and the first visit started on or after last of July of the same year, the imputed date = 1st of July
- If day and month are missing, and the recorded year is before the first visit year, the imputed date = 1st of July of the recorded year

See example below:

First Visit Date	Recorded Date	Imputed Date
June 20, 2002	2002	May 15, 2002
July15, 2002	2002	July 1, 2002
July 15, 2002	1999	July 1, 1999

- If year is missing, the imputed date is missing.

Adverse Event Dates

If incomplete dates exist after data cleaning and are used as the **Start Date,** the following imputation will be applied:

- If month and year known, and if the study treatment started in the same month, the imputed date = the date of the first treatment. Otherwise, imputed date = 1st of that month
- If day and month are missing, the imputed date = the first treatment date if the treatment started on the same year or Jan 1st if the treatment started before the recorded year.

For example:

Screen Visit Date	First treatment Date	Recorded Date	Imputed Date
May 20, 2002	June 6, 2002	June, 2002	June 6, 2002
May 20, 2002	June 1, 2002	2002	June 1, 2002
May 20, 2002	June 1, 2002	2003	Jan 1, 2003

- If year is missing, the imputed date is missing.

If incomplete dates exist after data cleaning and are used as the **End Date,** the following imputation will be applied:

- If month and year known, and if the month is the same as the month of the last treatment date plus 30 days or the death, then the imputed date = the earlier date of these two dates in that month. Otherwise, the imputed date = last of the same month.

For example:

Last treatment Date	Death Date	Recorded Date	Imputed Date
May 20, 2002	June 6, 2002	June, 2002	June 6, 2002
May 20, 2002	Dec 6, 2002	June, 2002	June 19, 2002
Nov 20, 2002	Aug 6, 2003	June, 2002	June 30, 2002

- If day and month are missing, and no death record in the same year, the imputed date = the last treatment date plus 30 days or the last date of the same year, whichever is earlier. If death did occur in the same year, the imputed date = the death date or the last treatment date plus 30 days, whichever is earlier.

For example:

Last treatment Date	Death Date	Recorded Date	Imputed Date
Dec 20, 2002	June 6, 2003	2002	Dec 31, 2002
August 1, 2002	Aug 15, 2002	2002	Aug 15, 2002
May 1, 2002	N/A	2002	May 31, 2002

- If year is missing, the imputed date is missing.

If any date involved in the summary calculation is still missing after the imputation, then the calculation result (such as AE duration, time from first diagnosis, etc.) will be missing. For AEs, if both the start date and the end date are missing, it will still be considered as treatment emergent AE. For the analysis of time to and duration of certain AEs, if the end date is missing, the end date will be censored at the last treatment date plus 30 days.

Follow-up therapy start date

In case the exact day of the start date of the follow-up therapy is missing, the following imputation rule will be applied:

The exact date is the latest of the following: Day 1 of that month, PD date during that month + 1 day, or study drug dosing date during that month + 1 day.

Date of death

In case the exact day of the date of death is missing, the following imputation rule will be applied:

The exact date is the latest of the following: Day 1 of that month, or the date of last clinical data point during that month + 1 day.