ONO-4538 Phase III Study

A Multicenter, Randomized, Open-label Study in Patients with esophageal Cancer refractory or intolerant to Combination Therapy with Fluoropyrimidine and Platinum-based Drugs

Protocol

Ono Pharmaceutical Co., Ltd.

This study is being conducted in collaboration with co-development partner, Bristol-Myers Squibb (BMS): BMS IND Number: 126,406; EUDRACT Number: 2015-003339-36.

Protocol No.: ONO-4538-24/BMS CA209473

Version No.: 9.0

Date of preparation: November 7, 2017

Confidentiality Statement

This document is confidential and proprietary information of the company sponsoring the study and its global affiliates. By reviewing this document, you agree to keep it confidential and to use and disclose it solely for the purpose of assessing whether your organization will participate in and/or perform the proposed study. Any permitted disclosures will be made only on a confidential "need to know" basis within your organization or to your independent ethics committee(s).

Any other use, copying, disclosure or dissemination of information in this document is strictly prohibited unless expressly authorized in writing by the sponsor. Any supplemental information (e.g., amendments) that may be added to this document is also confidential and proprietary to the sponsor and must be kept confidential in the same manner as the contents of this document. Any person who receives this document without due authorization from the sponsor is requested to return it to the sponsor or promptly destroy it. All other rights reserved. References to the sponsor in this protocol may apply to partners to which the sponsor has transferred obligations, e.g., a contract research organization (CRO).

Protocol Synopsis

1 Objective of the Study

This is a multicenter, randomized, open-label, docetaxel- or paclitaxel-controlled study to evaluate the efficacy and safety of ONO-4538 in patients with esophageal cancer refractory or intolerant to combination therapy with fluoropyrimidine and platinum-based drugs.

2 Study Design

A multicenter, randomized, open-label study

2.1 Study Design

This is a multicenter, randomized, open-label, docetaxel- or paclitaxel-controlled study where ONO-4538 is administered at 2-week intervals for the treatment of patients with esophageal cancer refractory or intolerant to combination therapy with fluoropyrimidine and platinum-based drugs, with 1 prior regimen to evaluate the efficacy and safety of ONO-4538. Patients will be randomized in a 1:1 ratio to the ONO-4538 group or control group (docetaxel or paclitaxel), and will be stratified by region (Japan vs. rest of world) and number of organs with metastases (≤1 vs. ≥2) and expression of PD-L1 (≥1% vs. <1% or indeterminate). After randomization, the ONO-4538 group will receive ONO-4538 (240 mg at 2-week intervals) and the control group will receive docetaxel (75 mg/m^2 at 3-week intervals) or paclitaxel (100 mg/m^2 weekly for 6 weeks in succession followed by a 2-week drug holiday). The primary endpoint is OS, based on which the superiority of ONO-4538 over the control group will be investigated. The study design overview is presented in Figure 2-1.

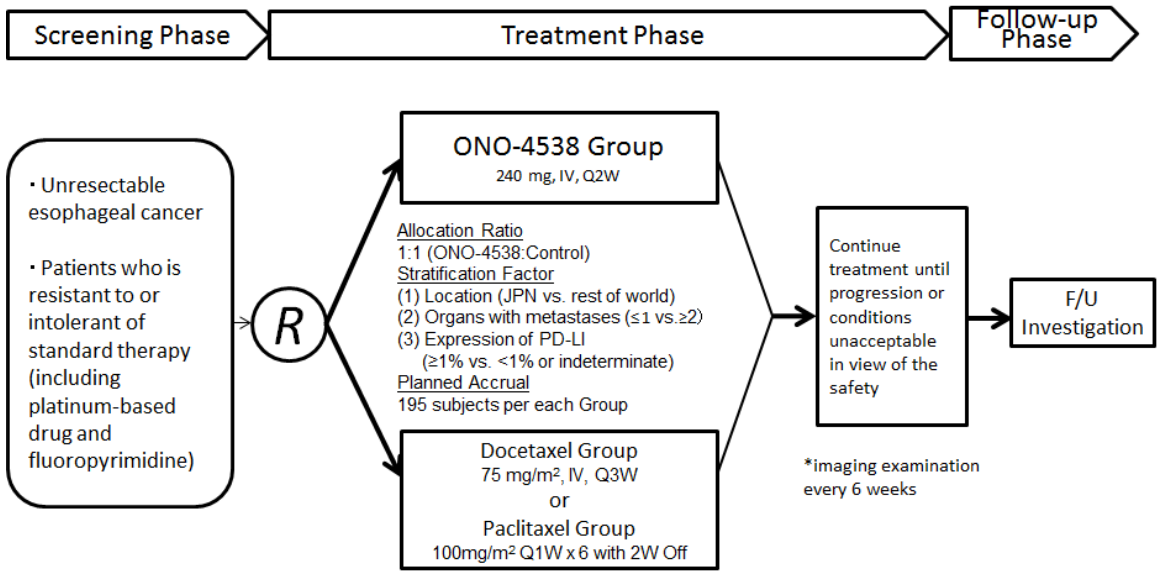


Figure 2-1 Outline of Study Design

3 Subjects

3.1 Subjects

Patients with esophageal cancer who are refractory or intolerant to combination therapy with fluoropyrimidine and platinum-based drugs will be enrolled.

3.2 Inclusion Criteria

Having provided written consent before participation in the study, patients must fulfill all of the following criteria to be eligible for randomization. If a randomized subject does not meet any one of the following criteria before the first dose of the investigational product after randomization, the subject will not be started on the study treatment and will be withdrawn from the study.

1. Sex: Male or female
2. Age (at the time of informed consent): 20 years and older
3. Patients with esophageal cancer and whose major lesion in the esophagus (if already resected, the major lesion in the esophagus prior to resection) satisfies the following criteria. For patients with multiple lesions, the deepest invasion by clinical diagnosis should be considered the major lesion. Lesions other than the major lesion should be considered as secondary lesions. Esophageal cancer in this study is defined as a cancer that has primarily developed from the esophagus.
    - Patients with a major lesion located in the cervical esophagus or thoracic esophagus (upper, middle, or lower thoracic region; including the esophagogastric junction)
    - Patients whose histological type of major lesion was pathologically proven squamous cell carcinoma or adenosquamous cell carcinoma (pathological diagnosis of secondary lesions in the esophagus is not mandatory). Note: adenosquamous is defined as an uncommon malignant esophageal neoplasm containing coexisting elements of infiltrating squamous cell carcinoma (SCC) and adenocarcinoma (AC)
4. Patients who are refractory or intolerant to combination therapy with fluoropyrimidine and platinum-based drugs for esophageal cancer, have previously received 1 treatment regimen, and are not indicated for a radical resection. The definition of refractory should be defined as follows; and a therapy applicable to the following should be counted as 1 regimen.
    - Patients whose PD or recurrence was confirmed by imaging during their initial chemotherapy (including chemoradiation) or within 8 weeks after the last dose[#1] of chemotherapy will be assessed as “refractory.”

- Patients who underwent a radical resection (R0 resection confirmed) in conjunction with chemotherapy including neo-adjuvant/adjuvant therapy and chemoradiation (including patients who underwent chemoradiation followed by salvage surgery) whose recurrence was confirmed by imaging within 24 weeks after the last dose[#1] of chemotherapy will be determined as "refractory."
- If a CR (≥2 consecutive CRs confirmed by imaging after an interval of ≥4 weeks) was assessed as a result of the initial chemotherapy (including chemoradiation), patients whose recurrence was confirmed by imaging during the initial chemotherapy (including chemoradiation) or within 24 weeks after the last dose[#1] of chemotherapy will be determined as "refractory."

[#1] In case of chemoradiation, this will be the last dose of chemotherapy or the last radiation treatment, whichever occurs later.

5. Patients who have at least 1 measurable or non-measurable lesion per the RECIST Guideline Ver. 1.1 as confirmed by imaging within 28 days before randomization. The following requirements should also be satisfied:
   - The primary esophageal cancer should be deemed to be non-measurable lesion.
   - If patients only have lesions that were previously treated with radiation, the lesion should be limited to one with confirmed aggravation by imaging after radiation.
   - If patients have pericardial or pleural effusion or ascites only, the lesion should be limited to one with cytologically confirmed malignancy.
6. ECOG Performance Status Score (see Appendix 3) 0 or 1
7. Patients with a life expectancy of at least 3 months
8. Patients must provide tumor tissue (stored tissue or tissue from the last biopsy) for analysis of PD-L1 expression. For patients who are unable to undergo another biopsy, stored tissue can be used for analysis. Tissue specimens must contain at least 100 evaluable tumor cells and must be available before the randomization.
9. Patients whose latest laboratory data meet the below criteria within 7 days before randomization. If the date of the laboratory tests at the time of randomization is not within 7 days before the first dose of the investigational product, testing must be repeated within 7 days before the first dose of the investigational product, and these latest laboratory tests must meet the following criteria. Of note, laboratory data will not be valid if the patient has received a granulocyte colony-stimulating factor (G-CSF) or blood transfusion within 14 days before testing.
   - White blood cells ≥2,000/mm^3 and neutrophils ≥1,500/mm^3

- Platelets ≥100,000/mm^3
- Hemoglobin ≥9.0 g/dL
- AST (GOT) and ALT (GPT) ≤3.0-fold the upper limit of normal (ULN) of the study site (or ≤5.0-fold the ULN of the study site in patients with liver metastases)
- Total bilirubin ≤1.5-fold the ULN of the study site
- Creatinine ≤1.5-fold the ULN of the study site or creatinine clearance (either the measured or estimated value using the Cockcroft-Gault equation) >45 mL/min

10. Women of childbearing potential (including women with chemical menopause or no menstruation for other medical reasons)[#1] must agree to use contraception[#2] from the time of informed consent until 5 months or more after the last dose of the investigational product. Also, women must agree not to breastfeed from the time of informed consent until 5 months or more after the last dose of the investigational product.
11. Men must agree to use contraception[#2] from the start of study treatment until 7 months or more after the last dose of the investigational product.

[#1] Women of childbearing potential are defined as all women after the onset of menstruation who are not postmenopausal and have not been surgically sterilized (e.g., hysterectomy, bilateral tubal ligation, bilateral oophorectomy). Postmenopause is defined as amenorrhea for ≥12 consecutive months without specific reasons. Women using oral contraceptives, intrauterine devices, or mechanical contraception such as contraceptive barriers are regarded as having childbearing potential.

[#2] The subject must consent to use any two of the following methods of contraception: vasectomy or condom for patients who are male or female subject's partner and tubal ligation, contraceptive diaphragm, intrauterine device, spermicide, or oral contraceptive for patients who are female or male subject's partner.

3.3 Exclusion Criteria

Patients meeting any of the following criteria at the time of assessment for randomization will be excluded. If patients meet any of the following criteria prior to the first dose of the investigational product after randomization, the study treatment should not be initiated.

1. Patients with significant malnutrition. Patients will be excluded if they are receiving intravenous hyperalimentation, or require continuous infusion therapy with hospitalization. Patients whose nutrition has been well controlled for ≥28 days prior to randomization may be enrolled.
2. Patients with apparent tumor invasion on organs located adjacent to the esophageal disease (e.g., the aorta or respiratory tract). Patients will be excluded if they are receiving stent therapy in esophagus or respiratory tract.
3. Patients with multiple primary cancers (with the exception of completely resected basal cell carcinoma, stage I squamous cell carcinoma, carcinoma in situ, intramucosal carcinoma, or superficial bladder cancer, or any other cancer that has not recurred for at least 5 years)
4. Patients with residual adverse effects of prior therapy or effects of surgery that would affect the safety evaluation of the investigational product in the opinion of the investigator or subinvestigator
5. Patients with current or past history of severe hypersensitivity to any other antibody products
6. Patients with concurrent autoimmune disease or history of chronic or recurrent autoimmune disease (see Appendix 4)
7. Patients with a current or past history of interstitial lung disease or pulmonary fibrosis diagnosed based on imaging or clinical findings. Patients with radiation pneumonitis may be randomized if the radiation pneumonitis has been confirmed as stable (beyond acute phase) without any concerns about recurrence.
8. Patients with concurrent diverticulitis or symptomatic gastrointestinal ulcerative disease
9. Patients with any metastasis in the brain or meninx that is symptomatic or requires treatment. Patients may be randomized if the metastasis is asymptomatic and requires no treatment.
10. Patients with pericardial fluid, pleural effusion, or ascites requiring treatment
11. Patients with uncontrollable, tumor-related pain
12. Patients who have experienced a transient ischemic attack, cerebrovascular accident, thrombosis, or thromboembolism (pulmonary arterial embolism or deep vein thrombosis) within 180 days before randomization
13. Patients with a history of uncontrollable or significant cardiovascular disease meeting any of the following criteria:
    - Myocardial infarction within 180 days before randomization
    - Uncontrollable angina pectoris within 180 days before randomization
    - New York Heart Association (NYHA) Class III or IV congestive heart failure

- Uncontrollable hypertension despite appropriate treatment (e.g., systolic blood pressure ≥150 mmHg or diastolic blood pressure ≥90 mmHg lasting 24 hours or more)
- Arrhythmia requiring treatment

14. Patients receiving or requiring anticoagulant therapy for a disease. Patients receiving antiplatelet therapy including low-dose aspirin may be enrolled.
15. Patients with uncontrollable diabetes mellitus
16. Patients with systemic infections requiring treatment
17. Patients with ≥ Grade 2 peripheral neuropathy
18. Patients who have received systemic corticosteroids (except for temporary use, e.g., for examination or prophylaxis of allergic reactions) or immunosuppressants within 28 days before randomization
19. Patients who have received antineoplastic drugs (e.g., chemotherapy agents, molecular-targeted therapy agents, or immunotherapy agents) within 28 days before randomization
20. Patients who have undergone surgical adhesion of the pleura or pericardium within 28 days before randomization
21. Patients who have undergone surgery under general anesthesia within 28 days before randomization
22. Patients who have undergone surgery involving local or topical anesthesia within 14 days before randomization
23. Patients who have received radiotherapy within 28 days before randomization, or radiotherapy to bone metastases within 14 days before randomization
24. Patients who have received any radiopharmaceuticals (except for examination or diagnostic use of radiopharmaceuticals) within 56 days before randomization
25. Patients with a positive test result for any of the following: HIV-1 antibody, HIV-2 antibody, HTLV-1 antibody, HBs antigen, or HCV antibody
26. Patients with a negative HBs antigen test but a positive test result for either HBs antibody or HBc antibody with a detectable level of HBV-DNA
27. Women who are pregnant or breastfeeding, or possibly pregnant
28. Patients who have received any other unapproved drug (e.g., investigational use of drugs, unapproved combined formulations, or unapproved dosage forms) within 28 days before randomization
29. Patients who have previously received taxane agents to treat esophageal cancer. Patients who were not proven refractory (see the definition of refractory in inclusion criteria 4) or intolerant to

taxane-based combination therapy and subsequently received fluoropyrimidine and platinum-based combination therapy, and then proven refractory (see the definition of refractory in inclusion criteria 4) or intolerant may be randomized.

30. Patients who are contraindicated to docetaxel and paclitaxel
31. Patients who have previously received ONO-4538 (MDX-1106 or BMS-936558), anti-PD-1 antibody, anti-PD-L1 antibody, anti-PD-L2 antibody, anti-CD137 antibody, anti-CTLA-4 antibody or other therapeutic antibodies or pharmacotherapies for regulation of T-cells
32. Patients judged to be incapable of providing consent for reasons such as concurrent dementia
33. Other patients judged by the investigator or subinvestigator to be inappropriate as subjects of this study

4 Investigational Product

4.1 Dosage, Administration, and Treatment Duration

4.1.1 Dosage, Administration, and Treatment Duration of ONO-4538

ONO-4538 240 mg will be intravenously administered over 30 minutes at 2-week intervals. Six weeks will count as 1 cycle of treatment. Treatment will be continued until PD is assessed by the investigator or subinvestigator according to the RECIST Guideline Ver. 1.1. ONO-4538 will be administered at 240 mg without any increase or decrease of the dose and when 11 or more days have passed since the last dose, i.e., at an interval of 10 days or more.

For information regarding appropriate storage, handling, dispensing, and administration of ONO-4538, refer to the latest Investigator's Brochure and/or Reference Sheet for Pharmacy.

4.1.2 Dosage, Administration, and Treatment Duration of Docetaxel

The initial dose of docetaxel is 75 mg/m^2, which will be intravenously administered at 3-week intervals. Each cycle of treatment lasts 3 weeks. Treatment will be continued until PD is assessed by the investigator or subinvestigator according to the RECIST Guideline Ver 1.1. If the subject has had a ≥10% change in body weight from that measured at the initial dose, the dose should be adjusted. If the change is observed on the day of administration, the dose can be adjusted from the next administration onwards. Each dose should be similarly adjusted if a further body weight change of ≥10% is observed after the previous change. The dose (mg) will be rounded to one decimal place. The recommended procedures for docetaxel administration are described below; however, it can be administered in accordance with the study site procedures. Docetaxel will be administered intravenously over at least 60 minutes in accordance with the package insert. Adverse reactions to

docetaxel must be treated appropriately with reference to local standards, such as the package insert and treatment guidelines.

4.1.3 Dosage, Administration, and Treatment Duration of Paclitaxel

The initial dose of paclitaxel is 100 mg/m^2, which will be administered once weekly for 6 consecutive weeks, followed by a 2-week drug holiday. This treatment cycle will be repeated until PD is assessed by the investigator or subinvestigator according to the RECIST Guideline Ver 1.1. There should be an interval of at least 5 days between the doses of paclitaxel. The dose on Day 1 of each cycle should be continued throughout that cycle as a rule. If the subject has had a ≥10% change in body weight from that measured at the initial dose, the dose should be adjusted. If the change is observed on the day of administration, the dose can be adjusted from the next administration onwards. Each dose should be similarly adjusted if a further body weight change of ≥10% is observed after the previous change. The dose (mg) will be rounded to one decimal place. The recommended procedures for paclitaxel administration are described below; however, it can be administered in accordance with the study site procedures.

Paclitaxel will be administered intravenously through an in-line filter using a membrane filter of ≤0.22 μm over 60 minutes in accordance with the package insert. Adverse reactions to paclitaxel must be treated appropriately with reference to local standards, such as the package insert and treatment guidelines.

5 Prior and Concomitant Therapies

5.1 Prohibited Therapies During the Study Period

5.1.1 Prohibited Therapies During the Study Period (ONO-4538 Group)

The following treatments are prohibited throughout the study period (i.e., after informed consent and until completion of the final examination stipulated by the protocol [with the exception of the outcome investigation]), unless absolutely necessary for medical reasons even after the examination at the end of the treatment phase (at discontinuation) has been completed.

1. Immunosuppressants and corticosteroids[1, 5]
2. Any antitumor therapies[5] (including chemotherapy, molecular target therapy, and immunotherapy[2])
3. Surgical treatment for malignant tumors[5]
4. (Chemo) radiotherapy[5]

5. Radiopharmaceuticals[3, 5]
6. Bisphosphonates and anti-RANKL antibody products[4]
7. Transplantation
8. Other unapproved drugs[5] (e.g., investigational use of drugs, unapproved combined formulations, and unapproved dosage forms)

[1] Corticosteroids may only be used topically (e.g., external, intra-articular, intranasal, ophthalmic, or inhalational use) or temporarily (e.g., for the treatment or prophylaxis of contrast medium allergy or AEs).

[2] These include local therapeutic agents such as picibanil.

[3] Use for examination or diagnostic use of radiopharmaceuticals is allowed.

[4] Treatment with bisphosphonates or anti-RANKL antibody products that has been ongoing since before the subject's first dose of study treatment may be continued only if there are no changes made to the dosage and mode of administration.

[5] Before randomization, these therapies may be used unless they are performed in the respective periods specified in the exclusion criteria #18～24, 28.

5.1.2 Prohibited Therapies During the Study Period (Control Group)

The following treatments are prohibited throughout the study period (i.e., after informed consent and until completion of the final examination stipulated by the protocol [with the exception of the outcome investigation]), unless absolutely necessary for medical reasons even after the examination at the end of the treatment phase (at discontinuation) has been completed.

1. Immunosuppressants and corticosteroids[1, 5]
2. Any antitumor therapies[5] (including chemotherapy, molecular target therapy, and immunotherapy [2])
3. Surgical treatment for malignant tumors[5]
4. (Chemo) radiotherapy[5]
5. Radiopharmaceuticals[3, 5]
6. Bisphosphonates and anti-RANKL antibody products[4]
7. Transplantation
8. Disulfiram, cyanamide, carmofur, and procarbazine hydrochloride (for the paclitaxel group only)

9. Other unapproved drugs[5] (e.g., investigational use of drugs, unapproved combined formulations, and unapproved dosage forms)

[1] Corticosteroids may only be used topically (e.g., external, intra-articular, intranasal, ophthalmic, or inhalational use) or temporarily (e.g., for the treatment or prophylaxis of contrast medium allergy or AEs).

[2] These include local therapeutic agents such as picibanil.

[3] Use for examination or diagnostic use of radiopharmaceuticals is allowed.

[4] Treatment with bisphosphonates or anti-RANKL antibody products that has been ongoing since before the subject's first dose of study treatment may be continued only if there are no changes made to the dosage and mode of administration.

[5] Before randomization, these therapies may be used unless they are performed in the respective periods specified in the exclusion criteria #18～24, 28.

5.2 Prophylactic Premedication

5.2.1 ONO-4538 Group

In subjects who may experience infusion-related reactions to ONO-4538, prophylactic premedication with acetaminophen or diphenhydramine is recommended before administration of ONO-4538 for subjects who have experienced infusion reaction previously.

5.2.2 Docetaxel Group

Follow recommendations per local SmPC/Package Insert.

5.2.3 Paclitaxel Group

The following pretreatments are recommended to prevent the occurrence of serious hypersensitivity symptoms caused by paclitaxel.

The following premedication should be completed approximately ≥30 minutes prior to administration of paclitaxel: intravenous doses of dexamethasone sodium phosphate injection (dexamethasone 8 mg), either ranitidine hydrochloride injection (ranitidine 50 mg) or famotidine injection (famotidine 20 mg), and oral diphenhydramine hydrochloride tablet (diphenhydramine hydrochloride 50 mg).

The initial dose of dexamethasone is 8 mg. If hypersensitivity symptoms do not occur until the next administration or are not clinically significant, the dexamethasone dose may be reduced to half the initial dose (4 mg) from Week 2. If hypersensitivity symptoms are absent or clinically insignificant in

the subsequent weeks of administration, the dexamethasone dose may be further reduced by half, down to 1 mg.

6 Study Schedule and Observation Items

6.1 Screening phase

The screening phase will start when the subject has signed the informed consent form. Subjects who have provided consent will be enrolled in the study. Among the enrolled subjects, the investigator or subinvestigator will proceed with randomization of those fulfilling the criteria in Section 4.2 "Inclusion Criteria" and not meeting any of the criteria in Section 4.3 "Exclusion Criteria" judged to be appropriate as subjects in this study. Protocol-specified evaluations and their schedule during this period are specified in Table 6-1, Table 6-2, and Table 6-3 (ONO-4538, docetaxel, and paclitaxel groups, respectively) in the protocol synopsis section.

6.2 Treatment phase

6.2.1 ONO-4538 Group

During the treatment phase, ONO-4538 will be administered every 2 weeks up to 3 doses. These 6 weeks will count as one cycle. Study treatment may be repeated in subjects fulfilling the criteria in Section 7.1.2.1 "Criteria for Administration of ONO-4538" and not meeting any of the criteria in Section 7.1.4.1 "Criteria for Discontinuation of ONO-4538." Protocol-specified evaluations and their schedule during this period are specified in Table 6-1 in the protocol synopsis section.

6.2.2 Docetaxel Group

During the treatment phase, docetaxel will be administered every 3 weeks. This is defined as 1 cycle. Study treatment may be repeated in subjects fulfilling the criteria in Section 7.1.2.2 "Criteria for Administration of Docetaxel" and not meeting any of the criteria in Section 7.1.4.2 "Discontinuation Criteria for Docetaxel or Paclitaxel." Protocol-specified evaluations and their schedule during this period are specified in Table 6-2 in the protocol synopsis section.

6.2.3 Paclitaxel Group

Paclitaxel will be administered once weekly for 6 consecutive weeks, followed by a 2-week drug holiday. This is defined as 1 cycle. Study treatment may be repeated in subjects fulfilling the criteria in Section 7.1.2.3 "Criteria for Administration of Paclitaxel" and not meeting any of the criteria in Section 7.1.4.2 "Discontinuation Criteria for Docetaxel or Paclitaxel." Protocol-specified evaluations and their schedule during this period are specified in Table 6-3 in the protocol synopsis section.

6.3 Follow-up phase

All subjects meeting any of the criteria in Section 7.1.4 "Discontinuation Criteria" after administration of the investigational product will undergo end-of-treatment examination (at completion/discontinuation) and proceed to the follow-up phase. Protocol-specified evaluations and their schedule during the follow-up phase are specified in Table 6-1, Table 6-2, and Table 6-3 (ONO-4538, docetaxel, and paclitaxel groups, respectively) in the protocol synopsis section.

6.4 Study Schedule

Table 6-1 Study Schedule for the ONO-4538 Group

Item	Screening phase[1]	Treatment phase										Post-treatment observation period	
		Cycle 1						Cycle ≥2			End (or discontinuation) of the treatment phase[2]	Examination 28 days after the end of treatment phase[2, 3]	Follow-up investigation
Study day		1		8	15	29	43	1	15, 29	43			
		Predose	Postdose										
Allowable window (day/days)	-7 to -1	±0		-3 to +3	-3 to +7	-6 to +7	-6 to +7	±0	-6 to +7	-6 to +7	±3	±7	—
Written informed consent	○[4]												
Demographic data, Eligibility check	○												
Administration of the investigational product[5]		○			○	○		○[6]	○				
Viral tests	○												
Pregnancy testing[7]	○	○						○[8]			○	○	
Performance Status	○			○	○[8]	○[8]	○	○[6, 8]	○[8]	○	○	○	
Vital signs and body weight measurement	○	○[2]	○[9]	○[9]	○[8, 9]	○[8, 9]	○	○[6, 8]	○[8, 9]	○	○	○	
Chest X-ray[10]	○						○			○	○	○	
12-lead ECG	○		○				○	○[8, 11]		○[11]	○	○	
Hematology, biochemistry, and urinalysis	○			○	○[8]	○[8]	○	○[6, 8]		○	○	○	
Immunological and hormone tests[12]	○						○			○	○	○	
Serum drug concentration		○				○[8]		○[8, 13]				○[14]	●[15]
Anti-ONO-4538 antibody		○				○[8]		○[8, 13]				○[14]	●[15]
[redacted]	[redacted]						[redacted]				[redacted]		
[redacted]	[redacted]						[redacted]				[redacted]		
[redacted]	[redacted]												
Tumor tissue examination (PD-L1)	○[18]										●[19]		
[redacted]	[redacted]												
Imaging (e.g., CT, MRI)[20]	○[21]	○[22]									○	○[23]	○[23]
Tumor markers[24]	As needed												
Concomitant treatment and AE monitoring			←								→		○[23, 25]
Outcome investigation													○[26]
Patient Reported Outcomes / Healthcare Resource Utilization	○						○			○[27]	○		○[27]

○ indicates mandatory items; ● indicates optional items.

1. For viral tests, the latest available results obtained within 1 year before randomization may be used as an alternative. For immunological tests and hormone tests, the latest available results obtained within 14 days before randomization may be used as an alternative. For imaging examinations (except chest X-ray), the latest available images obtained within 28 days before randomization may be used as an alternative. Tumor tissue must be obtained for analysis of PD-L1 expression. For the tumor tissue examination after informed consent, no allowable window is set for the timing of tumor tissue specimen collection and assessment before the start of study treatment, but a specimen collected after the end of prior treatment should be submitted as far as possible. [redacted]
2. The previous test results may be used as an alternative if the previous assessment has been done at the end of the treatment phase (at discontinuation) or within the allowable window of 28 days after the end of the treatment phase, unless new measurements are medically required. However, the data will be newly obtained if ≥2 days have passed since the previous vital signs measurement, ≥15 days have passed for imaging examinations (except chest X-ray), or ≥8 days have passed for other tests. Tests will be performed whenever medically indicated.
3. If post-study treatment for esophageal cancer is started due to clinical necessity until 28 days after the end of the treatment phase, an examination 28 days after the end of the treatment phase will be conducted prior to the initiation of the post-study treatment.
4. Administration of the investigational product should be started within 60 days after informed consent.
5. There should be an interval of at least 10 whole days between 2 doses of the investigational product. Considering the day following a dose as the first day, the next dose may be administered on the 11th day or later.
6. When continuation to the next cycle is decided based on examinations performed on Day 43 of the previous cycle, the next cycle should be started within the allowable window for the Day 43 time point of the previous cycle. The available data on Performance Status, vital signs, body weight, hematology tests, biochemistry tests, and urinalysis obtained on Day 43 of the previous cycle may be used as an alternative, unless new measurements are medically required. However, data will be newly obtained if ≥2 days have passed for vital signs or ≥8 days have passed for other tests. Tests will be performed whenever medically indicated.
7. Women of childbearing potential will undergo serum or urine pregnancy testing, and the same testing method should be used throughout the study as far as possible. From Cycle 2 onwards, pregnancy testing will be performed within 7 days before administration of the investigational product.
8. To be assessed before administration of the investigational product.
9. No body weight measurement will be required.
10. As necessary, unscheduled X-rays should be performed whenever an onset of respiratory disease is suspected based on clinical symptoms or findings or laboratory findings during the study period (i.e., after informed consent and until completion of the final examination stipulated by the protocol [with the exception of the outcome investigation]).
11. To be performed only in Cycle 4.
12. SP-D and KL-6 measurements at time points of evaluation as specifed in the screening, treatment and follow-up phases will be mandatory, but SP-D and KL-6 may be measured as necessary at other time points than these time points of evaluation in three phases.
13. To be performed on Day 1 of Cycles 4 and 9.
14. To be performed only in subjects proceeding to the follow-up phase by the end of Cycle 9.
15. To be also measured at 6-12 weeks after the last dose of the investigational product as far as possible.
16. [redacted]
17. This is applicable only to subjects discontinuing the study treatment before Cycle 2.
18. For determination of PD-L1 expression before randomization, a sample test should be performed by the central laboratory. Subjects with stored tumor tissue specimens collected prior to the start of the study treatment can provide the stored specimens to the sponsor. Results of the testing must be available prior to randomization and will be used for stratification through the IWRS. Results will not be made available to the investigator/site staff or the subject.
19. This tumor tissue specimen will be collected after completion of efficacy assessments for the subject.
20. Tumor response will be assessed based on a cervico-thoraco-abdomino-pelvic CT or other imaging examinations. The same method should be used for the assessment of tumor response throughout the study. The RECIST Guideline Ver. 1.1 will be used to assess tumor response. Depending on clinical symptoms, the brain or bone metastasis status should be assessed based on diagnostic imaging data.
21. If a brain or bone metastasis is clinically suspected before enrollment, a head CT/MRI or FDG-PET (or bone scintigraphy) must be performed to determine the presence or absence of brain or bone metastasis (within 28 days before randomization).
22. Imaging examinations will be performed every 6 weeks (-6 to +7 days) from the start of Cycle 1 during the treatment phase. After a period of one year from the start of Cycle 1 (after the end of Cycle 9), imaging examinations will be performed every 12 weeks (2 cycles). Depending on clinical symptoms, the brain or bone metastasis status should be assessed based on diagnostic imaging data.
23. For subjects who have an response assessment other than PD in accordance with the RECIST Guideline Ver. 1.1 (see Appendix 2) and have discontinued the treatment phase for safety reasons, imaging examinations will be continued as far as possible until either initiation of post-study treatment for esophageal cancer or assessment of PD or recurrence while continuing the follow-up investigations to collect data on any post-study treatment for esophageal cancer and the details of the treatment.
24. In subjects with tumor marker levels above the ULN range, measurements of the tumor marker will be continued as far as possible and as frequently as needed.
25. Monitoring for any AEs and concomitant treatments will be continued until 28 days after the end of the treatment phase. In subjects who, at the time of initiation of the follow-up investigation, have any AE for which a causal relationship to the investigational product cannot be ruled out or which led to discontinuation, investigation of the AE and treatment for the event will be continued at appropriate intervals until the event resolves, improves, or stabilizes and thus is judged not to require further follow-up. SAEs, IMAEs and any treatment for IMAEs are to be collected during the treatment period and for 100 days following the last dose of investigational product.
26. Outcome investigation (with collection of data on the date of death and cause of death if the patient died) should be conducted by direct contact or other means such as by telephone or letter, roughly every 6 to 8 weeks but as frequently as required based on the occurrence of events. Updated information on survival status will be entered into the eCRF. The follow-up investigation will also collect data on any post-study treatment for esophageal cancer, start date, and details as far as possible.
27. The Patient Reported Outcomes and Healthcare Resource Utilization will be administered every 6 weeks (-6 to +7 days) from the start of Cycle 1 until the end of the treatment phase and subsequently every 12 weeks (-14 to +14 days) during the follow-up phase.

Table 6-2 Study Schedule for the Docetaxel Group

Item	Screening phase[1]	Treatment phase						End (or discontinuation) of the treatment phase[2]	Post-treatment observation period	
		Cycle 1				Cycle ≥2			Examination 28 days after the end of treatment phase[2, 3]	Follow-up investigation
Study day		1		8	22	1	22			
		Pre dose	Post dose							
Allowable window (day/days)	-7 to -1	±0		-3 to +3	-6 to +7	±0	-6 to +7	±3	±7	—
Written informed consent	○[4]									
Demographic data, Eligibility check	○									
Administration of the investigational product		○				○[5]				
Serological tests	○									
Pregnancy testing[6]	○	○				○[7]		○	○	
Performance Status	○			○	○	○[5, 7]	○	○	○	
Vital signs and body weight	○	○[2]	○[8]	○[8]	○	○[5, 7]	○	○	○	
Chest X-ray [9]	○				○		○	○	○	
12-lead ECG	○		○		○	○[7, 10]	○[10]	○	○	
Hematology, biochemistry, and urinalysis	○			○	○	○[5, 7]	○	○	○	
Immunological and hormone tests[11]	○				○		○	○	○	
[redacted]										
[redacted]										
[redacted]										
Tumor tissue examination (PD-L1)	○[15]							●[16]		
[redacted]										
Imaging (e.g., CT, MRI)[17]	○[18]	○[19]						○	○[20]	○[20]
Tumor markers[21]	As needed									
Concomitant treatment and AE monitoring			←					→		○[20, 22]
Outcome investigation										○[23]
Patient Reported Outcomes / Healthcare Resource Utilization	○	○[24]						○		○[24]

○ indicates mandatory items; ● indicates optional items.

1. For serological tests, the latest available results obtained within 1 year before randomization may be used as an alternative. For immunological tests and hormone tests, the latest available results obtained within 14 days before randomization may be used as an alternative. For imaging examinations (except chest X-ray), the latest available images obtained within 28 days before randomization may be used as an alternative. Tumor tissue must be obtained for analysis of PD-L1 expression. For the tumor tissue examination after informed consent, no allowable window is set for the timing of tumor tissue specimen collection and assessment before the start of study treatment, but a specimen collected after the end of prior treatment should be submitted as far as possible. [redacted]
2. The previous test results may be used as an alternative if the previous assessment has been done at the end of the treatment phase (at discontinuation) or within the allowable window of 28 days after the end of the treatment phase, unless new measurements are medically required. However, the data will be newly obtained if ≥2 days have passed since the previous vital signs measurement, ≥15 days have passed for imaging examinations (except chest X-ray), or ≥8 days have passed for other tests. Tests will be performed whenever medically indicated.
3. If post-study treatment for esophageal cancer is started due to clinical necessity until 28 days after the end of the treatment phase, an examination 28 days after the end of the treatment phase will be conducted prior to the initiation of the post-study treatment.
4. Administration of the investigational product should be started within 60 days after informed consent.
5. When continuation to the next cycle is decided based on examinations performed on Day 22 of the previous cycle, the next cycle should be started within the allowable window for the Day 22 time point of the previous cycle. The available data on Performance Status, vital signs, body weight, hematology tests, biochemistry tests, and urinalysis obtained on Day 22 of the previous cycle may be used as an alternative, unless new measurements are medically required. However, data will be newly obtained if ≥2 days have passed for vital signs or ≥8 days have passed for other tests. Tests will be performed whenever medically indicated.
6. Women of childbearing potential will undergo serum or urine pregnancy testing, and the same testing method should be used throughout the study as far as possible. From Cycle 2 onwards, pregnancy testing will be performed within 7 days before administration of the investigational product.
7. To be assessed before administration of the investigational product.
8. No body weight measurement will be required.
9. As necessary, unscheduled X-rays should be performed whenever an onset of respiratory disease is suspected based on clinical symptoms or findings or laboratory findings during the study period (i.e., after informed consent and until completion of the final examination stipulated by the protocol [with the exception of the outcome investigation]).
10. To be performed only in Cycle 7.
11. SP-D and KL-6 measurements at time points of evaluation as specifed in the screening, treatment and follow-up phases will be mandatory , but SP-D and KL-6 may be measured as necessary at other time points than these time points of evaluation in three phases.
12. [redacted]
13. To be performed only in Cycle 2.
14. To be performed only when the study treatment is discontinued by Day 22 of Cycle 2.
15. For determination of PD-L1 expression before randomization, a sample test should be performed by the central laboratory. Subjects with stored tumor tissue specimens collected prior to the start of the study treatment can provide the stored specimens to the sponsor. Results of the testing must be available prior to randomization and will be used for stratification through the IWRS. Results will not be made available to the investigator/site staff or the subject.
16. This tumor tissue specimen will be collected after completion of efficacy assessments for the subject.
17. Tumor response will be assessed based on a cervico-thoraco-abdomino-pelvic CT or other imaging examinations. The same method should be used for the assessment of tumor response throughout the study. The RECIST Guideline Ver. 1.1 will be used to assess tumor response.
18. If a brain or bone metastasis is clinically suspected before enrollment, a head CT/MRI or FDG-PET (or bone scintigraphy) must be performed to determine the presence or absence of brain or bone metastasis (within 28 days before randomization).
19. Imaging examinations will be performed every 6 weeks (-6 to +7 days) from the start of Cycle 1 during the treatment phase and every 12 weeks (-14 to +14 days) after a period of one year from the start of Cycle 1. Depending on clinical symptoms, the brain or bone metastasis status should be assessed based on diagnostic imaging data.
20. For subjects who have an response assessment other than PD in accordance with the RECIST Guideline Ver. 1.1 (see Appendix 2) and have discontinued the treatment phase for safety reasons, imaging examinations will be continued as far as possible until either initiation of post-study treatment for esophageal cancer or assessment of PD or recurrence while continuing the follow-up investigations to collect data on any post-study treatment for esophageal cancer and the details of the treatment.
21. In subjects with tumor marker levels above the ULN range, measurements of the tumor marker will be continued as far as possible and as frequently as needed.
22. Monitoring for any AEs and concomitant treatments will be continued until 28 days after the end of the treatment phase. In subjects who, at the time of initiation of the follow-up investigation, have any AE for which a causal relationship to the investigational product cannot be ruled out or which led to discontinuation, investigation of the AE and treatment for the event will be continued at appropriate intervals until the event resolves, improves, or stabilizes and thus is judged not to require further follow-up. SAEs, IMAEs and any treatment for IMAEs are to be collected during the treatment period and for 100 days following the last dose of investigational product.
23. Outcome investigation (with collection of data on the date of death and cause of death if the patient died) should be conducted by direct contact or other means such as by telephone or letter, roughly every 6 to 8 weeks but as frequently as required based on the occurrence of events. Updated information on survival status will be entered into the eCRF. The follow-up investigation will also collect data on any post-study treatment for esophageal cancer, start date, and details as far as possible.
24. The Patient Reported Outcomes and Healthcare Resource Utilization will be administered every 6 weeks (-6 to +7 days) from the start of Cycle 1 until the end of the treatment phase and subsequently every 12 weeks (-14 to +14 days) during the follow-up phase.

Table 6-3 Study Schedule for the Paclitaxel Group

Item	Screening phase[1]	Treatment phase										Follow-up phase	
		Cycle 1						Cycle ≥2			End (or discontinuation) of the treatment phase[2]	Examination 28 days after the end of treatment phase[2, 3]	Follow-up investigation
Study day		1		8	15, 22, 29, 36	43	50	1	8, 15, 22, 29, 36	50			
		Pre dose	Post dose										
Allowable window (day/days)	-7 to -1	±0		-1 to +3	-2 to +3	-2 to +3	-6 to +7	±0	-2 to +3	-6 to +7	±3	±7	—
Written informed consent	○[4]												
Demographic data, Eligibility check	○												
Administration of the investigational product[5]		○		○	○			○[6]	○				
Serological tests	○												
Pregnancy testing[7]	○	○						○[8]			○	○	
Performance Status	○			○[8]			○	○[6, 8]		○	○	○	
Vital signs and body weight	○	○[2]	○[9]	○[8, 9]	○[8, 9, 10]		○	○[6, 8]	○[8, 9, 10]	○	○	○	
Chest X-ray[11]	○						○			○	○	○	
12-lead ECG	○		○				○	○[8, 12]		○[12]	○	○	
Hematology, biochemistry, and urinalysis	○			○[8, 10]	○[8, 10]		○	○[6, 8]	○[8, 10]	○	○	○	
Immunological and hormone tests[13]	○						○			○	○	○	
[redacted]													
[redacted]													
[redacted]													
Tumor tissue examination (PD-L1)	○[16]										●[17]		
[redacted]													
Imaging (e.g., CT, MRI)[18]	○[19]	○[20]									○	○[21]	○[21]
Tumor markers[22]		As needed											
Concomitant treatment and AE monitoring		←→											○[21, 23]
Outcome investigation													○[24]
Patient Reported Outcomes / and Healthcare Resource Utilization	○	○[25]									○		○[25]

○ indicates mandatory items; ● indicates optional items.

1. For serological tests, the latest available results obtained within 1 year before randomization may be used as an alternative. For immunological tests and hormone tests, the latest available results obtained within 14 days before randomization may be used as an alternative. For imaging examinations (except chest X-ray), the latest available images obtained within 28 days before randomization may be used as an alternative. Tumor tissue must be obtained for analysis of PD-L1 expression. For the tumor tissue examination after informed consent, no allowable window is set for the timing of tumor tissue specimen collection and assessment before the start of study treatment, but a specimen collected after the end of prior treatment should be submitted as far as possible. [redacted]
2. The previous test results may be used as an alternative if the previous assessment has been done at the end of the treatment phase (at discontinuation) or within the allowable window of 28 days after the end of the treatment phase, unless new measurements are medically required. However, the data will be newly obtained if ≥2 days have passed since the previous vital signs measurement, ≥15 days have passed for imaging examinations (except chest X-ray), or ≥8 days have passed for other tests. Tests will be performed whenever medically indicated.
3. If post-study treatment for esophageal cancer is started due to clinical necessity until 28 days after the end of the treatment phase, an examination 28 days after the end of the treatment phase will be conducted prior to the initiation of the post-study treatment.
4. Administration of the investigational product should be started within 60 days after informed consent.
5. There should be an interval of at least 5 days between the doses.
6. When continuation to the next cycle is decided based on examinations performed on Day 50 of the previous cycle, the next cycle should be started within the allowable window for the Day 50 time point of the previous cycle. The available data on Performance Status, vital signs, body weight, hematology tests, biochemistry tests, and urinalysis obtained on Day 50 of the previous cycle may be used as an alternative, unless new measurements are medically required. However, data will be newly obtained if ≥2 days have passed for vital signs or ≥8 days have passed for other tests. Tests will be performed whenever medically indicated.
7. Women of childbearing potential will undergo serum or urine pregnancy testing, and the same testing method should be used throughout the study as far as possible. From Cycle 2 onwards, pregnancy testing will be performed within 7 days before administration of the investigational product.
8. To be assessed before administration of the investigational product.
9. No body weight measurement will be required.
10. Check whether the dosing criteria and dose reduction criteria are met and whether the withdrawal criteria are fulfilled by performing vital signs measurements (systolic blood pressure, diastolic blood pressure, and body temperature), hematological tests (hemoglobin, white blood count, neutrophil count, and blood platelet count), and biochemical tests (albumin, AST, ALT, ALP, total bilirubin, Na, K, Ca, creatinine, and CRP) before treatment with the investigational product.
11. As necessary, unscheduled X-rays should be performed whenever an onset of respiratory disease is suspected based on clinical symptoms or findings or laboratory findings during the study period (i.e., after informed consent and until completion of the final examination stipulated by the protocol [with the exception of the outcome investigation]).
12. To be performed only in Cycle 3.
13. SP-D and KL-6 measurements at time points of evaluation as specifed in the screening, treatment and follow-up phases will be mandatory, but SP-D and KL-6 may be measured as necessary at other time points than these time points of evaluation in three phases.
14. [redacted]
15. To be performed only when the study treatment is discontinued by Day 43 of Cycle 1.
16. For determination of PD-L1 expression before randomization, a sample test should be performed by the central laboratory. Subjects with stored tumor tissue specimens collected prior to the start of the study treatment can provide the stored specimens to the sponsor. Results of the testing must be available prior to randomization and will be used for stratification through the IWRS. Results will not be made available to the investigator/site staff or the subject.
17. This tumor tissue specimen will be collected after completion of efficacy assessments for the subject.
18. Tumor response will be assessed based on a cervico-thoraco-abdomino-pelvic CT or other imaging examinations. The same method should be used for the assessment of tumor response throughout the study. The RECIST Guideline Ver. 1.1 will be used to assess tumor response. Depending on clinical symptoms, the brain or bone metastasis status should be assessed based on diagnostic imaging data.
19. Imaging examinations will be performed every 6 weeks (-6 to +7 days) from the start of Cycle 1 during the treatment phase and every 12 weeks (-14 to +14 days) after a period of one year from the start of Cycle 1. Depending on clinical symptoms, the brain or bone metastasis status should be assessed based on diagnostic imaging data.
20. After a lapse of one year from the start of Cycle 1 (after the end of Cycle 9), imaging examinations will be performed every 12 weeks (2 cycles).
21. For subjects who have an response assessment other than PD in accordance with the RECIST Guideline Ver. 1.1 (see Appendix 2) and have discontinued the treatment phase for safety reasons, imaging examinations will be continued as far as possible until either initiation of post-study treatment for esophageal cancer or assessment of PD or recurrence while continuing the follow-up investigations to collect data on any post-study treatment for esophageal cancer and the details of the treatment.
22. In subjects with tumor marker levels above the ULN range, measurements of the tumor marker will be continued as far as possible and as frequently as needed.
23. Monitoring for any AEs and concomitant treatments will be continued until 28 days after the end of the treatment phase. In subjects who, at the time of initiation of the follow-up investigation, have any AE for which a causal relationship to the investigational product cannot be ruled out or which led to discontinuation, investigation of the AE and treatment for the event will be continued at appropriate intervals until the event resolves, improves, or stabilizes and thus is judged not to require further follow-up. SAEs, IMAEs and any treatment for IMAEs are to be collected during the treatment period and for 100 days following the last dose of investigational product.
24. Outcome investigation (with collection of data on the date of death and cause of death if the patient died) should be conducted by direct contact or other means such as by telephone or letter, roughly every 6 to 8 weeks but as frequently as required based on the occurrence of events. Updated information on survival status will be entered into the eCRF. The follow-up investigation will also collect data on any post-study treatment for esophageal cancer, start date, and details as far as possible.
25. The Patient Reported Outcomes and Healthcare Resource Utilization will be administered every 6 weeks (-6 to +7 days) from the start of Cycle 1 until the end of the treatment phase and subsequently every 12 weeks (-14 to +14 days) during the follow-up phase.

7 Endpoints

7.1 Efficacy Endpoints - Primary Endpoint

Overall survival (OS)

7.2 Efficacy Endpoints - Secondary Endpoints

1. Objective response rate (ORR)
2. Disease control rate (DCR)
3. Progression free survival (PFS)
4. Duration of response
5. Time to response
6. Best overall response (BOR)
7. Maximum percent change from baseline in the sum of diameters of the target lesion

7.3 Safety Endpoints

1. Adverse events
2. Clinical laboratory tests
3. Vital signs
4. 12-lead ECG
5. Chest X-ray
6. Performance Status (ECOG)

7.4 Pharmacokinetic Endpoint

Serum ONO-4538 concentrations (for the ONO-4538 group only)

7.5 Anti-drug Antibody Endpoint

Anti-ONO-4538 antibody (for the ONO-4538 group only)

7.6 Exploratory Biomarkers

6. Tumor tissue examination (PD-L1 expression analysis, essential; [redacted])

7.7 Other Test Variables

1. Tumor markers
2. Patient reported outcomes (PROs)
3. Healthcare Resource Utilization

8 Planned Sample Size

The planned sample size for this study is 195 for each group, totaling 390.

Table of Contents

1 Background 34
1.1 Introduction 34
1.2 Summary of Important Findings from Non-clinical and Clinical Studies 36
1.2.1 Summary of Non-clinical Data 36
1.2.2 Summary of Clinical Data and Summary of Known and Potential Risks and Benefits to Subjects 38
2 Study Organization and Roles 40
3 Objective of the Study 40
3.1 Primary Objective of the Study 40
3.2 Secondary Objectives of the Study 40
4 Subjects 41
4.1 Subjects 41
4.2 Inclusion Criteria 41
4.3 Exclusion Criteria 44
5 Study Methods 47
5.1 Study Design 47
5.1.1 Study Design 47
5.1.2 Planned Sample Size 48
5.2 Method of Assigning Subjects to Treatment Groups 50
5.2.1 Method for investigational product allocation 50
5.3 Subject Enrollment Procedures 51
5.4 Blinding 52
5.5 Endpoints 52
5.5.1 Efficacy Endpoints 52
5.5.1.1 Primary Endpoint 52
5.5.1.2 Secondary Endpoints 52
5.5.2 Safety Endpoints 52
5.5.3 Pharmacokinetic Endpoint 53
5.5.4 Anti-drug Antibody Endpoint 53
5.5.5 Exploratory Biomarkers 53
5.5.6 Other Test Variables 53
6 Investigational Products 53

6.1 Handling and Supply of Investigational Product 54
6.2 Procedures for Investigational Product Accountability and Return of Unused Investigational products 55
6.2.1 Investigational Product Storage and Management 55
6.2.2 Disposal of Investigational Products 56
6.2.3 Return of Unused Investigational Products 57
7 Study Plan 57
7.1 Investigational Product 57
7.1.1 Dosage, Administration, and Treatment Duration 57
7.1.1.1 Dosage, Administration, and Treatment Duration of ONO-4538 57
7.1.1.2 Dosage, Administration, and Treatment Duration of Docetaxel 59
7.1.1.3 Dosage, Administration, and Treatment Duration of Paclitaxel 59
7.1.2 Criteria for Administration of Investigational Product 60
7.1.2.1 Criteria for Administration of ONO-4538 60
7.1.2.2 Criteria for Administration of Docetaxel 61
7.1.2.3 Criteria for Administration of Paclitaxel 62
7.1.3 Dosage Reduction Criteria 63
7.1.3.1 Dosage Reduction Criteria for ONO-4538 63
7.1.3.2 Dose Reduction Criteria for Docetaxel 63
7.1.3.3 Dose Reduction Criteria for Paclitaxel 64
7.1.4 Discontinuation Criteria 64
7.1.4.1 Criteria for Discontinuation of ONO-4538 64
7.1.4.2 Discontinuation Criteria for Docetaxel or Paclitaxel 66
7.2 Prior and Concomitant Therapies 67
7.2.1 Prohibited Therapies During the Study Period 67
7.2.1.1 Prohibited Therapies During the Study Period (ONO-4538 Group) 67
7.2.1.2 Prohibited Therapies During the Study Period (Control Group) 68
7.2.2 Prophylactic Premedication 69
7.2.2.1 ONO-4538 Group 69
7.2.2.2 Docetaxel Group 69
7.2.2.3 Paclitaxel Group 69
7.3 Discontinuation of the Supply of the Investigational Product 70
7.4 Post-Study Access to Therapy 70

8 Study schedule and Observation Items 70
8.1 Study Schedule 70
8.1.1 Screening Phase 70
8.1.2 Treatment Phase 71
8.1.2.1 ONO-4538 Group 71
8.1.2.2 Docetaxel Group 71
8.1.2.3 Paclitaxel Group 71
8.1.3 Follow-up Phase 71
8.2 Observation Items 72
8.2.1 Patient Demographics and Other Background Characteristics 72
8.2.2 Details of Study Treatment and Concomitant Treatments 73
8.2.2.1 Details of administration of the investigational product 73
8.2.2.2 Details of Administration of Other Drugs and Concomitant Therapies 73
8.2.3 Items related to Efficacy 73
8.2.3.1 Imaging Examination (with Measurements of Tumor Diameters and Assessment of Tumor Response) 74
8.2.4 Items related to Safety 75
8.2.4.1 Vital Signs, Body Weight, and 12-lead ECG 75
8.2.4.2 Chest X-ray 77
8.2.4.3 Laboratory Tests 78
8.2.4.4 Pregnancy Testing (Only in Women of Childbearing Potential) 80
8.2.4.5 Viral Tests 81
8.2.4.6 ECOG Performance Status Score 81
8.2.4.7 Adverse Events 82
8.2.5 Serum Drug Concentration Measurements (ONO-4538 group only) 83
8.2.6 Anti-drug Antibody Measurements (ONO-4538 group only) 84
8.2.7 Tumor Markers (As needed) 84
8.2.8 Patient Reported Outcomes (PROs), Healthcare Resource Utilization 85
8.2.9 Exploratory Biomarkers 85

[redacted]
8.2.9.6 Tumor Tissue Examinations (some are essential) 88
8.2.10 Follow-up Investigation 89
8.2.11 Lost to Follow-up 90
9 Adverse Events and Matters to Ensure the Safety of Subjects 90
9.1 Definition of Adverse Events 90
9.2 Definition of Immune-mediated Adverse Events 91
9.3 Definition of Serious Adverse Events 92
9.4 Grading Criteria of Adverse Events 93
9.5 Assessment Criteria of Causal Relationship to the Investigational Product 93
9.6 Event Outcome 94
9.7 Action Taken with the Investigational Product 94
9.8 Any Treatment for Adverse Events 95
9.9 Actions to be Taken Following Onset of Any Serious Adverse Event 95
9.10 Actions to be Taken Following Pregnancy 96
9.11 Actions to be Taken Following Overdose 96
9.12 Other Safety Considerations 97
9.13 Provision of New Information 97
10 Statistical Analysis 97
10.1 Description of Statistical Analysis Methods including the Interim Analysis 97
10.2 Independent Data Monitoring Committee 97
10.3 Significance Level to be Used 98
10.3.1 Efficacy 98
10.3.2 Safety 98
10.4 Determination of Subjects included in Analysis 98
10.4.1 Definitions of Analysis Sets 98
10.4.1.1 Enrolled Set（ENR） 98
10.4.1.2 Randomized Set（RND）／ITT 98
10.4.1.3 SAF 98
10.4.1.4 ADA 98
10.4.2 Criteria for Handling of Subjects 99
10.4.2.1 Disqualified Subjects 99
10.4.2.2 GCP Noncompiant Subjects 99

10.4.2.3 Untreated Subjects 99
10.4.2.4 Incomlete Anti-Drug Antibody Subjects 99
10.4.3 Criteria for Handling of Time Points of Evaluation 99
10.5 Analysis Methods 100
10.5.1 Reliability of the Study 100
10.5.1.1 Inclusion and Exclusion for Analyses 100
10.5.1.2 Withdrawals and Dropouts 100
10.5.2 Details of Treatment with the Investigational Product 100
10.5.3 Distribution of Patient Background Factors 100
10.5.3.1 Demographic Variables 100
10.5.3.2 Patient Background Characteristics 100
10.5.3.3 Baseline Values of Observation 101
10.5.3.4 Concomitant drugs and therapies 101
10.5.4 Primary Efficacy Endpoint 101
10.5.4.1 Purpose of Analysis 101
10.5.4.2 Hypothesis 101
10.5.4.3 Analytical Item and Data Handling 101
10.5.4.4 Analytical Methods 101
10.5.5 Secondary Efficacy Endpoints 102
10.5.5.1 Purpose of Analysis 102
10.5.5.2 Analytical Items and Data Handling 102
10.5.5.3 Analytical Methods, ... 104
10.5.6 Interaction Analysis 105
10.5.7 Adjusted Analysis 105
10.5.8 Analysis of safety 105
10.5.8.1 Purpose of Analysis 105
10.5.8.2 Analytical methods 105
10.5.9 Data Reviews 106
10.5.10 Other Issues 106
11 Ethical Conduct of the Study 106
11.1 Good Clinical Practice 106
11.2 Institutional Review Board or Independent Ethics Committee (IRB/IEC) 107
11.3 Privacy Protection for Subjects 107

12 Subject Informed Consent 107
12.1 Informed Consent 107
12.2 Withdrawal of Consent 109
13 Rules or Criteria for Withdrawing Individual Subjects from the Study or Discontinuing Parts of the Study or Entire Study 109
13.1 Rules or Criteria for Withdrawing Individual Subjects from the Study 109
14 Study Management 110
14.1 Compliance with the Protocol and Protocol Amendments 110
15 Completion of Electronic Case Report Forms (eCRFs) 110
15.1 eCRF 110
16 Monitoring 111
17 Procedures to Ensure Compliance of Subject Medication and Other Stipulations 111
18 Retention of Essential Documents 111
18.1 Record Retention 111
19 Publication of Study Results 112
20 Literature references 113

List of Tables

Table 6-1 Study Schedule for the ONO-4538 Group 16
Table 6-2 Study Schedule for the Docetaxel Group 18
Table 6-3 Study Schedule for the Paclitaxel Group 20
Table 1 List of abbreviations 31
Table 6-1 Description of Investigational Products 54
Table 7.1.1.1-1 Exposure of ONO-4538 at steady state 58
Table 7.1.3.2-1 Dose Criteria for Docetaxel 63
Table 7.1.3.2-2 Dose Reduction Criteria for Docetaxel 63
Table 7.1.3.3-1 Dose Criteria for Paclitaxel 64
Table 7.1.3.3-2 Dose Reduction Criteria for Paclitaxel 64
Table 9.2-1 Preferred Terms of IMAEs 91
* Irrespective of being treated with immunosuppressive medications or not. 91

List of Figures

Figure 2-1 Outline of Study Design 5

Appendix 1: Criteria and Definitions Used in This Study

Appendix 2: Tumor Response Assessment Criteria Used in This Study

Appendix 3: Performance Status Score (ECOG)

Appendix 4: List of Autoimmune Diseases

Table 1 List of abbreviations

Abbreviation	Unabbreviated term
ACTH	Adrenocorticotropic hormone
ADCC	Antibody-dependent cellular toxicity
ADL	Activities of daily living
AE	Adverse event
ALP	Alkaline phosphatase
ALT (GPT)	Alanine aminotransferase (glutamic pyruvic transaminase)
ANA	Antinuclear antibody
AST (GOT)	Aspartate aminotransferase (glutamic oxaloacetic transaminase)
AUC	Area under the plasma concentration-time curve
BMI	Body mass index
BOR	Best overall response
BUN	Blood urea nitrogen
CD	Cluster of differentiation
CDC	Complement-dependent cytotoxicity
C3	Beta-1C/beta-1A globulin
C4	Beta-1E globulin
CHO	Chinese hamster ovary
CH50	50% hemolytic complement
CK(CPK)	Creatine kinase (creatine phosphokinase)
Cmax	Maximum plasma concentration
CR	Complete response
CRP	C-reactive protein
CT	Computed tomography
CTCAE	Common Terminology Criteria for Adverse Events
CTLA-4	Cytotoxic T lymphocyte-associated antigen 4
DCR	Disease control rate
DNA	Deoxyribonucleic acid
ECOG	Eastern Cooperative Oncology Group
eCRF	Electronic case report form
EDC	Electronic data capture
EGFR	Epidermal growth factor receptor
EMR	Endoscopic mucosal resection
ENR	Enrolled Set
ESD	Endoscopic submucosal dissection
FDG-PET	^{18}F-fluorodeoxy glucose positron emission tomography
free T3	Free triiodothyronine
free T4	Free thyroxine
GAD	Glutamic acid decarboxylase
GCP	Good Clinical Practice
G-CSF	Granulocyte colony-stimulating factor
HBc	Hepatitis B core
HBs	Hepatitis B surface
HBV	Hepatitis B virus
HCV	Hepatitis C virus
HIV-1	Human immunodeficiency virus-1
HIV-2	Human immunodeficiency virus-2

Abbreviation	Unabbreviated term
[redacted]	[redacted]
HR	Heart rate
HTLV-1	Human T-lymphotropic virus-1
IDMC	Independent Data Monitoring Committee
IEC	Independent Ethics Committee
IFN-γ	Interferon-gamma
IgA	Immunoglobulin A
IRB	Institutional Review Board
ITT	Intention-to-treat
IWRS	Interactive Web Response System
JCOG	Japan Clinical Oncology Group
K_D	Dissociation constant
KL-6	Krebs von den Lungen-6
La	Anti-SSB antibody
LDH	Lactate dehydrogenase
LKM	Liver kidney microsome
MCH	Mean corpuscular hemoglobin
MCHC	Mean corpuscular hemoglobin concentration
MCV	Mean corpuscular volume
MLR	Human mixed lymphocyte reaction
MRI	Magnetic resonance imaging
NCI	National Cancer Institute
NE	Not evaluable
NYHA	New York Heart Association
OS	Overall survival
ORR	Objective response rate
PBMC	Peripheral blood mononuclear cell
PD	Progressive disease
PD-1	Programmed cell death-1
PD-L1	Programmed cell death-ligand 1
PD-L2	Programmed cell death-ligand 2
PS	Performance status
PFS	Progression-free survival
PK	Pharmacokinetics
PK/PD	Pharmacokinetics/pharmacodynamics
PR	Partial response
PR interval	PR interval on an electrocardiogram (atrioventricular conduction time)
PT	Preferred term
QRS width	QRS width on an electrocardiogram (ventricular activation time)
QT interval	QT interval on an electrocardiogram (duration of ventricular electrical systole)
RA	Rheumatoid factor
RANKL	Receptor activator of NF-κB ligand
RECIST	Response Evaluation Criteria in Solid Tumors

Abbreviation	Unabbreviated term
RNA	Ribonucleic acid
RND	Randomized set
Ro	Anti-SSA antibody
RR	Ventricular rate
SAE	Serious adverse event
SAF	Safety set
SD	Stable disease
SOC	System organ class
SP-D	Pulmonary surfactant protein D
SSA	Sjogren syndrome A
SSB	Sjogren syndrome B
TNM	Tumor node metastasis
TSH	Thyroid-stimulating hormone
UICC	Unio Internationalis Contra Cancrum
γ-GTP	Gamma-glutamyl transpeptidase

1 Background

1.1 Introduction

The global number of deaths due to malignant tumors was estimated to have exceeded 8,200,000 in 2012, of which the number of deaths due to esophageal cancer was 400,000, ranking 6th after lung, liver, gastric, colorectal, and breast cancers [1)]. The number of deaths due to malignant tumors in Japan, which ranks first among the causes of deaths in the country, was 364,872 (2013). Among these, the number of deaths due to esophageal cancer was 11,543 (2013), accounting for 3.2% of the total number of deaths due to malignant tumors [2)]. Esophageal cancer has a higher incidence in men than women at a ratio of about 6:1, and follows lung, gastric, colorectal, liver, and pancreatic cancers in terms of crude death rates in men. The number of patients with esophageal cancer was 23,119 in 2011 [3)], with high rates in patients in their 60s to 70s of age, accounting for approximately 68% of the overall incidence [4)]. The numbers of deaths and patients are expected to increase in both sexes, up to 13,100 and 24,900, respectively, from 2025-2029 [5)]. Moreover, epidemiological information in Korea and Taiwan reveals that the estimated number of patients with esophageal cancer in Korea was 8314 (2013) [6)], and the annual number of newly-diagnosed patients in Taiwan was 2238 (2010) [7)]. These numbers are reportedly increasing.

There are regional differences in the histological type and location of lesions of esophageal cancer. In Japan, the dominant histological type is squamous cell carcinoma, accounting for 88.7% of all esophageal cancers, and locations of lesions are most commonly found in the middle thoracic esophagus at a rate of 46.9% [4)]. Squamous cell carcinoma is also reportedly common in Asia including Taiwan. In contrast, adenocarcinomas are increasing in Western countries, which account for more than half of all histological types. The common locations of lesions include those of the lower chest, abdominal esophagus, and esophagogastric junction. Prognosis of squamous cell carcinoma is considered to be poorer than that of adenocarcinoma. The main risk factors of squamous cell carcinoma are drinking and smoking, and the risk is known to increase by a combination of both [8)-12)]. On the other hand, the main risk factors of adenocarcinoma are obesity and reflux esophagitis. It has been noted that adenocarcinomas may also be increasing in Japan due to its westernized lifestyle, etc [13)].

To diagnose and treat esophageal cancer, various imaging tests on the depth of tumor invasion, lymph node metastasis, and distal metastasis are performed to determine disease progression (stage and type), and multimodal treatments with a combination of endoscopic treatment, surgical treatment, chemotherapy, and radiotherapy are provided. Treatment in patients without distal metastases is aimed at a radical cure by methods appropriate for their stage of cancer. Meanwhile, for patients with distal

metastases or those with a postoperative distant recurrence, combination therapy with cisplatin and 5-FU (FP therapy) is recommended as the standard treatment, but it is not proven to prolong survival [15]. Survival prolongation by docetaxel or paclitaxel commonly used as second-line treatment has been confirmed in esophageal adenocarcinoma but not in esophageal squamous cell carcinoma in comparative studies[13) 16)-18) 27)]. The prognosis of patients treated solely by chemotherapy is poor with a 5-year survival rate of 8.6% [4]. Comparing the 5-year relative survival rate in patients with distal metastases by cancer type, the survival rate of all cancer types is 11.8%, while that of esophageal cancer is as low as 4.3% followed by pancreatic, biliary/bile duct, hepatic/intrahepatic bile duct, and lung cancers [14]. A triple-agent combination therapy with the addition of docetaxel to FP therapy has been recently investigated; however, chemotherapies including combination chemotherapy have their limits in terms of efficacy, so chemotherapy alone is only applied to patients with unresectable metastases [13]. There are no other standard therapies for esophageal cancer if a patient becomes refractory or intolerant to standard therapies including FP therapy. Hence, a new agent for esophageal cancer is needed.

ONO-4538, a human monoclonal antibody to the PD-1 (Programmed cell death-1, also known as CD279), was developed by Ono Pharmaceutical Co., Ltd. (Ono Pharmaceutical) and Medarex, Inc. (currently, Bristol-Myers Squibb Company [BMS]), and is now in the phase of clinical development by Ono Pharmaceutical and BMS. The PD-1 (or CD279) is a 55 kDa type 1 transmembrane protein and belongs to the CD28 immunoglobulin superfamily of T-cell co-stimulatory receptors including CD28, CTLA-4, ICOS, and BTLA. PD-1 is highly expressed on activated T-cells and B cells and is also expressed in varying degrees on memory T cell subsets. Two specific ligands for PD-1 have been identified: PD-L1 (also called B7-H1 or CD274) and PD-L2 (also called B7-DC or CD273). The binding of PD-1 with either of its ligands PD-L1 or PD-L2 has been suggested to negatively control the activation of T-cells in mice and humans [19)-21)]. ONO-4538 is approved for malignant melanoma and other indications in Japan, Korea, the US, Europe and so on, and also under development for various cancer types.

PD-L1 and PD-L2 are expressed in tumor tissues in 43.9% of patients with esophageal cancer. Patients with PD-L1 and PD-L2 expression are reported to show poorer prognoses than those without expression [22]. Based on phase I studies conducted in Japan and the US (ONO-4538-01 and CA209003), ONO-4538 is suggested to be effective for a wide range of tumor histological types, including squamous cell carcinomas and adenocarcinomas, the major histological types of esophageal cancer. A Japanese phase II study of ONO-4538 in patients with esophageal cancer refractory or intolerant to standard therapies (ONO-4538-07) is currently ongoing. This study has suggested that the

antitumor effect of ONO-4538 in esophageal cancer is sustainable for a long duration as in other types of cancer.

Based on the above, ONO-4538 is suggested to demonstrate a long-term, persistent tumor response to esophageal cancer refractory or intolerant to standard therapies, and thus is expected to prolong survival. Therefore, we believe ONO-4538, that inhibits the binding of PD-1 with its ligand, can be a therapeutic agent for esophageal cancer with a novel mode of action, and planned this study to confirm the efficacy and safety of ONO-4538 in patients with esophageal cancer.

1.2 Summary of Important Findings from Non-clinical and Clinical Studies

1.2.1 Summary of Non-clinical Data

ONO-4538 is a human monoclonal antibody that targets the human PD-1 receptor and is produced in Chinese hamster ovarian (CHO) cells via genetical combination technology. The drug substance is an aqueous solution of ONO-4538 that is clear or opalescent and colorless or pale yellow; fine particles may be slightly observed in the solution. The drug product is an aqueous injection containing 100 mg ONO-4538 in each vial and is stored away from light at 2 to 8°C. (refer to the Investigator's Brochure for details).

ONO-4538 specifically binds to the PD-1 receptor and has shown high affinity to human PD-1 and simian PD-1 with K_D values of 3.06 nmol/L and 3.92 nmol/L, respectively. ONO-4538 *in vitro* inhibited the binding of PD-1 with PD-1 ligands (i.e., PD-L1 and PD-L2), and human mixed lymphocyte reaction (MLR) showed increases in T-cell proliferation and IFN-gamma production. In an evaluation system in which human peripheral blood mononuclear cells (PBMCs) collected from donors previously infected with cytomegalovirus (CMV) were re-stimulated with CMV antigen, ONO-4538 increased the antigen-induced production of IFN-gamma in a concentration-dependent manner. In monkeys vaccinated with HBsAg, SKMel and DNP-Ficoll, ONO-4538 increased anti-SKMel cell specific antibody titers.

In multiple models of allogeneic tumor-bearing mice, an anti-mouse PD-1 antibody, 4H2, demonstrated a delay of tumor growth, thereby documenting the antitumor effect of anti-PD-1 antibody that blocks PD-1 ligand from binding to PD-1.

On the basis of the above, ONO-4538 is believed to exert its antitumor effect by inhibiting the binding between PD-1 and PD-1 ligand, thereby enhancing the activity of antigen-specific T-cells, resulting in an increased immune response.

Effects of ONO-4538 on the central nervous system and cardiovascular system were studied in conscious monkeys with single-dose intravenous administration up to 50 mg/kg, which showed no effects on the general conditions, body temperature, heart rate, blood pressure or electrocardiogram. Also in multiple-dose toxicity studies in monkeys with administration of ONO-4538 either once weekly for 4 weeks or twice weekly for 13 weeks, no effects were shown on the general condition, body temperature, heart rate, blood pressure or electrocardiogram up to the highest dose of 50 mg/kg. Effects of ONO-4538 on the respiratory system were studied in the 4-week multiple-dose toxicity study and the 13-week multiple-dose intravenous toxicity study in monkeys, which showed no effects up to the highest dose of 50 mg/kg.

After a single intravenous dose of ONO-4538 at 1, 10 and 50 mg/kg in monkeys, the AUC of ONO-4538 in serum increased in a near dose-proportional manner. The steady-state volume of distribution (Vss) was similar to the volume of plasma, thereby indicating the distribution of ONO-4538 primarily in circulating blood. Serum ONO-4538 concentrations over time were lower in the anti-ONO-4538 antibody-positive animals than in the anti-ONO-4538 antibody-negative animals. In a reproductive toxicity study in pregnant monkeys, ONO-4538 was detected in serum samples of neonates, indicating transmission of ONO-4538 to fetuses via the placenta of the mother.

In a single-dose toxicity study in monkeys, no deaths or toxic changes were observed with ONO-4538 up to the highest dose of 10 mg/kg. In a multiple-dose intravenous toxicity study in monkeys, no toxic changes were observed with ONO-4538 50 mg/kg administered once weekly for 4 weeks or twice weekly for 13 weeks. Thus, the no-observed-adverse-effect level in the 4-week and 13-week multiple dose toxicity studies was considered to be 50 mg/kg. In these studies, no changes suggestive of local irritation were observed at or around the intravenous administration site of ONO-4538 10 mg/mL given once weekly or twice weekly.

In expanded prenatal and postnatal development studies using pregnant monkeys, ONO-4538 at 10 mg/kg and above increased embryonic and fetal mortality during the third semester of gestationor increased neonatal mortality. However, no teratogenicity was observed, and no effects were shown on neonatal growth, behaviors, or immune function.

Crossreactivity with monkey and human normal tissues, antibody-dependent cellular toxicity (ADCC) activity, complement-dependent cytotoxicity (CDC) activity, and cytokine production in human whole

blood were studied. The study on crossreactivity using monkey and human normal tissues showed positive reactions of the lymphocyte membrane of humans and monkeys known to express PD-1, and also showed unexpected positive reactions of the cytoplasm of pituitary hormone-producing cells in monkeys and humans. However, ONO-4538, which is an antibody product and thus does not pass through the cell membrane, is not considered to have direct action on the cytoplasm of pituitary hormone-producing cells. Moreover, the 4-week and 13-week multiple-dose toxicity studies in monkeys showed no abnormalities in pituitary hormone levels (TSH), pituitary weight, or histopathological findings of the pituitary gland.

ONO-4538 did not show ADCC activity or CDC activity in activated human CD4-positive T-cells in the presence of effector cells (from human PBMCs) and human complements. ONO-4538 did not cause the production of cytokines in human whole blood.

In monkeys given ONO-4538 in combination with a human anti-CTLA-4 monoclonal antibody ipilimumab or an anti-LAG-3 monoclonal antibody BMS-986016 repeatedly for 4 weeks, inflammatory changes were observed in the large intestine or the central nervous system, which were not observed with ONO-4538 alone. These findings indicate that co-administration of a T-cell activating agent with ONO-4538 can lead to excessive potentiation of T-cell immune response.

1.2.2 Summary of Clinical Data and Summary of Known and Potential Risks and Benefits to Subjects

The efficacy, safety and pharmacokinetics (PK) of ONO-4538 have been evaluated in patients with cancers, including non-small-cell lung carcinoma, malignant melanoma, renal cell carcinoma, and esophageal carcinoma in phase I, II, and III studies sponsored by Ono Pharmaceutical or BMS. Ongoing clinical studies are evaluating ONO-4538 as monotherapy or in combination with chemotherapy agents, molecular-targeted agents or other immunotherapy agents.

After single-dose administration, the PK of ONO-4538 was linear with dose-proportional increase within the dose range of 0.3 to 10 mg/kg. After multiple doses, the PK of ONO-4538 was linear with dose-proportional increase of the exposure within the dose range of 0.1 to 10 mg/kg. The elimination and distribution of ONO-4538 were similar across different doses. A PPK analysis (interim data) indicated no dependence of the clearance of ONO-4538 on tumor types within the dose range of 0.1 to 10 mg/kg.

To date, approximately 8,600 subjects have received ONO-4538 as either monotherapy or in combination with other therapies. The safety profile of ONO-4538 monotherapy has been similar across different tumor types, and no dose dependency has been shown in the incidence of adverse events, severity, or causal relationship to the investigational product. However, inflammatory lung disease as an adverse event has been numerically more common in subjects with non-small-cell lung cancer than in subjects with other cancer types, in part because the relationship between ONO-4538 and lung symptoms or changes in imaging findings on X-ray could be equivocal in some subjects with non-small-cell lung cancer.

Ongoing clinical studies are evaluating the safety of the combination use of ONO-4538 with other agents, including ipilimumab, cytotoxic chemotherapy agents, angiogenesis inhibitors, and molecular-targeted agents. Of these, the one being most actively developed is the combination of ONO-4538 with ipilimumab for patients with non-small cell lung cancer, malignant melanoma and renal cell carcinoma. According to the data available to date, the safety profile of the combination therapy with ONO-4538 plus ipilimumab has been similar to that of either drug alone, but the incidence of adverse events was occasionally higher with the combination use of the two drugs.

Efficacy and safety data from completed and ongoing Japanese and overseas clinical studies are provided in the current investigator's brochure.

Some subjects have been reported to have required long-term high-dose corticosteroid therapy or immunosuppressive therapy for the treatment of adverse events that could be related to ONO-4538, with rare onset of opportunistic infection after immunosuppressive therapy in 3 subjects. One of these 3 subjects was a man with non-small-cell lung cancer, and during the use of ONO-4538 in combination with chemotherapy, he experienced corticosteroid-responsive pneumonitis. Immediately after completion of corticosteroid therapy, respiratory symptoms recurred, and a subsequent lung biopsy led to a diagnosis of invasive aspergillosis, without findings of pulmonary inflammation (pneumonitis). He died one week after the diagnosis of invasive aspergillosis. The second patient was a woman with non-small-cell lung cancer who, after receiving ONO-4538 monotherapy for more than 24 weeks, experienced corticosteroid-responsive pneumonitis. Subsequently, respiratory symptoms recurred and a lung biopsy led to a diagnosis of Aspergillus pneumonia, without findings of pulmonary inflammation (pneumonitis). Aspergillus pneumonia resolved after antifungal treatment. The remaining patient was a man with renal cell carcinoma who, during the use of ONO-4538 in

combination with pazopanib, experienced corticosteroid-responsive pneumonitis. High-dose corticosteroid therapy was given with dose tapering for more than 2 months. Subsequently, symptoms recurred, and a specimen collected during bronchoscopy led to a diagnosis of Pneumocystis jirovecii pneumonia, which completely resolved after antimicrobial therapy.

1. Japanese phase II study in patients with esophageal cancer (ONO-4538-07)

ONO-4538-07 is a Phase 2, multicenter, open-label, uncontrolled study of nivolumab (3 mg/kg Q2W) in esophageal cancer refractory or intolerant to combination therapy with fluoropyrimidine, platinum, and taxane-based drugs conducted solely in Japan. As of the data cut-off date of 28-Feb-2015, efficacy has been evaluated in 64 treated subjects. ORR according to RECIST v1.1 was 15.7% (10 of 64 subjects) by a central imaging assessment. The best overall response resulted in 1 CR (1.6%), 9 PR (14.1%), 17 SD (26.6%), 29 PD (45.3%), and 8 not estimable (12.5%) by central assessment at ONO imaging facilities.

2 Study Organization and Roles

See Annex 1 “Study Organization.”

Any revisions to Annex 1 will be made separately from protocol amendments.

3 Objective of the Study

This is a multicenter, randomized, open-label, docetaxel- or paclitaxel-controlled study to evaluate the efficacy and safety of ONO-4538 in patients with esophageal cancer refractory or intolerant to combination therapy with fluoropyrimidine and platinum-based drugs.

3.1 Primary Objective of the Study

To compare overall survival (OS) between the ONO-4538 group and control group (docetaxel or paclitaxel) in patients with esophageal cancer refractory or intolerant to combination therapy with fluoropyrimidine and platinum-based drugs

3.2 Secondary Objectives of the Study

- To compare progression-free survival (PFS) between the ONO-4538 group and control group (docetaxel or paclitaxel) in patients with esophageal cancer refractory or intolerant to combination therapy with fluoropyrimidine and platinum-based drugs.

- To compare objective response rates between the ONO-4538 group and control group (docetaxel or paclitaxel) in patients with esophageal cancer refractory or intolerant to combination therapy with fluoropyrimidine and platinum-based drugs.

4 Subjects

4.1 Subjects

Patients with esophageal cancer who are refractory or intolerant to combination therapy with fluoropyrimidine and platinum-based drugs will be enrolled.

4.2 Inclusion Criteria

Having provided written consent before participation in the study, patients must fulfill all of the following criteria to be eligible for randomization. If a randomized subject is found not to meet any of the following criteria before the first dose of the investigational product, the subject will not be started on the study treatment and will be withdrawn from the study.

1. Sex: Male or female
2. Age (at the time of informed consent): 20 years and older
3. Patients with esophageal cancer and whose major lesion in the esophagus (if already resected, the major lesion in the esophagus prior to resection) satisfies the following criteria. For patients with multiple lesions, the deepest invasion by clinical diagnosis should be considered the major lesion. Lesions other than the major lesion should be considered as secondary lesions. Esophageal cancer in this study is defined as a cancer that has primarily developed from the esophagus.
    - Patients with a major lesion located in the cervical esophagus or thoracic esophagus (upper, middle, or lower thoracic region; including the esophagogastric junction)
    - Patients whose histological type of major lesion was pathologically proven squamous cell carcinoma or adenosquamous cell carcinoma (pathological diagnosis of secondary lesions in the esophagus is not mandatory). Note: adenosquamous is defined as an uncommon malignant esophageal neoplasm containing coexisting elements of infiltrating squamous cell carcinoma (SCC) and adenocarcinoma (AC)
4. Patients who are refractory or intolerant to combination therapy with fluoropyrimidine and platinum-based drugs for esophageal cancer, have previously received 1 treatment regimen, and are not indicated for a radical resection. The definition of refractory should be defined as follows; and a therapy applicable to the following should be counted as 1 regimen.

- Patients whose PD or recurrence was confirmed by imaging during their initial chemotherapy (including chemoradiation) or within 8 weeks after the last dose[#1] of chemotherapy will be assessed as “refractory.”
- Patients who underwent a radical resection (R0 resection confirmed) in conjunction with chemotherapy including neo-adjuvant/adjuvant therapy and chemoradiation (including patients who underwent chemoradiation followed by salvage surgery) whose recurrence was confirmed by imaging within 24 weeks after the last dose[#1] of chemotherapy will be determined as “refractory.”
- If a CR (≥2 consecutive CRs confirmed by imaging after an interval of ≥4 weeks) was assessed as a result of the initial chemotherapy (including chemoradiation), patients whose recurrence was confirmed by imaging during the initial chemotherapy (including chemoradiation) or within 24 weeks after the last dose[#1] of chemotherapy will be determined as “refractory.”

[#1] In case of chemoradiation, this will be the last dose of chemotherapy or the last radiation treatment, whichever occurs later.

5. Patients who have at least 1 measurable or non-measurable lesion per the RECIST Guideline Ver. 1.1 as confirmed by imaging within 28 days before randomization. The following requirements should also be satisfied:
   - The primary esophageal cancer should be deemed to be non-measurable lesion.
   - If patients only have lesions that were previously treated with radiation, the lesion should be limited to one with confirmed aggravation by imaging after radiation.
   - If patients have pericardial or pleural effusion or ascites only, the lesion should be limited to one with cytologically confirmed malignancy.
6. ECOG Performance Status Score 0 or 1
7. Patients with a life expectancy of at least 3 months
8. Patients must provide tumor tissue (stored tissue or tissue from the last biopsy) for analysis of PD-L1 expression. For patients who are unable to undergo another biopsy, stored tissue can be used for analysis. Tissue specimens must contain at least 100 evaluable tumor cells and must be available before the randomization.
9. Patients whose latest laboratory data meet the below criteria within 7 days before randomization. If the date of the laboratory tests at the time of randomization is not within 7 days before the first dose of the investigational product, testing must be repeated within 7 days before the first dose of the investigational product, and these latest laboratory tests must meet the following criteria. Of

note, laboratory data will not be valid if the patient has received a granulocyte colony-stimulating factor (G-CSF) or blood transfusion within 14 days before testing.

- White blood cells ≥2,000/mm^3 and neutrophils ≥1,500/mm^3
- Platelets ≥100,000/mm^3
- Hemoglobin ≥9.0 g/dL
- AST (GOT) and ALT (GPT) ≤3.0-fold the upper limit of normal (ULN) of the study site (or ≤5.0-fold the ULN of the study site in patients with liver metastases)
- Total bilirubin ≤1.5-fold the ULN of the study site
- Creatinine ≤1.5-fold the ULN of the study site or creatinine clearance (either the measured or estimated value using the Cockcroft-Gault equation) >45 mL/min

10. Women of childbearing potential (including women with chemical menopause or no menstruation for other medical reasons)[#1] must agree to use contraception[#2] from the time of informed consent until 5 months or more after the last dose of the investigational product. Also, women must agree not to breastfeed from the time of informed consent until 5 months or more after the last dose of the investigational product.
11. Men must agree to use contraception[#2] from the start of study treatment until 7 months or more after the last dose of the investigational product.

[#1] Women of childbearing potential are defined as all women after the onset of menstruation who are not postmenopausal and have not been surgically sterilized (e.g., hysterectomy, bilateral tubal ligation, bilateral oophorectomy). Postmenopause is defined as amenorrhea for ≥12 consecutive months without specific reasons. Women using oral contraceptives, intrauterine devices, or mechanical contraception such as contraceptive barriers are regarded as having childbearing potential.

[#2] The subject must consent to use any two of the following methods of contraception: vasectomy or condom for patients who are male or female subject's partner and tubal ligation, contraceptive diaphragm, intrauterine device, spermicide, or oral contraceptive for patients who are female or male subject's partner.

Rationale

1. No need to limit the gender of study subjects was considered and set.
2. The age limit was set at the age of voluntary consent to participate in the study.

3. This criterion was set to select patients with primary esophageal cancer as the subjects of the study, and those with esophageal cancer of a histological type amenable to efficacy and safety evaluation.
4. This criterion was set to include patients who do not benefit from chemotherapy with fluoropyrimidine and platinum-based drugs.
5. This criterion was set to identify patients evaluable for efficacy in this study.
6. This criterion was set to enroll patients with a favorable performance status.
7. This criterion was set to select patients evaluable for efficacy and safety in this study.
8. This criterion was set to select patients evaluable for PD-L1 expression.
9. This criterion was set to select patients with adequate functioning of major organs.
10, 11. The safety of ONO-4538 in pregnant women and fetuses has not been established.

4.3 Exclusion Criteria

Patients who meet any of the following criteria at the time of assessment for randomization will be excluded. If a randomized subject is found to meet any of the following criteria before the first dose of the investigational product, the subject will not be started on the study treatment and will be withdrawn from the study.

1. Patients with significant malnutrition. Patients will be excluded if they are receiving intravenous hyperalimentation, or require continuous infusion therapy with hospitalization. Patients whose nutrition has been well controlled for ≥28 days prior to randomization may be enrolled.
2. Patients with apparent tumor invasion on organs located adjacent to the esophageal disease (e.g., the aorta or respiratory tract). Patients will be excluded if they are receiving stent therapy in esophagus or respiratory tract.
3. Patients with multiple primary cancers (with the exception of completely resected basal cell carcinoma, stage I squamous cell carcinoma, carcinoma in situ, intramucosal carcinoma, or superficial bladder cancer, or any other cancer that has not recurred for at least 5 years)
4. Patients with residual adverse effects of prior therapy or effects of surgery that would affect the safety evaluation of the investigational product in the opinion of the investigator or subinvestigator
5. Patients with current or past history of severe hypersensitivity to any other antibody products
6. Patients with concurrent autoimmune disease or history of chronic or recurrent autoimmune disease (see Appendix 4)

7. Patients with a current or past history of interstitial lung disease or pulmonary fibrosis diagnosed based on imaging or clinical findings. Patients with radiation pneumonitis may be randomized if the radiation pneumonitis has been confirmed as stable (beyond acute phase) without any concerns about recurrence.
8. Patients with concurrent diverticulitis or symptomatic gastrointestinal ulcerative disease
9. Patients with any metastasis in the brain or meninx that is symptomatic or requires treatment. Patients may be randomized if the metastasis is asymptomatic and requires no treatment.
10. Patients with pericardial fluid, pleural effusion, or ascites requiring treatment
11. Patients with uncontrollable, tumor-related pain
12. Patients who have experienced a transient ischemic attack, cerebrovascular accident, thrombosis, or thromboembolism (pulmonary arterial embolism or deep vein thrombosis) within 180 days before randomization
13. Patients with a history of uncontrollable or significant cardiovascular disease meeting any of the following criteria:
    - Myocardial infarction within 180 days before randomization
    - Uncontrollable angina pectoris within 180 days before randomization
    - New York Heart Association (NYHA) Class III or IV congestive heart failure
    - Uncontrollable hypertension despite appropriate treatment (e.g., systolic blood pressure ≥150 mmHg or diastolic blood pressure ≥90 mmHg lasting 24 hours or more)
    - Arrhythmia requiring treatment
14. Patients receiving or requiring anticoagulant therapy for a disease. Patients receiving antiplatelet therapy including low-dose aspirin may be enrolled.
15. Patients with uncontrollable diabetes mellitus
16. Patients with systemic infections requiring treatment
17. Patients with ≥ Grade 2 peripheral neuropathy
18. Patients who have received systemic corticosteroids (except for temporary use, e.g., for examination or prophylaxis of allergic reactions) or immunosuppressants within 28 days before randomization
19. Patients who have received antineoplastic drugs (e.g., chemotherapy agents, molecular-targeted therapy agents, or immunotherapy agents) within 28 days before randomization
20. Patients who have undergone surgical adhesion of the pleura or pericardium within 28 days before randomization
21. Patients who have undergone surgery under general anesthesia within 28 days before

randomization

22. Patients who have undergone surgery involving local or topical anesthesia within 14 days before randomization
23. Patients who have received radiotherapy within 28 days before randomization, or radiotherapy to bone metastases within 14 days before randomization
24. Patients who have received any radiopharmaceuticals (except for examination or diagnostic use of radiopharmaceuticals) within 56 days before randomization
25. Patients with a positive test result for any of the following: HIV-1 antibody, HIV-2 antibody, HTLV-1 antibody, HBs antigen, or HCV antibody
26. Patients with a negative HBs antigen test but a positive test result for either HBs antibody or HBc antibody with a detectable level of HBV-DNA
27. Women who are pregnant or breastfeeding, or possibly pregnant
28. Patients who have received any other unapproved drug (e.g., investigational use of drugs, unapproved combined formulations, or unapproved dosage forms) within 28 days before randomization
29. Patients who have previously received taxane agents to treat esophageal cancer. Patients who were not proven refractory (see the definition of refractory in inclusion criteria 4) or intolerant to taxane-based combination therapy and subsequently received fluoropyrimidine and platinum-based combination therapy, and then proven refractory or intolerant may be randomized.
30. Patients who are contraindicated to docetaxel and paclitaxel
31. Patients who have previously received ONO-4538 (MDX-1106 or BMS-936558), anti-PD-1 antibody, anti-PD-L1 antibody, anti-PD-L2 antibody, anti-CD137 antibody, anti-CTLA-4 antibody or other therapeutic antibodies or pharmacotherapies for regulation of T-cells
32. Patients judged to be incapable of providing consent for reasons such as concurrent dementia
33. Other patients judged by the investigator or subinvestigator to be inappropriate as subjects of this study

Rationale

1 to 17. These criteria were set to secure the safety of patients and to avoid possible effects on evaluations of the efficacy and safety of ONO-4538.

18 to 24. These criteria were set to avoid possible effects of prior and concomitant treatments on evaluations of the efficacy and safety of ONO-4538.

25, 26. These criteria were set to secure the safety of patients and to avoid possible effects on evaluations of the efficacy and safety of ONO-4538.

27. The safety of ONO-4538 in pregnant women, fetuses, and infants has not been established.

28. This criterion was set to avoid effects of unapproved drugs with unestablished drug efficacy.

29, 31. This criterion was set to avoid possible effects on the efficacy evaluation.

30. This criterion was set to avoid possible effects on the safety evaluation.

32. This criterion was set to protect patients' human rights.

33. This criterion was set to exclude patients not meeting any of the exclusion criteria 1 to 27 but who are inappropriate for other reasons.

5 Study Methods

5.1 Study Design

5.1.1 Study Design

This is a multicenter, randomized, open-label, docetaxel- or paclitaxel-controlled study where ONO-4538 is administered at 2-week intervals for the treatment of subjects with esophageal cancer refractory or intolerant to combination therapy with fluoropyrimidine and platinum-based drugs, with up to 1 prior regimen to evaluate the efficacy and safety of ONO-4538.

<Rationale for the Study Design>

This study will be conducted in patients with esophageal cancer refractory or intolerant to combination therapy with fluoropyrimidine and platinum-based drugs, the standard chemotherapy treatment for esophageal cancer, with up to 1 prior regimen.

Esophageal cancer patients without distal metastases are currently provided according to their stage of cancer with endoscopic treatment, surgical treatment, radiation therapy, and chemotherapy selected in a multimodal manner for radical cure. On the other hand, esophageal cancer patients with unresectable distal metastases are treated mainly with chemotherapy in which a combination of fluoropyrimidine and a platinum-based drug, 5-FU and cisplatin (FP therapy) is recommended as the standard therapy. Second-line treatments after FP therapy are generally taxanes such as docetaxel and paclitaxel. Among patients with esophageal cancer for which chemotherapy is indicated, those with esophageal adenocarcinoma benefit from docetaxel in terms of survival prolongation; however, no chemotherapy shown to prolong survival in esophageal squamous cell carcinoma currently exists.

Currently, a phase II multicenter, open-label, uncontrolled study (ONO-4538-07) is underway to explore the efficacy and safety of multiple-dose intravenous administration of ONO-4538 (3 mg/kg) at

2-week intervals in patients with esophageal cancer refractory or intolerant to fluoropyrimidine-, platinum-, or taxane-based standard therapy. This study has suggested that the antitumor effect of ONO-4538 in esophageal cancer is sustainable for a long duration, as in other types of cancer. In addition, since squamous cell carcinoma is the major histological type of esophageal cancer in Asia, a global phase III study of ONO-4538 (CA209017) was conducted on a similar histological type, squamous non-small cell lung cancer, where the target population was patients with stage IIIB/IV or recurrent squamous non-small cell lung cancer who received prior platinum-containing chemotherapy. The study demonstrated that ONO-4538 has a significantly higher prolongation of survival (HR=0.59; p=0.00025) over existing therapy (docetaxel). Based on the above, a randomized open-label study, with docetaxel and paclitaxel controls that are widely used in 2nd-line treatment, is planned in patients with esophageal cancer after FP therapy to evaluate the efficacy and safety of ONO-4538.

5.1.2 Planned Sample Size

195 subjects in each group, totaling 390 subjects

<Rationale for the Sample Size>

This study is intended to verify the superiority of ONO-4538 group over the control groups (docetaxel or paclitaxel) in terms of OS, the primary endpoint, in patients with esophageal cancer refractory or intolerant to fluoropyrimidine and platinum-based agents, the standard chemotherapy regimen for the treatment of esophageal cancer. Subjects will be randomized to either the ONO-4538 group or the control group in a 1:1 ratio and stratified by region (Japan vs. rest of world) and number of organs with metastases (≤1 vs. ≥2) and PD-L1 expression (≥1% vs. <1% or indeterminate) [22)-24)].

As for OS, assuming an exponential distribution for the control group (7.2 months median OS) and a piecewise mixture model for the ONO-4538 group with 5.0% long term survival rate (the ratio that long term survival is expected due to the effect of ONO-4538), the hazard ratio of the ONO-4538 group relative to the control group was assumed to be 1.0 for the first 3 months and 0.65 for non-long term survival population thereafter (10.0 months overall median OS for the ONO-4538 group), the expected average hazard ratio of the ONO-4538 group to the control groups was assumed to be 0.70 in this study. The number of events required to detect superiority of the ONO-4538 group over the control groups with two-sided significance level of 5% and 90% or more power by the log-rank test was calculated to be 331.

Assuming the enrollment period to be 16 months and the follow-up period after the last patient's enrollment to be 18 months, the number of subjects required to ensure the required 331 events was

estimated to be 390. For the calculation of the required events and sample size at the time of planning the study, the statistical analysis software SAS (version 9.3) was used.

<Rationale for assuming the median OS in the control group to be 7.2 months>

The phase II single-arm study of docetaxel and the retrospective study of paclitaxel as 2nd-line treatment for patients with esophageal cancer reported that the median OS of docetaxel and paclitaxel was 5.5-8.1 months and 6.1-10.4 months, respectively, [17), 18), 25), 26)]. Based on this information, the median OS in the control groups of this study is assumed to be 7.2 months.

<Rationale for assuming long term survival rate (the ratio that long term survival is expected due to the effect of ONO-4538) as 5.0% and the hazard ratio of the ONO-4538 group relative to the control groups was assumed to be 1.0 for the first 3 months>

We set long term survival rate and the hazard ratio referring to the results of a randomized phase 3 clinical trial of nivolumab vs investigator's choice in recurrent or metastatic platinum-refractory squamous cell carcinoma of the head and neck (CA209141) [28)] and a phase 2 uncontrolled study of nivolumab in esophageal cancer refractory or intolerant to standard therapy (ONO-4538-07)[29)].

<Rationale for setting the hazard ratio of the ONO-4538 group relative to the control groups was assumed to be 0.65 after first 3 months in non-long term survival population>

In the global phase III study conducted in patients with stage IIIB/IV or recurrent squamous non-small cell lung cancer who received a prior platinum-containing chemotherapy regimen (CA209017) to verify the superiority of ONO-4538 over docetaxel, one of the controls in this study, the hazard ratio of ONO-4538 relative to docetaxel was 0.59 [96.85% confidence interval: 0.43, 0.81].

If the same histological type of cancer and control are used in the present study, the efficacy on ONO-4538 is expected to be similar. However, this study includes paclitaxel as well as docetaxel as controls. Efficacy data on ONO-4538 vs. paclitaxel has not been obtained previously. Based on these reasons, the expected average hazard ratio of the ONO-4538 group to the control groups was assumed to be 0.70 in this study.

In this study, since we assumed that 5.0% of long term survival rate and the hazard ratio of the ONO-4538 group relative to the control group was to be 1.0 for the first 3 months, the hazard ratio of the ONO-4538 group relative to the control groups after first 3 months in non-long term survival population was calculated as 0.65 to become the expected average hazard ratio of the ONO-4538 group relative to the control groups was 0.70.

5.2 Method of Assigning Subjects to Treatment Groups

5.2.1 Method for investigational product allocation

Subjects who meet all inclusion criteria will be assigned to the ONO-4538 group or control group (docetaxel group or paclitaxel group) in a 1:1 ratio using the following assignment factors.

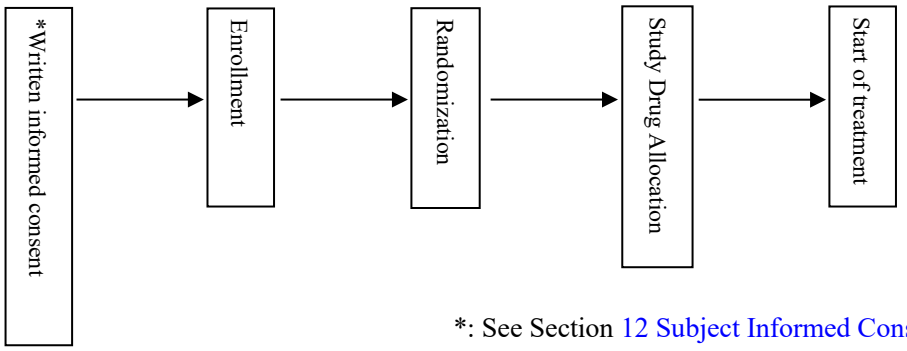
- Location (Japan vs. rest of world)
- Number of organs with metastases (≤1 vs. ≥2) (at randomization)
- Expression of PD-L1 (≥1% vs. <1% or indeterminate)

Location (Japan vs. rest of world) was used as a stratification factor in order to uniformly assign subjects to the ONO-4538 or control group (docetaxel group or paclitaxel group) in each location. The number of organs with metastases (≤1 vs. ≥2) and PD-L1 expression (≥1% vs. <1% or indeterminate) were selected based on a published report on prognostic factors in esophageal cancer patients [22)-24)] as well as specialist advice.

Stratified randomization will be carried out using a permuted block method. Administration of the investigational product (ONO-4538, docetaxel, or paclitaxel) will be started within 7 days after the randomization.

5.3 Subject Enrollment Procedures

Subjects will be enrolled according to the following procedures.



*: See Section 12 Subject Informed Consent

Each subject who has provided written informed consent will be enrolled in the study by the IWRS (enrollment). Detailed procedures for the use of the IWRS will be specified in a separate document provided to each study site. The investigator or designee will enroll subjects according to predetermined procedures. The following data are required for enrollment after informed consent:

- Date of informed consent
- Date of birth
- Sex

Subjects who meet all inclusion criteria after enrollment are eligible for randomization by the IWRS (randomization). The following data are required for randomization.

- Subject identification code (ID)
- Date of birth
- Location (Japan vs. rest of world)
- Number of organs with metastases (≤1 or ≥2) (at randomization)
- PD-L1 expression (≥1% vs. <1% or indeterminate)

Administration of the investigational product must be started within 7 days following randomization.

If an enrolled subject is withdrawn without randomization, the investigator or designee will communicate the withdrawal before randomization using IWRS. If an eligible subject is withdrawn after randomization for specific reasons before administration of the investigational product, the investigator or designee will communicate the withdrawal before administration of study drug using IWRS.

5.4 Blinding

Not applicable.

5.5 Endpoints

5.5.1 Efficacy Endpoints

5.5.1.1 Primary Endpoint

Overall survival (OS)

Rationale

OS is defined as the time from randomization until death from any cause. OS was selected as the primary endpoint of this study because it is generally regarded as the most reliable cancer endpoint and is usually the preferred endpoint for studies that can be conducted to adequately assess survival.

5.5.1.2 Secondary Endpoints

1. Objective response rate (ORR)
2. Disease control rate (DCR)
3. Progression free survival (PFS)
4. Duration of response
5. Time to response
6. Best overall response (BOR)
7. Maximum percent change from baseline in the sum of diameters of the target lesion

Rationale

These endpoints were set to assess the efficacy of ONO-4538 and docetaxel or paclitaxel from various perspectives in patients with esophageal cancer.

5.5.2 Safety Endpoints

1. Adverse events
2. Clinical laboratory tests
3. Vital signs
4. 12-lead ECG
5. Chest X-ray
6. Performance Status (ECOG)

Rationale

These endpoints were set to assess the safety of ONO-4538 and docetaxel or paclitaxel from various perspectives in patients with esophageal cancer.

5.5.3 Pharmacokinetic Endpoint

Serum ONO-4538 concentrations (for the ONO-4538 group only)

5.5.4 Anti-drug Antibody Endpoint

Anti-ONO-4538 antibody (for the ONO-4538 group only)

5.5.5 Exploratory Biomarkers

[redacted]

6. Tumor tissue examination (PD-L1 expression analysis, essential; [redacted])

5.5.6 Other Test Variables

1. Tumor markers
2. Patient reported outcomes (PROs)
3. Healthcare Resource Utilization

Rationale

These variables have been employed for exploratory examination of biomarkers that may be related to the efficacy or safety of ONO-4538, including PD-L1 expression, [redacted]

6 Investigational Products

An investigational product, also known as an investigational medicinal product in some regions, is defined as a pharmaceutical form of an active substance or placebo being tested or used as a reference in a clinical study, including products already with marketing authorization but used or assembled

(formulated or packaged) differently from the authorized form or used for an unauthorized indication, or when used to gain further information about the authorized form.

The investigational product should be stored in a secure area according to local regulations. The investigator has the responsibility to ensure that investigational product is only dispensed to study subjects. The investigational product must be dispensed only from official study sites by authorized personnel according to local regulations.

In this protocol, the investigational products are ONO-4538 (nivolumab), docetaxel, and paclitaxel.

Table 6-1 Description of Investigational Products

Name and dosage form of the investigational product	Contents of the primary packaging (small box)	No. of vials in the secondary packaging (large box)	Appearance	Storage conditions (as per the indication on the label)
ONO-4538 (Nivolumab) For injection	100 mg/10 mL Vials	6	Colorless to pale yellow clear to opalescent liquid, occasionally containing microparticles	Store refrigerated at 2 to 8°C with protection from light.
Docetaxel For i.v. drip infusion	80 mg/2 mL Vials	3	Yellow to orange–yellow clear viscous liquid	Store at 1 to 30°C with protection from light.
(Accompanying reconstitution diluent) Ethanol	764.4 mg/6 mL Vials	3	Colorless clear liquid	Store at 1 to 30°C with protection from light.
Paclitaxel For injection	100 mg/16.7 mL Vials	6	Colorless to pale yellow clear, viscous oily liquid	Store at 1 to 30°C with protection from light.
Docetaxel concentrate for solution for infusion*	80 mg/vial (10 mg/mL)	1	Clear, pale yellow to brownish-yellow solution.	Store below 25°C. Store in original package and protect from light.
Paclitaxel Solution for Injection*	100 mg/vial (6 mg/mL)	4	Clear, colorless or slightly yellow viscous solution	Store at 15 to 30°C. Protect from light.

*Comparators used in "United States and European countries. These products may be obtained by the investigational sites as local commercial product in certain countries if allowed by local regulations. In these cases, product may be a different pack size/potency than listed in this table.

6.1 Handling and Supply of Investigational Product

The investigational product will be supplied to individual study sites as listed in table 6-1.

The investigator should confirm that the investigational products are being stored under the environmental conditions (temperature, light, and humidity) specified by the sponsor. If concerns arise

regarding the quality or appearance of the investigational product, the product should not be dispensed and the sponsor should be contacted immediately.

The investigator should maintain investigational product documentation that includes all of the processes required to ensure that the investigational product is administered accurately. This will include documentation of investigational product storage, administration and, as applicable, storage temperatures, reconstitution, and use of required processes (e.g., a required diluent, administration sets).

For further details regarding preparation/administration and storage of nivolumab, docetaxel, and paclitaxel please refer to the current Investigator Brochure, pharmacy reference sheet and/or appropriate package insert/SPC.

It is the responsibility of the investigator/site to provide any supplies required to prepare/administer Investigational Product (e.g., filters, diluents, IV bags, administration sets) and any required pre-medication treatments (e.g., for paclitaxel).

6.2 Procedures for Investigational Product Accountability and Return of Unused Investigational products

6.2.1 Investigational Product Storage and Management

The investigator is responsible for ensuring that a current status record of the investigational product (supplied by the sponsor) is maintained at each study site where the investigational product is inventoried and dispensed. Records or logs must comply with applicable regulations and guidelines and should include:

- amount received and placed in storage area
- amount currently in storage area
- label identification number or batch number
- amount dispensed to and returned by each subject, including Subject Identification Codes
- amount transferred to another area/site for dispensing or storage
- non-study status (e.g., lost, wasted)
- amount destroyed at study site, if applicable
- amount returned to the sponsor
- samples retained for bioavailability/bioequivalence, if applicable
- dates and initials of person responsible for investigational product dispensing/accountability, as per the delegation of authority form

The sponsor will provide forms to facilitate inventory control if the study site does not have an established system that meets these requirements.

6.2.2 Disposal of Investigational Products

In this study, the investigational products (those supplied by the sponsor or sourced by the investigator), such as partially used investigational product containers, vials and syringes, may be destroyed on site.

Any unused investigational product will not be destroyed unless investigational product containers must be immediately destroyed as required for safety or to meet local regulations (e.g., cytotoxics or biologics).

On-site destruction is allowed provided the following minimal standards are met:

- On-site disposal practices must not expose humans to risks from the drug.
- On-site disposal practices and procedures are in agreement with applicable laws and regulations, including any special requirements for controlled or hazardous substances.
- Written procedures for on-site disposal are available and followed. The procedures must be filed with the site's Standard Operating Procedures (SOPs) and a copy provided to the sponsor upon request.
- Records are maintained that enable the traceability of each container, including the date of disposal, quantity disposed, and the person disposing of the containers. The method of disposal, i.e., incinerator, licensed sanitary landfill, or licensed waste disposal vendor must be documented.
- Investigational product accountability and disposal records must be complete, up-to-date, and accessible to the sponsor for review throughout the clinical study period.

If the conditions for destruction of the investigational product cannot be met, the responsible monitor assigned by the sponsor will make arrangements to return the investigational product.

Provided that procedures for proper disposal have been established according to applicable local and institutional guidelines and procedures, and provided that appropriate records of disposal are kept, the investigator will arrange the disposal of all empty containers.

6.2.3 Return of Unused Investigational Products

If the investigational product will not be destroyed upon completion or termination of the study, all unused and/or partially used investigational product that was supplied by the sponsor must be returned to the sponsor.

The investigator is responsible for arranging the disposal of all empty containers, provided that procedures for proper disposal have been established according to applicable local and institutional guidelines and procedures, and provided that appropriate records of disposal are kept.

7 Study Plan

7.1 Investigational Product

7.1.1 Dosage, Administration, and Treatment Duration

7.1.1.1 Dosage, Administration, and Treatment Duration of ONO-4538

ONO-4538 240 mg will be administered intravenously over 30 minutes at 2-week intervals. Six weeks will count as 1 cycle of treatment. Treatment will be continued until PD is assessed by the investigator or subinvestigator according to the RECIST Guideline Ver. 1.1. ONO-4538 will be administered at 240 mg without any increase or decrease of the dose and when 11 or more days have passed since the last dose, i.e., at an interval of 10 days or more. For information regarding appropriate storage, handling, dispensing, and administration of ONO-4538, refer to the latest Investigator's Brochure and/or Reference Sheet for Pharmacy.

<Rationale for Setting the Dosage and Administration of ONO-4538>

In the global phase I multiple-dose study (CA209003), multiple doses of ONO-4538 at 0.1-10 mg/kg was intravenously administered every 2 weeks to 306 patients with malignant tumors, including non-small-cell lung cancer, malignant melanoma, and renal cell carcinoma. Favorable antitumor effects with ONO-4538 ≥3 mg/kg have been shown in the study.

In terms of safety, the tolerability of multiple intravenous doses of ONO-4538 monotherapy every 2 weeks has been demonstrated up to 20 mg/kg in Japan and up to 10 mg/kg in overseas countries in phase I studies (ONO-4538-01 and CA209003). In the global phase I multiple-dose study (CA209003), the type, incidence, and severity of adverse events (AEs) for which a causal relationship to ONO-4538 could not be ruled out were similar among the evaluated doses.

These efficacy and safety data led to the use of ONO-4538 at a dose of 3 mg/kg every 2 weeks in ongoing phase II and III studies in patients with non-small-cell lung cancer, malignant melanoma, and renal cell carcinoma.

Based on these data taken together, “multiple intravenous doses of ONO-4538 at 3 mg/kg every 2 weeks” was selected as the clinically recommended dosage for esophageal cancer, in line with the recommended dosage regimen for other cancer types. In addition, the ongoing Japanese phase II study of ONO-4538 3 mg/kg (given at a 2-week interval) in patients with esophageal cancer refractory or intolerant to standard therapies (ONO-4538-07) has suggested the efficacy and safety of this regimen in esophageal cancer.

As described above, multiple clinical studies have been conducted in which ONO-4538 is administered at a dose per body weight (mg/kg). The results of population pharmacokinetic (PPK) analysis demonstrated a linear PK profile of ONO-4538 and a dose-proportional increase in exposure to ONO-4538 within the range of 0.1 to 10 mg/kg. No difference has been suggested in the PK of ONO-4538 among cancer types. The clearance and distribution volume of ONO-4538 increased as body weight increased; however, the increase in these parameters was less than proportional to the increase in body weight. On the other hand, considering that the PK of ONO-4538 correlates with body weight, administration of ONO-4538 at a fixed dose (mg), which is not adjusted for body weight, is expected to provide lower exposure in heavy subjects than in light subjects. The predicted trough, peak, and mean concentrations at steady state (Cminss, Cmaxss, and Cavgss, respectively) after administration of ONO-4538 at a dose of 3 mg/kg or 240 mg to patients with non-small-cell lung cancer (NSCLC) are shown in Table 7.1.1.1-1. The median body weight in the three phase II and III studies of ONO-4538 monotherapy (CA209017, CA209057, and CA209063) was approximately 80 kg. ONO-4538 240 mg is equivalent to 3 mg/kg ONO-4538 in subjects weighing 80 kg. As shown in Table 7.1.1.1-1, the geometric means of the Cminss, Cmaxss, and Cavgss after administration of ONO-4538 at a fixed dose (240 mg) were slightly higher (<15%) than those at 3 mg/kg, and the variation coefficient at 240 mg was also slightly higher (<10%) than that at 3 mg/kg. Since no difference has been suggested in the PK of ONO-4538 among cancer types, the predicted values in patients with NSCLC in Table 7.1.1.1-1 are applicable to patients with other cancer types.

Table 7.1.1.1-1 Exposure of ONO-4538 at steady state

Dose of ONO-4538	Cminss [μg/mL]	Cmaxss [μg/mL]	Cavgss [μg/mL]
240 mg	61.5 (44.6)	133.7 (35.0)	84.2 (38.2)
3 mg/kg	54.7 (41.9)	118.9 (31.8)	73.3 (35.6)

Geometric mean (variation coefficient %)

The safety and tolerance of ONO-4538 have been confirmed up to 10 mg/kg. The relationship between the exposure and efficacy of ONO-4538 was relatively consistent at 3 mg/kg. The above PK,

safety, and efficacy data of ONO-4538 suggest that the safety and efficacy profile of ONO-4538 at 240 mg is similar to that at 3 mg/kg.

Based on the above information taken together, a dose of 240 mg (multiple-dose intravenous administration at a 2-week interval) was selected for global clinical studies of ONO-4538 in esophageal cancer.

7.1.1.2 Dosage, Administration, and Treatment Duration of Docetaxel

Docetaxel will be intravenously administered at a dose of 75 mg/m^2 every 3 weeks. Each treatment cycle lasts 3 weeks. Treatment will be continued until PD is assessed by the investigator or subinvestigator according to the RECIST Guideline Ver 1.1. To calculate the first dose of the docetaxel, the body weight measured at study enrollment can be used instead of that on the day of administration (predose). If the subject has had a ≥10% change in body weight from that measured at the initial dose, the dose should be adjusted. If the change is observed on the day of administration from the initial dose, the dose can be adjusted from the next administration onwards. Each dose should be similarly adjusted if a further body weight change of ≥10% is observed after the previous change. The dose (mg) will be rounded to one decimal place. Although administration should be performed per the procedure of the study site, the recommended procedure is as follows: Docetaxel will be intravenously administered over at least 60 minutes in accordance with the package insert. Adverse reactions to docetaxel must be treated appropriately with reference to local standards, such as the package insert and treatment guidelines.

<Rationale for Setting Dosage and Administration of Docetaxel>

The standardized dosing regimen of docetaxel was established for this study with reference to local standards, such as the package insert of docetaxel and treatment guidelines in each country.

7.1.1.3 Dosage, Administration, and Treatment Duration of Paclitaxel

Paclitaxel will be administered at a 100 mg/m^2 dose weekly for 6 weeks followed by a 2-week drug holiday. This treatment cycle will be repeated until PD is assessed by the investigator or subinvestigator in accordance with the RECIST Guideline Ver 1.1. There should be an interval of at least 5 days between the doses of paclitaxel. To calculate the first dose of paclitaxel, the body weight measured at study enrollment can be used instead of that on the day of administration (predose). The dose on Day 1 of each cycle should be continued throughout that cycle as a rule. If the subject has had a ≥10% change in body weight from that measured at the initial dose, the dose should be adjusted. If

the change is observed on the day of administration from the initial dose, the dose can be adjusted from the next administration onwards. Each dose should be similarly adjusted if a further body weight change of ≥10% is observed after the previous change. The dose (mg) will be rounded to one decimal place. Although administration should be performed per procedure of the study site, recommended procedure is as follows:

Paclitaxel will be administered intravenously through an in-line filter using a membrane filter of ≤0.22 μm over 60 minutes in accordance with the package insert. Adverse reactions to paclitaxel must be treated appropriately with reference to local standards, such as the package insert and treatment guidelines.

<Rationale for Setting Dosage and Administration of Paclitaxel>

The standardized dosing regimen of paclitaxel was established for this study with reference to local standards, such as the package insert of paclitaxel and treatment guidelines in each country.

7.1.2 Criteria for Administration of Investigational Product

Subjects who meet all of the criteria for administration of the investigational product and meet none of the criteria for its discontinuation will be allowed to continue study treatment.

The following terms are defined for the administration criteria.

Discontinuation: The study treatment is discontinued permanently without resuming the treatment.

Delay: The interval between administrations is extended; or administration is delayed from the predetermined schedule.

Skip: Administration of the investigational products is skipped to the next dosing schedule.

7.1.2.1 Criteria for Administration of ONO-4538

Before the start of each infusion, the subject must fulfill the following criteria to receive the dose on the basis of the latest data available at that time. If the subject does not meet any of the following criteria at the start of each infusion, the scheduled administration of ONO-4538 will be skipped. However, subjects will be allowed to continue study treatment if the treatment is expected to lead to clinical benefits and is assessed to be administered safely unless investigator-assessed disease progression is observed. In this case, the subject's consent to continue study treatment must be obtained prior to the continuation of study treatment and recorded in the medical record. The continuation should be reported to the Sponsor as in a timely manner as possible. The protocol-

specified diagnostic imaging for the assessment of tumor response and other examinations for safety assessment will be performed as scheduled, irrespective of the administration of ONO-4538.

1. No Grade ≥3 adverse events for which a causal relationship to ONO-4538 cannot be ruled out (i.e., adverse drug reactions).
2. No Grade ≥2 diarrhea or colitis for which a causal relationship to ONO-4538 cannot be ruled out.
3. No Grade ≥2 creatinine increased for which a causal relationship to ONO-4538 cannot be ruled out.
4. No Grade ≥2 neurotoxicity for which a causal relationship to ONO-4538 cannot be ruled out.
5. No increase of ≥2 grades from baseline (predose of ONO-4538) in AST (GOT), ALT (GPT), or total bilirubin for which a causal relationship to ONO-4538 cannot be ruled out.
6. No suspected autoimmune disease based on signs/symptoms or clinical laboratory tests.
7. The investigator or subinvestigator judges that the dose of ONO-4538 may be administered to the subject.

7.1.2.2 Criteria for Administration of Docetaxel

Before the start of each infusion, subjects must fulfill all of the following criteria to receive the dose on the basis of the latest data available at that time. At the start of the treatment cycles, if the subject does not meet any of the following criteria, the scheduled dose of docetaxel should be delayed.

If no worsening of clinical symptoms determined due to disease progression is observed, and continuation of the study treatment is expected to lead clinical benefits, and the dose of docetaxel can be continued safely without contraindications, treatment with docetaxel may be continued. In this case, the subject's willingness to continue with docetaxel treatment must be confirmed prior to continuation and must be documented in the medical record. In addition, it should be reported to the sponsor as in a timely manner as possible. Imaging examinations to assess tumor response should be performed as per the schedule specified in the protocol, irrespective of the administration of docetaxel. When docetaxel administration is resumed after the dose delay, the criteria for dose reduction should be referred. Switching to paclitaxel during the study period (after informed consent and until completion of the final examination stipulated by the protocol [with the exception of the outcome investigation]) is not permitted.

1. No Grade ≥3 adverse events (i.e., adverse drug reactions) for which a causal relationship to docetaxel cannot be ruled out. Patients with controllable hypertension may continue docetaxel
2. Neutrophils ≥1500/mm^3
3. None of the following adverse events (i.e., adverse drug reactions) for which a causal relationship to docetaxel cannot be ruled out:
    - AST (GOT) or ALT (GPT) >3.0-fold ULN of the study site (or >5.0-fold the ULN of the study site in patients with liver metastases)
    - Total bilirubin >1.5-fold ULN of the study site
4. The investigator or subinvestigator judges that the dose of docetaxel may be administered to the subject.

7.1.2.3 Criteria for Administration of Paclitaxel

Before the start of each infusion, the subject must fulfill the following criteria to receive the dose on the basis of the latest data available at that time.

At the start of the treatment cycles, if the subject does not meet any of the following criteria, the planned dose of paclitaxel should be delayed. In the same cycle, if the subject does not meet any of the following criteria, the planned dose of paclitaxel will be skipped. If no worsening of clinical symptoms determined due to disease progression is observed, and continuation of study treatment is expected to lead clinical benefits, and the dose of paclitaxel can be continued safely without contraindications, treatment with paclitaxel may be continued. In this case, the subject's willingness to continue with paclitaxel treatment must be confirmed prior to continuation and must be documented in the medical record. In addition, it should be reported to the Sponsor as in a timely manner as possible. Imaging examinations to assess tumor response should be performed as per the schedule specified in the protocol, irrespective of the administration of paclitaxel. When paclitaxel administration is resumed after the dose delay or skip, the criteria for dose reduction should be referred. Switching to docetaxel during the study period (after informed consent and until completion of the final examination stipulated by the protocol [with the exception of the outcome investigation]) is not permitted.

1. No Grade ≥3 adverse events (i.e., adverse drug reactions) for which a causal relationship to paclitaxel cannot be ruled out. Patients with controllable hypertension may continue paclitaxel.
2. Neutrophils ≥1500/mm^3 (in the same cycle, neutrophils ≥1000/mm^3)

3. None of the following adverse events (i.e., adverse drug reactions) for which a causal relationship to paclitaxel cannot be ruled out:
   - AST (GOT) or ALT (GPT) >3.0-fold ULN of the study site (or >5.0-fold the ULN of the study site in patients with liver metastases)
   - Total bilirubin >1.5-fold ULN of the study site
4. The investigator or subinvestigator judges that the dose of paclitaxel may be administered to the subject.

7.1.3 Dosage Reduction Criteria

7.1.3.1 Dosage Reduction Criteria for ONO-4538

The study has no dosage reduction criteria for ONO-4538.

7.1.3.2 Dose Reduction Criteria for Docetaxel

The dose of docetaxel should be reduced in accordance with Table 7.1.3.2-1 based on the event listed in Table 7.1.3.2-2 and the other reasonable reason. After the dose reduction of docetaxel, safety measures such as frequent laboratory testing should be taken for the patient's safety. Subsequent treatment cycles should be continued at the same dose or a further reduced dose; the dose should not be increased.

Table 7.1.3.2-1 Dose Criteria for Docetaxel

Dose Level	Dose of Docetaxel
Initial dose	75 mg/m^2
Level 1 Reduction	60 mg/m^2
Level 2 Reduction	45 mg/m^2
Level 3 Reduction	To be discontinued

Table 7.1.3.2-2 Dose Reduction Criteria for Docetaxel

Adverse Event (related to docetaxel)	Docetaxel
Grade 4 neutropenia (<500/mm^3)	Reduce by 1 dose level
Grade ≥3 febrile neutropenia	Reduce by 1 dose level
Grade 4 thrombocytopenia (<25000/mm^3)	Reduce by 1 dose level
Grade ≥2 peripheral neuropathy	Reduce by 1 dose level
Grade ≥3 non-hematological toxicity other than the above (except for controllable hypertension)	Reduce by 1 dose level

7.1.3.3 Dose Reduction Criteria for Paclitaxel

The dose of paclitaxel should be reduced in accordance with Table 7.1.3.3-1 based on the event listed in Table 7.1.3.3-2 and the other reasonable reason. If the administration of paclitaxel is skipped for ≥2 consecutive times, the dose should be reduced by 1 dose level. Dose reduction by 1 dose level should also be considered if the administration is skipped for 1 time. After the dose reduction of paclitaxel, safety measures such as frequent laboratory testing should be taken for patient's safety. Subsequent cycles should be continued at the same dose or a further reduced dose; the dose should not be increased.

Table 7.1.3.3-1 Dose Criteria for Paclitaxel

Dose Level	**Dose of Paclitaxel**
Initial dose	100 mg/m^2
Level 1 Reduction	80 mg/m^2
Level 2 Reduction	60 mg/m^2
Level 3 Reduction	To be discontinued

Table 7.1.3.3-2 Dose Reduction Criteria for Paclitaxel

Adverse Event (related to paclitaxel)	**Paclitaxel**
Grade 4 neutropenia (<500/mm^3)	Reduce by 1 dose level
Grade ≥3 febrile neutropenia	Reduce by 1 dose level
Grade 4 thrombocytopenia (<25,000/mm^3)	Reduce by 1 dose level
Grade ≥2 peripheral neuropathy	Reduce by 1 dose level
Grade ≥3 non-hematological toxicity other than the above (except for controllable hypertension)	Reduce by 1 dose level

7.1.4 Discontinuation Criteria

If any of the following criteria are met during the treatment phase, the study treatment will be discontinued for the subject. The subject will undergo end-of-treatment examinations (at completion/discontinuation) and proceed to the follow-up phase.

7.1.4.1 Criteria for Discontinuation of ONO-4538

1. Has an overall response of PD as assessed by the investigator or subinvestigator according to the RECIST Guideline Ver. 1.1.

2. Apparent worsening of clinical symptoms determined to be due to disease progression.
3. Onset of Grade ≥2 interstitial lung disease with or without a causal relationship to ONO-4538.
4. Onset of Grade ≥2 eye disorder (eye pain or reduced visual acuity) which does not improve to Grade ≤1 with topical treatment and for which a causal relationship to ONO-4538 cannot be ruled out.
5. Onset of Grade ≥3 bronchospasm, diarrhea, colitis, neurologic toxicity, hypersensitivity reaction, infusion reaction (e.g., fever, chills, nausea, pain, headache, cough, itching, or rash) or uveitis for which a causal relationship to ONO-4538 cannot be ruled out.
6. Grade 3 drug-related thrombocytopenia > 7 days or associated with bleeding requires discontinuation.
7. Any drug-related liver function test (LFT) abnormality that meets the following criteria require discontinuation:
   - AST or ALT > 5-10x ULN for > 2 weeks
   - AST or ALT >10x ULN
   - Total bilirubin > 5x ULN
   - Concurrent AST or ALT > 3x ULN and total bilirubin > 2x ULN
8. Not received a dose of ONO-4538 within the past 6 weeks for specific reasons, such as the onset of an adverse event, unless study treatment is withheld for >6 weeks for steroid tapering, for which a consultation with the Sponsor is required prior to resuming study treatment.
9. The investigator or subinvestigator judges that the continuation of study treatment for the subject is inappropriate for other reasons from the viewpoint of efficacy or safety.

However, for subjects who meet either the above criterion 1 or 8 and who fulfill all of the following criteria in the opinion of the investigator or subinvestigator, treatment with ONO-4538 may be continued after consultation with the Sponsor. However, if continuing the study treatment is expected to be beneficial and ONO-4538 is promptly required, an after-the-fact report to the Sponsor is allowed. Per criterion 1 above, the subject's willingness to continue with the study treatment must be confirmed and a written re-consent must be obtained using the subject information and an informed consent form separately prepared for the re-consent.

If further PD is assessed in accordance with the RECIST Guideline Ver. 1.1 following the decision to continue treatment after the first PD, no additional continuation will be allowed and the study treatment must be discontinued.

1. No rapid disease progression and the continuation of study treatment is expected to lead to clinical benefits
2. ONO-4538 was tolerated
3. A stable ECOG Performance Status Score
4. Continuation of study treatment will not cause a delay of any prophylactic intervention for serious complications associated with disease progression (e.g., brain metastasis)

7.1.4.2 Discontinuation Criteria for Docetaxel or Paclitaxel

1. Despite dose reductions of docetaxel or paclitaxel being performed twice, a third round of dose reduction is required due to the onset of an adverse event
2. Has an overall response of PD as assessed by the investigator or subinvestigator according to the RECIST Guideline Ver. 1.1
3. Apparent worsening of clinical symptoms determined to be due to disease progression.
4. Onset of Grade ≥3 peripheral neuropathy with or without a causal relationship to docetaxel or paclitaxel
5. Onset of Grade ≥2 interstitial lung disease with or without a causal relationship to docetaxel or paclitaxel
6. Onset of the following adverse events (adverse drug reactions) for which a causal relationship to docetaxel or paclitaxel cannot be ruled out:
   - Grade ≥3 thrombocytopenia with hemorrhage
   - Increase in AST (GOT) or ALT (GPT) >5.0- to 10.0-fold ULN of the study site persisting for more than 2 weeks
   - Increase in AST (GOT) or ALT (GPT) >10.0-fold ULN of the study site
   - Increase in Total bilirubin >5.0-fold ULN of the study site
   - Increase in AST (GOT) or ALT (GPT) >3.0-fold ULN of the study site and an increase in total bilirubin >2.0-fold ULN of the study site
7. Not received docetaxel or paclitaxel within the past 6 weeks for specific reasons, such as the onset of an adverse event, unless study treatment is withheld for >6 weeks for steroid tapering, for which a consultation with the Sponsor is required prior to resuming study treatment
8. The investigator or subinvestigator judges that continuation of study treatment for the subject is inappropriate for other reasons from the viewpoint of efficacy and safety.

However, for subjects who meet the above criterion 7 and who fulfill all of the following criteria in the opinion of the investigator or subinvestigator, treatment with docetaxel or paclitaxel may be continued after consultation with the Sponsor. However, if continuing the study treatment is expected to be beneficial and docetaxel or paclitaxel is promptly required, an after-the-fact report to the Sponsor is allowed.

1. No rapid disease progression and the continuation of study treatment is expected to lead to clinical benefits
2. Docetaxel or paclitaxel were tolerated
3. A stable ECOG Performance Status Score
4. Continuation of the study treatment will not cause a delay of any prophylactic intervention for serious complications associated with disease progression (e.g., brain metastases)

7.2 Prior and Concomitant Therapies

7.2.1 Prohibited Therapies During the Study Period

7.2.1.1 Prohibited Therapies During the Study Period (ONO-4538 Group)

The following treatments are prohibited throughout the study period (i.e., after informed consent and until completion of the final examination stipulated by the protocol [with the exception of the outcome investigation]), unless absolutely necessary for medical reasons even after the examination at the end of the treatment phase (at discontinuation) has been completed.

1. Immunosuppressants and corticosteroids[1, 5]
2. Any antitumor therapies[5] (including chemotherapy, molecular target therapy, and immunotherapy[2])
3. Surgical treatment for malignant tumors[5]
4. (Chemo) radiotherapy[5]
5. Radiopharmaceuticals[3, 5]
6. Bisphosphonates and anti-RANKL antibody products[4]
7. Transplantation
8. Other unapproved drugs[5] (e.g., investigational use of drugs, unapproved combined formulations, and unapproved dosage forms)

[1] Corticosteroids may only be used topically (e.g., external, intra-articular, intranasal, ophthalmic, or inhalational use) or temporarily (e.g., for the treatment or prophylaxis of contrast medium allergy or AEs).

[2] These include local therapeutic agents such as picibanil.

[3] Use for examination or diagnostic use of radiopharmaceuticals is allowed.

[4] Treatment with bisphosphonates or anti-RANKL antibody products that has been ongoing since before the subject's first dose of the study treatment may be continued only if there are no changes made to the dosage and mode of administration.

[5] Before randomization, these therapies may be used unless they are performed in the respective periods specified in the exclusion criteria #18～24, 28.

<Rationale for Selection of the Prohibited Therapies>

The above mentioned therapies have been selected as prohibited therapies during the study period because they are likely to have a potential impact on the evaluation of safety and efficacy of the study.

7.2.1.2 Prohibited Therapies During the Study Period (Control Group)

The following treatments are prohibited throughout the study period (i.e., after informed consent and until completion of the final examination stipulated by the protocol [with the exception of the outcome investigation]), unless absolutely necessary for medical reasons even after the examination at the end of the treatment phase (at discontinuation) has been completed.

1. Immunosuppressants and corticosteroids[1, 5]
2. Any antitumor therapies[5] (including chemotherapy, molecular target therapy, and immunotherapy [2])
3. Surgical treatment for malignant tumors[5]
4. (Chemo) radiotherapy[5]
5. Radiopharmaceuticals[3, 5]
6. Bisphosphonates and anti-RANKL antibody products[4]
7. Transplantation
8. Disulfiram, cyanamide, carmofur, and procarbazine hydrochloride (for the paclitaxel group only)
9. Other unapproved drugs[5] (e.g., investigational use of drugs, unapproved combined formulations, and unapproved dosage forms)

[1] Corticosteroids may only be used topically (e.g., external, intra-articular, intranasal, ophthalmic, or inhalational use) or temporarily (e.g., for the treatment or prophylaxis of contrast medium allergy or adverse events).

[2] These include local therapeutic agents such as picibanil.

[3] Use for examination or diagnostic use of radiopharmaceuticals is allowed.

[4] Treatment with bisphosphonates or anti-RANKL antibody products that has been ongoing since before the subject's first dose of the study treatment may be continued only if there are no changes made to the dosage and mode of administration.

[5] Before randomization, these therapies may be used unless they are performed in the respective periods specified in the exclusion criteria #18～24, 28.

<Rationale for Selection of the Prohibited Therapies>

The above mentioned therapies have been selected as prohibited therapies during the study period because they are likely to have a potential impact on the evaluation of safety and efficacy of the study.

7.2.2 Prophylactic Premedication

7.2.2.1 ONO-4538 Group

In subjects who may experience infusion-related reactions to ONO-4538, prophylactic premedication with acetaminophen or diphenhydramine is recommended before administration of ONO-4538 for subjects who have experienced infusion reaction previously.

7.2.2.2 Docetaxel Group

Follow recommendations per local SmPC/Package Insert.

7.2.2.3 Paclitaxel Group

The following pretreatments are recommended to prevent the occurrence of serious hypersensitivity symptoms caused by paclitaxel.

The following premedication should be completed approximately ≥30 minutes prior to administration of paclitaxel: intravenous doses of dexamethasone sodium phosphate injection (dexamethasone 8 mg), either ranitidine hydrochloride injection (ranitidine 50 mg) or famotidine injection (famotidine 20 mg), and oral diphenhydramine hydrochloride tablet (diphenhydramine hydrochloride 50 mg).

The initial dose of dexamethasone is 8 mg. If hypersensitivity symptoms do not occur until the next administration or are not clinically significant, the dexamethasone dose may be reduced to half the

initial dose (4 mg) from Week 2. If hypersensitivity symptoms are absent or clinically insignificant in the subsequent weeks of administration, the dexamethasone dose may be further reduced by half, down to 1 mg.

7.3 Discontinuation of the Supply of the Investigational Product

Subjects will be supplied with the investigational product until they meet the criteria in Section 7.1.4 "Discontinuation Criteria" or in Section 13.1 "Rules or Criteria for Withdrawing Individual Subjects from the Study." However, the sponsor will discontinue the supply of the investigational product in the following instances.

1. When the regulatory authority of each country has rejected the approval application for ONO-4538 for the indication of unresectable advanced or recurrent esophageal cancer
2. When the entire study is discontinued
3. When the subject is able to receive ONO-4538 for the same indication as this study under public or private health insurance coverage
4. When the subject has become able to receive alternative treatment

7.4 Post-Study Access to Therapy

At the conclusion of the study, subjects who continue to demonstrate clinical benefit will be eligible for sponsor supplied study drug. Study drug will be provided via an extension of the study, a rollover study requiring approval by responsible health authority and ethics committee or through another mechanism at the discretion of sponsor as per local health authority requirements.

8 Study schedule and Observation Items

8.1 Study Schedule

The entire study period will consist of three phases: screening phase, treatment phase, and follow-up phase. A schema of this study is shown in Figure 2-1 in the protocol synopsis section.

8.1.1 Screening Phase

The screening phase will start when the subject has signed the informed consent form. Subjects who have provided consent will be enrolled in the study. Among the enrolled subjects, the investigator or subinvestigator will proceed with randomization of those fulfilling the criteria in Section 4.2 "Inclusion Criteria" and not meeting any of the criteria in Section 4.3 "Exclusion Criteria" judged to

be appropriate as subjects in this study. Protocol-specified evaluations and their schedule during this phase are specified in Table 6-1, Table 6-2, and Table 6-3 (ONO-4538, docetaxel, and paclitaxel groups, respectively) in the protocol synopsis section.

8.1.2 Treatment Phase

8.1.2.1 ONO-4538 Group

During the treatment phase, ONO-4538 will be administered every 2 weeks up to 3 doses. These 6 weeks will count as one cycle. Study treatment may be repeated in subjects fulfilling the criteria in Section 7.1.2.1 "Criteria for Administration of ONO-4538" and not meeting any of the criteria in Section 7.1.4.1 "Criteria for Discontinuation of ONO-4538." Protocol-specified evaluations and their schedule during this phase are specified in Table 6-1 in the protocol synopsis section.

8.1.2.2 Docetaxel Group

During the treatment phase, docetaxel will be administered every 3 weeks. This is defined as 1 cycle. Study treatment may be repeated in subjects fulfilling the criteria in Section 7.1.2.2 "Criteria for Administration of Docetaxel" and not meeting any of the criteria in Section 7.1.4.2 "Discontinuation Criteria for Docetaxel or Paclitaxel." Protocol-specified evaluations and their schedule during this phase are specified in Table 6-2 in the protocol synopsis section.

8.1.2.3 Paclitaxel Group

During the treatment phase, paclitaxel will be administered every 1 week up to 6 doses, followed by a 2-week rest period. This is defined as 1 cycle. Study treatment may be repeated in subjects fulfilling the criteria in Section 7.1.2.3 "Criteria for Administration of Paclitaxel" and not meeting any of the criteria in Section 7.1.4.2 "Discontinuation Criteria for Docetaxel or Paclitaxel." Protocol-specified evaluations and their schedule during this phase are specified in Table 6-3 in the protocol synopsis section.

8.1.3 Follow-up Phase

All subjects meeting any of the criteria in Section 7.1.4 "Discontinuation Criteria" after administration of the investigational product will undergo end-of-treatment examination (at discontinuation) and proceed to the follow-up phase. Protocol-specified evaluations and their schedule during the follow-up phase are specified in Table 6-1, Table 6-2, and Table 6-3 (ONO-4538 group, docetaxel group, and paclitaxel group, respectively) in the protocol synopsis section.

8.2 Observation Items

8.2.1 Patient Demographics and Other Background Characteristics

During the screening phase, the investigator or designee will collect and enter the following data into the eCRF. The TNM Classification of Malignant Tumours, Seventh Edition, published by the Union for International Cancer Control (UICC) (hereinafter referred to as "the TNM classification") will be used for determination of the lesion sites (anatomical subsites), clinical classification, and disease stage (see Appendix 1).

- Subject identification code (ID)
- Date of written informed consent
- Sex
- Date of birth
- Race
- Ethnicity
- Height
- Weight
- Medical history (i.e., medical history within 1 year before randomization and other clinically relevant medical history)
- Concurrent diseases
- History of alcohol consumption
- History of smoking
- Performance Status Score (ECOG)
- Date of diagnosis of the primary disease
- Lesion sites (anatomical subsites according to the TNM classification) Note)
- Clinical classification (TNM clinical classification) Note)
- Disease stage (Stage grouping: TNM classification) Note)
- Histological classification Note)
- Number of organs with metastases Note)
- Date of surgery (including EMR or ESD) for the primary lesion (for subjects with recurrent esophageal cancer)
- Residual tumor (R) classification (for subjects with recurrent esophageal cancer)
- Site of recurrence (for subjects with recurrent esophageal cancer)
- Diameter of the target lesions Note)

- Past treatments for cancer (e.g., surgery, radiotherapy, pharmacotherapy)

Note): Based on the results at randomization

8.2.2 Details of Study Treatment and Concomitant Treatments

8.2.2.1 Details of administration of the investigational product

The investigator or designee will enter details of study treatment (i.e., dose, infusion start time, infusion end time) into the eCRF.

8.2.2.2 Details of Administration of Other Drugs and Concomitant Therapies

For any drugs and therapies used after the start of study treatment and before completion of the final examination stipulated by the protocol (with the exception of the outcome investigation), the investigator or designee will enter details of administration (i.e., name of treatment, therapy dates, route of administration, reason for use and, if required, dosage and mode of administration) into the eCRF.

During the follow-up investigation, an outcome investigation will be performed, and an investigation will also be performed on any post-study treatment for esophageal cancer, with its starting date and details.

In subjects who, at the time of initiation of the follow-up investigation, have any adverse event for which a causal relationship to the investigational product cannot be ruled out or which led to discontinuation, treatment for the adverse event will be continued at appropriate intervals until the event resolves, improves, or stabilizes and thus is judged not to require further follow-up.

Data from investigation of the details of any post-study treatment for esophageal cancer and concomitant therapies will be entered into the eCRF.

Details of follow-up investigation are described in Section 8.2.10 "Follow-up Investigation."

8.2.3 Items related to Efficacy

The investigator or designee will perform the following measurements, examinations, and investigations at predetermined time points. For subjects who have discontinued study treatment, the subject's safety will be prioritized, and measurements, examinations, and investigations will be performed as far as possible.

8.2.3.1 Imaging Examination (with Measurements of Tumor Diameters and Assessment of Tumor Response)

Neck-thoraco-abdomino-pelvic CT or other imaging examinations will be performed and the tumor response will be assessed. For the assessment of tumor response, in principle, the same method should be used throughout the study. If brain or bone metastasis is clinically suspected during the screening phase, a head CT/MRI or FDG-PET (or bone scintigraphy) must be performed to determine the presence of the brain or bone metastasis. Also after the start of study treatment, whenever brain or bone metastasis is suggested by clinical symptoms, an imaging examination should be performed to determine the presence or absence of brain or bone metastasis.

In accordance with the RECIST Guideline Ver. 1.1 (see Appendix 2), the diameters of target lesions will be measured and the tumor response will be assessed, and the results of both the measurement and assessment will be entered into the eCRF. If previous data were obtained within the allowable window for the end-of-treatment examination (at completion/discontinuation) or the examination 28 days after the end of the treatment phase, the previous data may be used as an alternative, unless new data are medically required. However, data will be newly obtained if 15 days have passed for imaging examination. Tests will be performed whenever medically indicated.

Imaging examinations (with measurements of tumor diameters and assessment of tumor response) will be performed at the time points specified below. The decision to continue the study treatment will be made by the investigator or subinvestigator based on the imaging assessment in accordance with the RECIST Guideline Ver. 1.1.

Time points of evaluation

Screening phase

Treatment phase: Every 6 weeks starting from the start of Cycle 1, every 12 weeks after a period of 1 year from the start of Cycle 1, and at the end (discontinuation) of the treatment phase[Note 1)]

Follow-up phase: Day 28 post-treatment[Note 1)] and during the follow-up investigation (subjects who have a response assessment other than PD in accordance with the RECIST Guideline Ver. 1.1 [see Appendix 2] and have discontinued the treatment phase for safety reasons should undergo imaging examinations as far as possible until either initiation of post-study treatment for esophageal cancer, or a diagnosis of PD or recurrence.)

Note 1): The previous data may be used as an alternative, unless new data are medically required. However, data will be newly obtained if 15 days have passed for imaging examination.
Tests will be performed whenever medically indicated.

8.2.4 Items related to Safety

For the items specified in the following subsections 8.2.4.1 to 8.2.4.7, the investigator or designee will perform measurements, examinations, and investigations at predetermined time points. For subjects who have discontinued study treatment, the subject's safety will be prioritized, and measurements, examinations, and investigations will be performed as far as possible.

8.2.4.1 Vital Signs, Body Weight, and 12-lead ECG

8.2.4.1.1 Vital Signs (Systolic and Diastolic Blood Pressure, Pulse Rate, and Body Temperature), and Body Weight

Vital signs and body weight will be measured at the time points specified below. Resting systolic and diastolic blood pressure and pulse rate will be measured in principle using the same body position throughout the study period.

Time points of evaluation

Screening phase (the allowance window starts from the day of the first dose, ranging between -7 and -1.)

Treatment phase: [ONO-4538 group]: Before dosing on Day 1 of each cycle[Note 1)], after dosing on Day 1 (measured in Cycle 1 only, except for body weight), Day 8 (measured in Cycle 1 only, except for body weight), before dosing on Day 15 (except for body weight), before dosing on Day 29 (except for body weight), Day 43, and at the end (or discontinuation) of the treatment phase[Note 2)]

[Docetaxel group]: Before dosing on Day 1 in each cycle[Note 1)], after dosing on Day 1 (measured in Cycle 1 only, except for body weight), Day 8 (measured in Cycle 1 only, except for body weight), Day 22, and at the end (or discontinuation) of the treatment phase[Note 2)]

[Paclitaxel group]: Before dosing on Day 1 in each Cycle[Note 1)], after dosing on Day 1 (measured in Cycle 1 only, except for body weight), before dosing on Day 8 (except for body weight and pulse rate measurement [Cycle 2 and beyond for pulse rate]), before dosing on Days 15, 22, 29, and 36 (except for body weight and

pulse rate measurement), Day 50, and at the end (or discontinuation) of the treatment phase[Note 2)]

Follow-up phase: At the examination 28 days after the end of the treatment phase [Note 2)]

[Note 1)]: If a dose of the investigational product is administered within the allowed window for the last day of the previous cycle (Day 43 in the ONO-4538 group, Day 22 in the docetaxel group, and Day 50 in the paclitaxel group), the measurements obtained on the last day of the previous cycle (screening phase for Cycle 1) may be used unless a new test is deemed medically indicated. However, vital signs will be newly measured if ≥2 days have passed since the previous measurement, and body weight will be newly measured if ≥8 days have passed since the previous measurement. Measurements will be performed whenever medically indicated.

[Note 2)]: The previous measurements may be used as an alternative, unless new measurements are medically required. However, vital signs will be newly measured if ≥2 days have passed since the previous measurement, and body weight will be newly measured if ≥8 days have passed since the previous measurement. Measurements will be performed whenever medically indicated.

8.2.4.1.2 12-lead ECG (Heart Rate [HR], PR Interval, RR Interval, QRS Width and QT Interval)

At the time point specified below, resting 12-lead ECG will be performed using the electrocardiograph lent by the Sponsor, with measurements of HR, PR interval, RR interval, QRS width, and QT interval. The ECG data will be forwarded to the core ECG analysis laboratory. The written procedure prepared separately will be followed for measurements of 12-lead ECG parameters and data transmission to the core ECG analysis laboratory. After a period of one year from the start of Cycle 1, resting ECG will be performed using the electrocardiograph at the study site, with measurements of HR, PR interval, RR interval, QRS width, and QT interval. Data transmission to the core ECG analysis laboratory is not necessary. Immediately after measurements of 12-lead ECG parameters and after receipt of ECG analysis results from the core ECG analysis laboratory, the investigator or subinvestigator will check the results for any abnormal findings.

Time points of evaluation

Screening phase [Note 1)] (or before dosing on Day 1 in Cycle 1)

Treatment phase: [ONO-4538 group]: After dosing on Day 1 and Day 43 in Cycle 1, before dosing on Day 1 and Day 43 in Cycle 4, and at the end (or discontinuation) of the treatment phase[Note 2)]

[Docetaxel group]: After dosing on Day 1 and Day 22 in Cycle 1, before dosing on Day 1 and Day 22 in Cycle 7, and at the end (or discontinuation) of the treatment phase[Note 2)]

[Paclitaxel group]: After dosing on Day 1 and Day 50 in Cycle 1, before dosing on Day 1 and Day 50 in Cycle 3, and at the end (or discontinuation) of the treatment phase[Note 2)]

Follow-up phase: At the examination 28 days after the end of the treatment phase [Note 2)]

[Note 1)]: The allowance window starts from the day of the first dose, ranging between -7 and -1.

[Note 2)]: The previous ECG data may be used as an alternative unless new ECG data are medically required. However, an ECG will be newly performed if ≥8 days have passed since the previous ECG. ECGs will be performed whenever medically indicated.

8.2.4.2 Chest X-ray

Chest X-ray will be performed at the time points specified below. The investigator or subinvestigator will check if there are any abnormal findings. As necessary, unscheduled X-rays should be performed whenever onset of respiratory disease is suspected based on clinical symptoms or findings, or laboratory findings during the study period (i.e., after informed consent and until completion of the final examination stipulated by the protocol [with the exception of the outcome investigation]).

Time points of evaluation

Screening phase [Note 1)]

Treatment phase: The last day of each cycle (Day 43 in the ONO-4538 group, Day 22 in the docetaxel group, Day 50 in the paclitaxel group), and at the end (or discontinuation) of the treatment phase[Note 2)]

Follow-up phase: At the examination 28 days after the end of the treatment phase [Note 2)]

[Note 1)]: The allowance window starts from the day of the first dose, ranging between -7 and -1.

[Note 2)]: The previous chest X-ray may be used as an alternative unless a new chest X-ray is medically required. However, chest X-ray will be newly performed if ≥8 days have passed since the previous chest X-ray. Chest X-rays will be performed whenever medically indicated.

8.2.4.3 Laboratory Tests

During the treatment phase and the follow-up phase, the investigator or subinvestigator will perform hematology tests, biochemistry tests, urinalysis, immunological tests and hormone tests as specified in the following subsections 8.2.4.3.1 and 8.2.4.3.2, and assess each test result to determine whether it is normal or abnormal.

8.2.4.3.1 Hematology Tests, Biochemistry Tests, and Urinalysis

Hematology tests, biochemistry tests, and urinalysis will be performed at the time points specified below. In principle, these tests will be performed locally at individual study sites.

Time points of evaluation

Screening phase (the allowance window starts from the day of the first dose, ranging between -7 and -1.)

Treatment phase: [ONO-4538 group]: Day 8, before dosing on Day 15, before dosing on Day 29 and Day 43 in Cycle 1; before dosing on Day 1 [Note 1)] and Day 43 in Cycle 2 and beyond, and at the end (or discontinuation) of the treatment phase[Note 2)]

[Docetaxel group]: On Days 8 and 22 in Cycle 1; before dosing on Day 1[Note 1)] and Day 22 in and after Cycle 2, and at the end (or discontinuation) of the treatment phase[Note 2)]

[Paclitaxel group]: Before dosing on Days 8, 15, 22, 29, and 36 [Note 3)], and Day 50 in Cycle 1; before dosing on Days 1 [Note 1)], 8, 15, 22, 29, and 36 [Note 3], and Day 50 in and after Cycle 2, and at the end (or discontinuation) of the treatment phase[Note 2)]

Follow-up phase: At the examination 28 days after the end of the treatment phase [Note 2)]

[Note 1)]: If the investigational product is administered within the allowed window for the last day of the previous cycle (Day 43 in the ONO-4538 group, Day 22 in the docetaxel group, Day 50 in the paclitaxel group), the measurements obtained on the last day of the previous cycle may be used unless a new test is deemed medically indicated. However, tests will be newly performed if ≥8 days have passed since the previous tests. Tests will be performed whenever medically indicated.

Note 2): The previous results may be used as an alternative, unless new results are medically required. However, tests will be newly performed if ≥8 days have passed since the previous tests. Tests will be performed whenever medically indicated.

Note 3): For the paclitaxel group, the criteria for administration, dose reduction and discontinuation should be confirmed with the results of hemoglobin, white blood cells, neutrophils, and platelets of hematological parameters and albumin, AST, ALT, ALP, total bilirubin, Na, K, Ca, creatinine, and CRP of biochemical parameters.

Test parameters

Hematology:	Red blood cell count, MCV, MCH, MCHC, hemoglobin, hematocrit, white blood cell count, white blood cell differential count (neutrophils, lymphocytes, eosinophils, basophils, monocytes), platelet count
Biochemistry:	Albumin, ALP, AST (GOT), ALT (GPT), total bilirubin, direct bilirubin, gamma-GTP, total protein, creatinine, blood glucose, LDH, BUN, uric acid, CK (CPK), P, Ca, Na, K, Cl, C-reactive protein (CRP)
Urinalysis:	Specific gravity, protein, glucose, occult blood, sediment (white blood cells, red blood cells)

8.2.4.3.2 Immunological and Hormone Tests

Immunological and hormone tests will be performed at the time points specified below. Blood specimens will be collected in the morning as far as possible, and the time of blood collection will be recorded for each day of testing. These tests will be performed at the designated central laboratory. Measurements of SP-D and KL-6 at the time points specified below will be mandatory, but SP-D and KL-6 may be measured as necessary at other time points than these specified below.

Time points of evaluation

Screening phase[Note 1)]

Treatment phase: The last day of each cycle (Day 43 in the ONO-4538 group, Day 22 in the docetaxel group, Day 50 in the paclitaxel group) and at the end (or discontinuation) of the treatment phase[Note 2]

Follow-up phase: At the examination 28 days after the end of the treatment phase [Note 2)]

Note 1): Latest results obtained within 14 days before randomization may be used as an alternative.

Note 2): The previous results may be used as an alternative, unless new results are medically required. However, tests will be newly performed if ≥8 days have passed since the previous tests. Tests will be performed whenever medically indicated.

Test parameters (volume of blood collection at a time: 9 mL)

Immunological tests: Rheumatoid factor (RA), antinuclear antibody (ANA), surfactant protein D (SP-D), Krebs von den Lungen-6 (KL-6)

Hormone tests: Thyroid-stimulating hormone (TSH), free triiodothyronine (free T3), free thyroxine (free T4)

If a clinically meaningful change in RA, ANA, TSH, or free T4 is shown as a result of these measurements, the following "Additional test parameters" will be measured as soon as possible. Other additional test parameters, such as adrenocorticotropic hormone (ACTH; volume of blood collection at a time, 2 mL), should also be performed as appropriate.

Additional test parameters (volume of blood collection at a time: 15 mL)

Anti-DNA antibody, anti-SSA antibody (Ro), anti-SSB antibody (La), anti-thyroglobulin antibody, anti-LKM antibody, anti-phospholipid antibody, anti-GAD antibody, anti-neutrophil cytoplasmic antibody, beta-1C/beta-1A globulin (C3), beta-1E globulin (C4), 50% hemolytic complement (CH50) activity in serum, TSH stimulating receptor antibody, anti-thyroid peroxidase antibody

8.2.4.4 Pregnancy Testing (Only in Women of Childbearing Potential)

Serum or urine pregnancy testing will be performed in women of childbearing potential (including women with chemical menopause or no menstruation for other medical reasons).[Note)] The same testing method should be used throughout the study as far as possible. In principle, pregnancy testing will be performed locally at individual study sites.

Note): Women of childbearing potential are defined as all women after onset of menstruation who are not post-menopausal and have not been surgically sterilized (e.g., hysterectomy, bilateral tubal ligation, bilateral oophorectomy). Postmenopause is defined as amenorrhea for ≥12 consecutive months without specific reasons. Women using contraceptive methods, i.e., oral contraceptives or other hormonal contraceptives (vaginal contraceptives, injectable

contraceptives, etc.) or mechanical contraception (intrauterine devices, contraceptive barriers), are regarded as having childbearing potential.

Time points of evaluation

Screening phase (the allowance window starts from the day of the first dose, ranging between -7 and -1.)

Treatment phase: Before dosing on Day 1 in each cycle[Note 1)] and at the end (or discontinuation) of the treatment phase[Note 2)]

Follow-up phase: At the examination 28 days after the end of the treatment phase [Note 2)]

[Note 1)]: For Cycle 2 and beyond, pregnancy testing may be performed within 7 days before administration of the investigational product.

[Note 2)]: The previous result may be used as an alternative, unless a new result is medically required. However, testing will be newly performed if ≥8 days have passed since the previous test. Pregnancy testing will be performed whenever medically indicated.

8.2.4.5 Viral Tests

In the screening phase, viral tests will be performed on the following parameters. If these tests have been performed within 1 year before randomization, the available results may be used as an alternative, unless new results are medically required. HBV-DNA quantification assay will be additionally performed in subjects with a negative HBs antigen test but with a positive test result for either HBs antibody or HBc antibody. In principle, these tests will be performed locally at individual study sites.

Test parameters

HIV-1 antibody, HIV-2 antibody, HTLV-1 antibody, HBs antigen, HBs antibody, HBc antibody, HCV antibody

8.2.4.6 ECOG Performance Status Score

ECOG Performance Status (PS) Score will be assessed at the time points specified below.

Time points of evaluation

Screening phase (the allowance window starts from the day of the first dose, ranging between -7 and -1.)

Treatment phase: [ONO-4538 group]: Before dosing on Day 1 in each cycle[Note 1)] (Cycle 2 and beyond), on Day 8 (in Cycle 1 only), before dosing on Day 15, before dosing on Day 29, on Day 43, and at the end (or discontinuation) of the treatment phase[Note 2)]

[Docetaxel group]: Before dosing on Day 1 in each cycle[Note 1)] (Cycle 2 and beyond), Day 8 (in Cycle 1 only), Day 22, and at the end (or discontinuation) of the treatment phase[Note 2)]

[Paclitaxel group]: Before dosing on Day 1 in each cycle[Note 1)] (Cycle 2 and beyond), before dosing on Day 8 (in Cycle 1 only), Day 50, and at the end (or discontinuation) of the treatment phase[Note 2)]

Follow-up phase: At the examination 28 days after the end of the treatment phase [Note 2)]

[Note 1)]: If the investigational product is administered within the allowed window for the last day of the previous cycle (Day 43 in the ONO-4538 group, Day 22 in the docetaxel group, Day 50 in the paclitaxel group), the measurements obtained on the last day of the previous cycle may be used unless a new test is deemed medically indicated. If more than 8 days have passed since the previous test, a new test should be performed. Also, the test should be performed whenever it is deemed medically indicated.

[Note 2)]: The previous PS score may be used as an alternative, unless a new assessment is medically required. However, PS will be newly assessed if ≥8 days have passed since the previous assessment. PS assessment will be performed whenever medically indicated.

8.2.4.7 Adverse Events

All adverse events will be assessed according to the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI-CTCAE), Version 4.0. For signs or symptoms not listed in the CTCAE, the investigator or subinvestigator will use his/her judgment in determining whether a given sign/symptom should be handled as an adverse event.

For each adverse event occurring after the start of the initial dosing of the investigational product up to the time of completion of the final examination stipulated by the protocol (with the exception of the outcome investigation), the following data will be entered into the eCRF: event term, date of onset, grade, decision made on whether to continue the investigational product, any treatment for the adverse event, seriousness, causal relationship to the investigational product, and event outcome at the time of final assessment (or time of resolution if the event resolved before the final assessment). Immune-

mediated adverse events (IMAEs) and adverse events likely to be IMAEs (see Table 9.2-1) and any treatment for IMAEs are to be collected during the treatment period and for 100 days following the last dose of investigational product. For subjects proceeding to the follow-up phase, any adverse events for which a causal relationship to the investigational product cannot be ruled out (i.e., adverse drug reactions) or which led to termination of the treatment phase should be followed up and appropriately treated, with the observation/measurement of clinical symptoms, general laboratory tests, or other parameters, until no further follow-up is deemed necessary for reasons of recovery, improvement, stabilization, or for any other reason. Serious adverse events should be reported to the sponsor (or designee) according to Section 9.9 "Actions to be Taken Following Onset of Any Serious Adverse Event."

If the investigator or subinvestigator concludes that if an adverse event of subjective symptoms/objective findings is related to an abnormal change in a laboratory test value, the event is in principle to be reported as a single adverse event (adverse event of subjective symptoms/objective findings only). If these events are considered to be mutually unrelated events, the subjective symptom/objective finding and the abnormal laboratory change are to be reported as separate independent adverse events.

8.2.5 Serum Drug Concentration Measurements (ONO-4538 group only)

Serum ONO-4538 concentrations will be measured using an electrochemiluminescence assay by Pharmaceutical Product Development, LLC (PPD). Detailed measurement methods will be specified in the assay protocol prepared by the person responsible for serum ONO-4538 concentration measurements. Results of measurements will be reported in the serum ONO-4538 concentration measurement report. The specimens will be retrieved and transported in accordance with a written procedure prepared separately.

Blood specimens for serum drug concentration measurements will be collected at the time points specified below (volume of blood collection at a time: 2 mL). As far as possible, blood will also be collected 6-12 weeks after the last dose of the investigational product.

For the purpose of technical exploration, ONO-4538 and anti-ONO-4538 antibody may be measured in an exploratory manner using specimens for measuring serum drug concentration, but the results will not be reported.

Time points of evaluation

Treatment phase: Before dosing on Days 1 and 29 in Cycle 1; and before dosing on Day 1 in Cycles 4 and 9

Follow-up phase: At the examination 28 days after the end of the treatment phase (only in subjects proceeding to the follow-up phase by the end of Cycle 9), and 6–12 weeks after the last dose of the investigational product (as far as possible)

8.2.6 Anti-drug Antibody Measurements (ONO-4538 group only)

Anti-ONO-4538 antibody will be measured using an electrochemiluminescence assay by Pharmaceutical Product Development, LLC (PPD). As necessary, neutralizing activity of antibody will be assessed by BMS. Detailed measurement methods will be specified in the assay protocol prepared by the person responsible for anti-ONO-4538 antibody measurements. The specimens will be retrieved and transported in accordance with a separate written procedure.

Blood specimens for anti-ONO-4538 antibody measurements will be collected at the time points specified below (volume of blood collection at a time: 2 mL). As far as possible, blood will also be collected 6–12 weeks after the last dose of ONO-4538.

If the volume of the specimen for anti-ONO-4538 antibody measurement is insufficient, the specimen collected at the same time for serum drug concentration measurement may be used. For the purpose of technical exploration, serum ONO-4538 and anti-ONO-4538 antibodies may be measured in an exploratory manner using samples for measuring serum anti-ONO-4538 antibodies, but these results will not be reported.

Time points of evaluation

Treatment phase: Before dosing on Days 1 and 29 in Cycle 1; and before dosing on Day 1 in Cycles 4 and 9

Follow-up phase: At the examination 28 days after the end of the treatment phase (only in subjects proceeding to the follow-up phase by the end of Cycle 9), and 6–12 weeks after the last dose of ONO-4538 (as far as possible)

8.2.7 Tumor Markers (As needed)

In subjects with tumor marker levels above the upper limit of the normal range, measurements of the tumor marker will be continued as far as possible, and as frequently as needed. In principle, tumor marker measurements will be performed locally at individual study sites.

8.2.8 Patient Reported Outcomes (PROs), Healthcare Resource Utilization

The evaluation of health-related quality of life is an increasingly important aspect of a clinical efficacy in oncology trials. Quality of life data provide an understanding of the impact of treatment from a patient's perspective and offer insights into the patient experience that may not be captured through physician reporting.

General health status will be measured using the EQ-5D. The EQ-5D is a standardized instrument for use as a measure of self-reported health status. The EQ-5D comprises 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) and a visual analog rating scale (VAS). The utility data generated from the EQ-5D is recommended for and commonly used in cost effectiveness analysis.

Healthcare resource utilization data will be collected for all randomized subjects, including information about medical care encounters such as hospital admissions and their duration, outpatient visits, diagnostic tests and procedures, concomitant medications, and reasons for encounters. Data will be used for subsequent cost-effectiveness analyses.

Time points of evaluation

Screening phase (the allowance window starts from the day of the first dose, ranging between -7 and -1.)

Treatment phase: Every 6 weeks starting from the start of Cycle 1 and at the end (discontinuation) of the treatment phase[Note 1)]

Follow-up phase: Every 12 weeks

[Note 1)]: The previous data may be used as an alternative, unless new data are medically required. However, data will be newly obtained if 15 days have passed for the previous evaluation. Tests will be performed whenever medically indicated.

8.2.9 Exploratory Biomarkers

[redacted]

[redacted]

[redacted]

[redacted] [redacted]

[redacted]

[redacted]

[redacted]

[redacted] [redacted]

[redacted]

[redacted]

[redacted] [redacted]

[redacted]

[redacted]

[redacted] [redacted]

[redacted]

[redacted]

8.2.9.6 Tumor Tissue Examinations (some are essential)

Tumor tissues must be obtained from each subject for analysis of PD-L1 expression during the screening phase to ensure the selection of subjects who meet all of the eligibility criteria (specified inclusion criteria). The collected tumor tissue specimens should be sent to the central laboratory before randomization. Collection of tumor tissues for biomarker analysis (expression analysis of the exploratory biomarkers) during the screening phase and during the follow-up phase is optional. For subjects in whom sampling of tumor tissue during the screening phase is possible and safe in the investigator's opinion, newly collected specimens will be provided to the central laboratory. If subjects are unable to undergo another tissue biopsy, the stored specimens can be used. During the follow-up phase, tumor tissues will be collected after the end of efficacy evaluation.

Biopsy specimens collected from bone lesions are not suitable for any of the analyses planned for this study because decalcification of specimens compromises some of the planned biomarker analyses.

Details of collection of tumor specimens

For evaluation of PD-L1 expression, at least one formalin-fixed paraffin-embedded (FFPE) tumor-tissue block (recommended) or at least five FFPE unstained slides are required. Tissue specimens

must contain at least 100 evaluable cancerous cells. PD-L1-stained tissue specimens will be analyzed by a pathologist at the sponsor-designated central laboratory. Specimens with insufficient cancerous cells are **not evaluable** and such subjects will not be randomized in the study. The subject will test **positive** for PD-L1 when cancerous cells with stained cell membranes account for ≥1% of at least 100 evaluable cancerous cells and will test **negative** for PD-L1 when such cells account for <1%. If specimens meet the evaluation criteria, but assessment of the staining state of the cell membranes is difficult due to the biological conditions of the specimens (not due to their inappropriate preparation or handling), the specimens are classified as **Indeterminate**. These reasons include, but are not limited to: (1) a high melanin content in cells and (2) intensely stained cytoplasm.

[redacted]

Procedures for the handling, storage and analysis of the specimens and provision of the genetic test results to patients will be in accordance with a separate written procedural manual.

8.2.10 Follow-up Investigation

For subjects who have an overall response assessment other than PD and have discontinued the treatment phase for safety reasons in accordance with the RECIST Guideline Ver. 1.1 (see Appendix 2), imaging examinations will be continued as far as possible, until either initiation of the post-study treatment for esophageal cancer or assessment of PD or recurrence. The obtained imaging findings will be assessed to determine the tumor response status according to the RECIST Guideline Ver. 1.1 (see Appendix 2) and the results will be entered into the eCRF.

For subjects who, at the time of initiation of the follow-up investigation, develop an adverse event for which a causal relationship to the investigational product cannot be ruled out or that led to discontinuation, the investigation of the adverse event and treatment for the event will be continued at appropriate intervals until the event resolves, improves, or stabilizes, etc. and thus is judged not to require further follow-up. The results of the follow-up investigation will be entered into the eCRF.

The outcome investigation will be part of the follow-up investigation and will be conducted for all subjects enrolled in this study who have completed the examination 28 days after the end of the

treatment phase (if the patient died, data on the date and cause of death will also be collected). The outcome investigation (with collection of data on the date and cause of death if the patient died) should be conducted by direct contact or other means of communication, such as telephone or letter, roughly every 6 to 8 weeks but as frequently as required by the occurrence of events. Updated information on the survival status will be entered into the eCRF, as appropriate. The follow-up investigation will also collect data on any post-study treatment for esophageal cancer, its start date, and its details as far as possible.

8.2.11 Lost to Follow-up

All reasonable efforts must be made to locate subjects to determine and report their ongoing status. This includes follow-up with persons previously authorized by subject to provide this information. Lost to follow-up is defined by the inability to reach the subject after a minimum of three documented phone calls, faxes, or emails as well as lack of response by subject to one registered mail letter. All attempts should be documented in the subject's medical records. If it is determined that the subject has died, the site will use permissible local methods to obtain the date and cause of death.

If investigator's use of third-party representative to assist in the follow-up portion of the study has been included in the subject's informed consent, then the investigator may use a Sponsor-retained third-party representative to assist site staff with obtaining subject's contact information or other public vital status data necessary to complete the follow-up portion of the study. The site staff and representative will consult publicly available sources, such as public health registries and databases, in order to obtain updated contact information. If after all attempts, the subject remains lost to follow-up, then the last known alive date as determined by the investigator should be reported and documented in the subject's medical records.

9 Adverse Events and Matters to Ensure the Safety of Subjects

9.1 Definition of Adverse Events

An adverse event is any unfavorable or unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of the investigational product, regardless of whether the event is causally related to the investigational product. Worsening of the primary disease, of a symptom associated with the primary disease, or of a concurrent disease will be regarded as an adverse event if the worsening is medically judged to be beyond the predictable range of the natural course of the disease. Also, tumor enlargement or appearance of a new lesion (except for a malignant tumor histologically different from the primary lesion) after initiation of the study will be regarded as

an adverse event if the enlargement/appearance is medically judged to be beyond the predictable range of the natural course of the disease.

9.2 Definition of Immune-mediated Adverse Events

IMAEs include events, regardless of causality, occurring within 100 days of the last dose. IMAEs are specific events that include pneumonitis, diarrhea/colitis, hepatitis, nephritis/renal dysfunction, rash, and endocrine abnormalities (adrenal insufficiency, hypothyroidism/thyroiditis, hyperthyroidism, diabetes mellitus, and hypophysitis) for which subjects received immunosuppressive medication for treatment of the event, with the exception of endocrine events (hypothyroidism/thyroiditis, hyperthyroidism, hypophysitis, diabetes mellitus, adrenal insufficiency) but any inflammation confirmed by Histopathology, positive autoantibodies, findings of endoscope or imaging may also support the diagnosis of IMAEs.

Table 9.2-1 below provides a summary of the IMAEs category and their respective PTs.

Table 9.2-1 Preferred Terms of IMAEs

*** Irrespective of being treated with immunosuppressive medications or not.**

IMAE Category	PTs included under IMAE Category
Pneumonitis	Pneumonitis, Interstitial lung disease
Diarrhoea/Colitis	Diarrhoea, Colitis, Enterocolitis
Hepatitis	Hepatotoxicity, Hepatitis, Hepatitis acute, Autoimmune Hepatitis, AST Increased, ALT increased, Hyperbilirubinaemia, Blood Bilirubin Increased, Blood ALP increased
Adrenal Insufficiency*	Adrenal Insufficiency
Hypothyroidism*/Thyroiditis*	Hypothyroidism, Thyroiditis Thyroiditis acute (collapsed with thyroiditis for frequency), Autoimmune Thyroiditis (collapsed with thyroiditis for frequency)
Hyperthyroidism*	Hyperthyroidism
Hypophysitis*	Hypophysitis
Diabetes Mellitus*	Diabetes Mellitus, Diabetic ketoacidosis
Nephritis and renal dysfunction	Nephritis, Nephritis allergic, Tubulointerstitial nephritis, Acute renal failure, Renal Failure, Increased Creatinine
Rash	Rash, Rash maculo-papular

9.3 Definition of Serious Adverse Events

Among events meeting the definition of adverse events, an event will be assessed as "serious" when it involves any of the following:

1. Results in death
2. Is life-threatening
3. Requires inpatient hospitalization or prolongation of existing hospitalization [Note 1)]
4. Results in persistent or significant disability or incapacity
5. Is a congenital anomaly
6. Other medically important events [Note 2)]

[Note 1)]: This does not include hospitalization for this study or any of the following:

- Hospitalization or prolongation of existing hospitalization for a procedure (e.g., surgery, examination) that had been planned since before the study
- Hospitalization or prolongation of existing hospitalization for examination or education
- Hospitalization or prolongation of existing hospitalization for treatment (e.g., surgery) of a concurrent disease or associated symptom because of a change in the treatment strategy despite no worsening of the concurrent disease or associated symptom
- Hospitalization or prolongation of existing hospitalization for follow-up observation of an already healed or improved condition
- Admission to a hospice facility, nursing care facility, or rehabilitation facility
- Hospitalization or prolongation of existing hospitalization for post-study treatment or other treatments for the primary disease
- Hospitalization or prolongation of existing hospitalization for social reasons (e.g., leave of a caregiver or temporary absence of a family member)

[Note 2)]: Other medically important events include events that may not be immediately life-threatening or result in death or hospitalization but may jeopardize the subject or may require intervention to prevent one of the other outcomes listed in definitions 1 to 5 above. For such events, medical and scientific judgment should be exercised in deciding whether expedited reporting is appropriate. In other situations, these should also usually be considered serious. Examples of such events are intensive treatment in an emergency room, etc. for bronchospasm; blood disorder or convulsions that do not result in hospitalization;

development of drug dependency or drug abuse; or onset of a malignant tumor that is histologically different from the primary lesion. If a spread of any infectious factor mediated by the investigational product is suspected, it must be reported as a medically significant event.

9.4 Grading Criteria of Adverse Events

Adverse events not listed in the CTCAE will be graded using the following 5-point scale from 1 to 5, using the grade whose definition most closely applies to the event. If the grade definition text includes specific treatment, the clinical requirement for the treatment will be considered in the grading.

Grade 1 Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated.

Grade 2 Moderate; minimal, local, or noninvasive intervention indicated; limiting age-appropriate instrumental ADL [Note 1)].

Grade 3 Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self-care ADL [Note 2)].

Grade 4 Life-threatening consequences; urgent intervention indicated.

Grade 5 Death related to adverse events

A semi-colon (;) indicates “or” within the description of the grade.

Note 1): Instrumental ADL refers to preparing meals, shopping for groceries or clothes, using the telephone, managing money, etc.

Note 2): Self-care ADL refer to bathing, dressing and undressing, feeding self, using the toilet, taking medications, and not being bedridden

9.5 Assessment Criteria of Causal Relationship to the Investigational Product

Causal relationship to the investigational product will be assessed using the following 2-point scale, in light of the subject’s condition, medical history, concomitant drugs, and temporal relationship with the event onset. For all adverse events for which a causal relationship to the investigational product cannot be ruled out, the basis for the assessment will be entered into the comment field of the eCRF.

1. Related: There is a reasonable causal relationship between the use of the investigational product and the adverse event.
2. Not related: There is no reasonable causal relationship between the use of the investigational product and the adverse event.

A reasonable causal relationship means that there is evidence to suggest a causal relationship between the adverse event and the investigational product.

9.6 Event Outcome

The outcome of adverse events (i.e., symptoms and abnormal laboratory test values) will be assessed using the following criteria. The date of outcome will be the date of recovery from adverse events with an outcome of "Resolved," or the date of final assessment of adverse events with other outcomes.

Event outcome	Assessment Criteria (reference)
Resolved	The symptom resolved, or the test value normalized or returned to the pre-therapy level.
Resolving	The event improved in intensity, or the symptom showed an improving tendency.
Not resolved	The symptom or the test value has not improved.
Resolved with sequelae	The event resolved or was resolving, but with remaining symptoms of disability or other sequela
Death	The subject died.
Unknown (Lost to follow-up)	The subject was lost to follow-up.

9.7 Action Taken with the Investigational Product

For each adverse event experienced, the investigator or subinvestigator will determine the action taken with the investigational product, and record the action using one of the following seven categories. Note that the action here refers to the action taken by the investigator/subinvestigator, not by the subject.

Action taken by the investigator/ subinvestigator	Assessment Criteria
1. Does not changed	As a result of the adverse event, the investigator or subinvestigator has judged that administration of the investigational product may be continued.
2. Drug interrupted	As a result of the adverse event, the investigator or subinvestigator has judged that administration of the investigational product needs to be interrupted. Investigational product has begun but has temporarily ceased or investigational product infusion was started and was prematurely stopped and is expected to be re-introduced.

Action taken by the investigator/ subinvestigator	Assessment Criteria
3. Drug delayed	As a result of the adverse event, the investigator or subinvestigator has judged that administration of the investigational product needs to be delayed (or skipped). Administration of investigational product at the next scheduled time point has not started and has been delayed, but is expected to be re-introduced.
4. Dose reduced	As a result of the adverse event, the investigator or subinvestigator has judged that administration of the investigational product needs to be dose reduced.
5. Drug withdrawn	As a result of the adverse event, the investigator or subinvestigator has judged that administration of the investigational product needs to be discontinued.
6. Not applicable	The adverse event developed after the end of administration of the investigational product.
7. Other	None of the above apply, but as a result of the adverse event the investigator or subinvestigator has changed the mode of administration of the investigational product.

9.8 Any Treatment for Adverse Events

Any treatments given for the adverse event will be specified.

9.9 Actions to be Taken Following Onset of Any Serious Adverse Event

Following the subject's written consent to participate in the study, all serious adverse events that occur between the start of administration of the investigational product and 100 days after the last dose of investigational product, whether related or not related to the investigational product, must be noted, including those thought to be associated with protocol-specified procedures. If applicable, serious adverse events must be noted that relate to any later protocol-specified procedure (e.g., biopsy after the end of the treatment phase).

The investigator should report any serious adverse event that occurs after these time periods and that is believed to be related to the investigational product or protocol-specified procedure.

An SAE report should be completed for any event where doubt exists regarding its seriousness.

If the investigator believes that a serious adverse event is not related to the investigational product, but is potentially related to the conditions of the study (such as withdrawal of previous therapy or a complication of a study procedure), the relationship should be specified in the narrative section of the SAE Report Form.

Serious adverse events including overdose, whether related or not to the investigational product, and pregnancies must be reported to the sponsor (or designee) within 24 hours. Serious adverse events must be recorded on the SAE Report Form; pregnancies on the Pregnancy Report Form.

The report forms are to be transmitted via email or confirmed fax transmission to the following, and the original report forms are to remain on site:

SAE Email Address: Refer to Contact Information list.

SAE Fax Number: Refer to Contact Information list.

SAE Telephone Contact (required for serious adverse event and pregnancy reporting): Refer to Contact Information list.

If only limited information is initially available, follow-up reports are required. (Note: Follow-up SAE reports should include the same investigator term(s) reported initially.)

If an ongoing serious adverse event changes in its intensity or causal relationship to the investigational product or if new information becomes available, a follow-up SAE report should be sent within 24 hours to the sponsor (or designee) using the same procedure used for transmitting the initial SAE report.

All serious adverse events should be followed until resolution, resolving, or stabilization.

9.10 Actions to be Taken Following Pregnancy

Between the start of administration of the investigational product and 28 days after the end (i.e., completion/discontinuation) of the treatment phase, if it is subsequently discovered that a study subject is pregnant or may have been pregnant at the time of investigational product exposure, the investigational product will be permanently discontinued in an appropriate manner (e.g., dose tapering if necessary for subject safety).

The investigator will immediately notify the sponsor (or designee) of this event and complete and forward the Pregnancy Report Form to the sponsor (or designee) within 24 hours and in accordance with SAE reporting procedures described in Section 9.9.

Follow-up information regarding the course of the pregnancy, including perinatal and neonatal outcome and, where applicable, offspring information must be reported on the Pregnancy Report Form.

Any pregnancy that occurs in a female partner of a male study participant should be reported to the sponsor. Information on this pregnancy will be collected on the Pregnancy Report Form.

9.11 Actions to be Taken Following Overdose

An overdose is defined as the accidental or intentional administration of any dose of a product that is considered both excessive and medically important. All occurrences of overdose must be reported as a serious adverse event (see Section 9.9 for reporting details).

9.12 Other Safety Considerations

Any significant worsening noted during interim or final physical examinations, electrocardiogram, chest X-ray, and any other potential safety assessment required or not by the protocol should also be recorded as a non-serious or serious adverse event, as appropriate, and reported accordingly.

9.13 Provision of New Information

The sponsor, upon obtaining any information that might adversely affect the safety of the subject, impact the study conduct, or require a change in the approval of the IRB regarding study continuation, or information on serious unpredictable adverse drug reactions, will report the information promptly to all investigators engaged in the study and to the head of each study site. If it is deemed necessary to revise the informed consent form and written subject information, these will be revised promptly. Also, the sponsor will revise the clinical study protocol and investigator's brochure as appropriate.

10 Statistical Analysis

Ono Pharmaceutical Co., Ltd. will perform the statistical analysis. The analysis principles are outlined below, while details of the analytical methods are described in the statistical analysis plan prepared by the person responsible for statistical analysis and the statistical analysis personnel.

The statistical analysis plan Version 1.0 will be finalized before randomization for the first subject. The statistical analysis plan will be amended after Version 1.0 of it was fixed, as needed.

Additional analyses not stated in the statistical analysis plan may be performed, and will be reported in the clinical study report.

10.1 Description of Statistical Analysis Methods including the Interim Analysis

Stopping criteria from statistical viewpoint is not set because formal interim analysis is not planned in this study.

10.2 Independent Data Monitoring Committee

An Independent Data Monitoring Committee (IDMC) will be established to independently monitor safety data as well as operation of the study. The IDMC will have periodic meetings to ensure careful monitoring of subject safety. Ad hoc meetings may also be held as required. After each meeting, the IDMC will provide advice on whether the study should be continued, modified, or discontinued on the basis of the toxicities observed. Detailed tasks of the IDMC will be described in a written procedure prepared separately.

10.3 Significance Level to be Used

10.3.1 Efficacy

Two-sided test will be performed with 5% significance level. If superiority in OS is determined, a hierarchical hypothesis testing approach for the key secondary endpoints will be used to preserve a study-wise type I error rate at 5%. The key secondary endpoints will be tested in the following hierarchical order:

1. ORR
2. PFS

Interactions will be tested with two-sided 15% significance level.

10.3.2 Safety

Not applicable because no statistical tests will be performed on safety.

10.4 Determination of Subjects included in Analysis

10.4.1 Definitions of Analysis Sets

For efficacy endpoints, the Intention-to-Treat (ITT) population will be the analysis set.

For safety endpoints, the Safety Set (SAF) will be the analysis set.

For anti-drug antibody endpoint, the Anti-Drug Antibody Set (ADA) will be the analysis set.

Individual sets of subjects are defined below.

10.4.1.1 Enrolled Set（ENR）

The Enrolled Set (ENR) will consist of all subjects enrolled (enrollment) in this study via IWRS.

10.4.1.2 Randomized Set（RND）／ITT

The Randomized Set (RND) and ITT will consist of all randomized (randomization) subjects.

10.4.1.3 SAF

The SAF will consist of all subjects given at least one dose of the investigational product.

10.4.1.4 ADA

The ADA will consist of all subjects which meet all of the following items in the SAF population.

1. It is not GCP noncompliant subject defined in 10.4.2 “Criteria for Handling of Subjects.”

2. It has measurements of anti-ONO-4538-antibody (the sample which screening result is potential positive but confirmatory result is not available will be excluded) at both baseline and at least one post-baseline anti-drug antibody assessment.

10.4.2 Criteria for Handling of Subjects

Criteria for handling of subjects are specified below.

10.4.2.1 Disqualified Subjects

Disqualified subjects are defined as those who failed to fulfill 4.2 "Inclusion Criteria" and conflict with 4.3 "Exclusion Criteria."

10.4.2.2 GCP Noncompiant Subjects

GCP noncompliant subjects are defined as those who were enrolled based on materials not reviewed by the IRB/IEC, who were enrolled at a medical institution where no study contract had been signed, or who did not provide informed consent appropriately.

10.4.2.3 Untreated Subjects

Untreated subjects are defined as those who received no dose of the investigational product.

10.4.2.4 Incomlete Anti-Drug Antibody Subjects

Incomplete anti-drug antibody subjects are defined as those who were without baseline anti-ONO-4538-antibody measurements, and/or all post-baseline anti-ONO-4538-antibody measurements.

For any other subject with unexpected problems, the handling in the analysis will be determined by the sponsor before fixing the data for the interim analysis, on the basis of discussion with the medical officer and the biostatistical advisor.

10.4.3 Criteria for Handling of Time Points of Evaluation

If the actual date of examination is not the protocol-specified date, analysis by the time point of evaluation will be performed with inclusion of only the data obtained within the respective allowable windows, as shown in Table 6-1, Table 6-2 and Table 6-3. Of note, Day 1 of each cycle of treatment during the treatment period is defined as the first day of each cycle.

10.5 Analysis Methods

10.5.1 Reliability of the Study

10.5.1.1 Inclusion and Exclusion for Analyses

For each analysis set, the frequency of inclusion and exclusion for analysis as well as the frequency by reason for exclusion will be summarized by treatment group.

10.5.1.2 Withdrawals and Dropouts

The number of subjects withdrawn from the study will be summarized by treatment group.

10.5.2 Details of Treatment with the Investigational Product

Details of treatment with the investigational product will be summarized by treatment group.

10.5.3 Distribution of Patient Background Factors

In each treatment group, frequency distributions and summary statistics will be calculated for the following variables, and comparability between the two treatment groups will be assessed.

10.5.3.1 Demographic Variables

Categorical variables such as sex, age, race, and ethnicity will be summarized in terms of frequency and continuous variables such as height, weight, and BMI will be summarized using summary statistics (mean, standard deviation, median, minimum and maximum value) by treatment group.

10.5.3.2 Patient Background Characteristics

Medical history, concurrent diseases, history of alcohol consumption, history of smoking, performance status score（ECOG）, date of diagnosis of the primary disease, lesion sites (anatomical subsites according to the TNM classification), clinical classification (TNM clinical classification), disease stage (stage grouping: TNM classification), histological classification, number of organs with metastases, date of surgery (including EMR or ESD) for the primary lesion (for subjects with recurrent esophageal cancer), residual tumor (R) classification (for subjects with recurrent esophageal cancer), site of recurrence (for subjects with recurrent esophageal cancer), and past treatments for cancer (e.g., surgery, radiotherapy, pharmacotherapy) will be summarized by treatment group.

10.5.3.3 Baseline Values of Observation

Diameters of the target lesions will be summarized using summary statistics (mean, standard deviation, median, minimum and maximum value) by treatment group.

10.5.3.4 Concomitant drugs and therapies

The presence or absence of analgesics will be summarized in terms of frequency by treatment group.

10.5.4 Primary Efficacy Endpoint

10.5.4.1 Purpose of Analysis

To assess the efficacy of the ONO-4538 group compared to the control group (docetaxel or paclitaxel) on basis of overall survival (OS) as the primary endpoint in patients with esophageal cancer refractory to or intolerant of combination therapy with fluoropyrimidine and platinum-based drug .

10.5.4.2 Hypothesis

The ONO-4538 group is superior to the control group in terms of OS.

10.5.4.3 Analytical Item and Data Handling

1. Analytical Item

   OS

2. Handling of data

   OS will be calculated from the following equation:

   OS (Months) = (“Date of death from any cause” - “Date of randomization” + 1) / 30.4375

For subjects lost to follow-up and subjects who are alive at the time of data cutoff date, data will be censored at the time the subject was last confirmed to be alive.

10.5.4.4 Analytical Methods

The following analyses will be performed on OS:

1. Primary analytical method

   Data will be compared between the two treatment groups using the stratified log-rank test with the location (Japan vs rest of the world), the number of organs with metastases (≤1 vs. ≥2), and PD-L1 expression (≥1% vs. <1% or indeterminate) which are the randomization factors as the stratification factors.

2. Secondary analytical method

1） The hazard ratio and the corresponding two-sided 95% confidence interval (CI) for the ONO-4538 group relative to the control group will be estimated using a stratified Cox proportional hazards model with the randomization factors as the stratification factors.

2） The Kaplan-Meier curve will be plotted for each treatment group. Using the Kaplan-Meier method, the median OS and the corresponding two-sided 95% CI will be estimated for each treatment group. CI for median OS will be calculated using the Brookmeyer and Crowley method based on a log-log transformed CI for the survivor function.

3） For each treatment group, the survival rates at months 6, 9, and 12 will be derived from the Kaplan-Meier method and corresponding two-sided 95% confidence interval will be calculated based on Greenwood formula for variance derivation and on log-log transformation applied on the survivor function.

10.5.5 Secondary Efficacy Endpoints

10.5.5.1 Purpose of Analysis

To assess the efficacy of ONO-4538 from various aspects by exploratory analysis on other endpoints than the primary endpoint

10.5.5.2 Analytical Items and Data Handling

1. Analytical items

   1）Objective response rate

   2）Disease control rate

   3）Progression free survival

   4）Duration of response

   5）Time to response

   6）Best overall response

   7）Maximum percent change from baseline in the sum of diameters of target lesions

2. Handling of data

Overall response and best overall response will be determined solely by imaging assessment according to the RECIST Guideline Version 1.1, and will not take into account any clinical/symptomatic progression. Evaluable imaging data will be overall response without an overall response of “NE.”

For each analytical item, the handling of data will be as follows:

1）Objective response rate is defined as the percentage of subjects whose best overall response is assessed as either CR or PR.

2）Disease control rate is defined as the percentage of subjects whose best overall response is assessed as CR, PR or SD.

3）Progression free survival will be calculated from the following equation:

Progression free survival (Months) = (“Earlier date on which either the overall response was assessed as PD or the subject died of any cause” - “Date of randomization” + 1) / 30.4375

For subjects who are alive without an overall response of PD, data will be censored on the day of last evaluable imaging. For subjects who are alive without evaluable imaging, data will be censored on the day of randomization. For subjects started on post-study esophageal cancer treatment before an overall response of PD or death, data will be censored on the day of last evaluable imaging before initiation of post-study esophageal cancer treatment.

4）Duration of response will be calculated from the following equation:

Duration of response (Months) = (“Earlier date on which either the overall response was assessed as PD for the first time after confirmed response or the subject died of any cause” - “Date of first assessment of confirmed CR or PR” + 1) / 30.4375

This will be calculated in subjects who had confirmed CR or PR during the study.

5）Time to response will be calculated from the following equation:

Time to response (Months) = (“Date of first assessment of confirmed CR or PR” - “Date of randomization” + 1) / 30.4375

For subjects without a response in individual treatment groups, the data on the time to response will be censored at the “longest time to response + 1 day among subjects with response in the treatment group”.

6）Maximum percent change from baseline in the sum of diameters of target lesions

In subjects with target lesions, on the basis of the diameters of target lesions as assessed according to the RECIST Guideline Version 1.1, the maximum percent change from baseline in the sum of diameters of target lesions will be calculated from the following equation. However, this calculation

will exclude the diameter data obtained after an overall response of PD and after new treatment for cancer.

$$\text{Maximum percent change from baseline in the sum of diameters of target lesions (\%)} = \left(\frac{\text{"Smallest sum of diameters of target lesions after study treatment"}}{\text{"Baseline sum of diameters of target lesions}} - 1 \right) \times 100$$

For progression free survival and duration of response, the details of judgment criteria of event and censor will be provided in the statistical analysis plan.

10.5.5.3 Analytical Methods,

1. For the analytical item 1) and 2), the following analyses will be performed:
   （1） Data will be compared between the two treatment groups using the Cochran-Manten-Haenszel test with the randomization factors as the stratification factors. The associated odds ratio and the corresponding two-sided 95% CI for the ONO-4538 group relative to the control group and the estimate of the difference in objective response rate and corresponding two-sided 95% CI will be calculated using Cochran-Manten-Haenszel method and adjusted by the same stratification factors.
   （2） The proportions and the corresponding two-sided 95% CI will be estimated using the Clopper-Pearson method for each treatment group.
2. For the analytical item 3), the following analyses will be performed:
   （1） Data will be compared between the two treatment groups using stratified log-rank test with the randomization factors as the stratification factors. The hazard ratio and the corresponding two-sided 95% CI for the ONO-4538 group relative to the control group will be estimated using a stratified Cox proportional hazards model with the randomization factors as the stratification factors.
   （2） The Kaplan-Meier curve will be plotted for each treatment group. Using the Kaplan-Meier method, the median OS and the corresponding two-sided 95% CI will be estimated for each treatment group. CI for median PFS will be calculated using the Brookmeyer and Crowley method based on a log-log transformed CI for the survivor function.
   （3） For each treatment group, the progression free survival rates at months 3, 6, 9, and 12 will be derived from the Kaplan-Meier estimates and the corresponding two-sided 95% confidence intervals will be derived based on Greenwood’s formula for variance derivation and on log-log transformation applied on the survivor function.

3. For the analytical item 4), the following analyses will be performed:

（1）The Kaplan-Meier curve will be plotted for each treatment group. Using the Kaplan-Meier method, the median OS and the corresponding two-sided 95% CI will be estimated for each treatment group. CI for median duration of response will be calculated using the Brookmeyer and Crowley method based on a log-log transformed CI for the survivor function..

4. For the analytical item 5), the following analyses will be performed:

（1）Summary statistics will be calculated for subjects whose best overall response is CR or PR.

（2）The Kaplan-Meier curve will be plotted for each treatment group.

5. For the analytical item 6), the percentage of CR, PR, SD, PD and NE will be calculated for each treatment group. For the percentages of CR, PR and SD, the corresponding 95% CI will be estimated using the Clopper-Pearson method for each treatment group.

6. For the analytical item 7), waterfall plot will be displayed.

10.5.6 Interaction Analysis

For the primary efficacy endpoint, the interaction between treatment group and the each demographic variable will be assess using a stratified Cox proportional-hazards model with the randomization factors as the stratification factors. The demographic factors which will be used for interaction analysis will be provided in the statistical analysis plan.

10.5.7 Adjusted Analysis

If any factors are found to be imbalanced between the two treatment groups or to affect evaluation in the study, adjusted analysis for the primary endpoint will be performed using a stratified Cox proportional-hazards model with the randomization factors as the stratification factors and those factors as adjustment factors . The adjustment factors for adjusted analysis will be provided in the statistical analysis plan.

10.5.8 Analysis of safety

10.5.8.1 Purpose of Analysis

To assess the safety of ONO-4538.

10.5.8.2 Analytical methods

1. Numbers of subjects with adverse events and adverse drug reactions in each treatment group will be summarized

2. Adverse events and adverse drug reactions in each treatment group will be summarized by System Organ Class (SOC), Preferred Term (PT), Grade, etc.
3. For laboratory parameters whose Grade were defined in CTCAE v4.0, the baseline Grade and the worst CTC Grade after treatment in each treatment group will be evaluated in the shift table.
4. Numbers of subjects with abnormal hepatic function and thyroid dusfunction in each treatment group will be summarized.

10.5.9 Data Reviews

After data collection from this study, using the data before fixed, a preliminary analysis will be performed in terms of the distribution of background factors etc., to determine the appropriateness of planned analytical methods.

10.5.10 Other Issues

Subgroup analyses will be performed on the efficacy and safety endpoints.

Analysis for anti-ONO-4538 antibody will be performed.

Analysis for Patient Reported Outcomes will be performed.

Exploratory analyses may be additionally performed on the efficacy or safety of ONO-4538 and biomarkers as necessary. Pharmacokinetic (PK) and pharmacokinetic/pharmacodynamic (PK/PD) analyses may be performed as necessary. If these analyses are performed, reports will be separately prepared.

11 Ethical Conduct of the Study

11.1 Good Clinical Practice

This study will be conducted in accordance with the ethical principles that are consistent with the Good Clinical Practice (GCP) guideline developed by the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH).

This study will be conducted in compliance with the protocol. Prior to initiation of the study, the protocol and any amendments, and the subject information and informed consent form must have received approval/favorable opinion from the institutional review board or independent ethics committee (IRB/IEC).

All potential serious breaches, if any, must be immediately reported to the sponsor. A serious breach is a breach of the conditions and principles of GCP in connection with the study or the protocol, which is

likely to affect, to a significant degree, the safety or physical or mental integrity of the subjects of the study or the scientific value of the study.

Personnel involved in conducting this study will be qualified by education, training, and experience to perform their respective tasks.

This study will not use the services of study personnel where sanctions have been invoked or where there has been scientific misconduct or fraud (e.g., loss of medical licensure, debarment).

11.2 Institutional Review Board or Independent Ethics Committee (IRB/IEC)

Before study initiation, the investigator must have written and dated approval from the IRB/IEC for the protocol, subject information and informed consent form, subject recruitment materials or method of recruitment (e.g., advertisements), and any other written information to be provided to subjects. The investigator or sponsor should also provide the IRB/IEC with a copy of the Investigator's Brochure or product labeling information to be provided to subjects and any updates.

The investigator or sponsor should provide the IRB/IEC with reports, updates, and other information (e.g., expedited safety reports, amendments, and administrative letters) according to regulatory requirements or institution procedures.

11.3 Privacy Protection for Subjects

The utmost consideration should be given to the protection of subjects' privacy. Subject Identification Codes will be used in place of subjects' names to distinguish individuals on the eCRFs and in other data. Completed eCRFs will not be used for purposes other than this study. Any information made available to the sponsor via this study will never be released to a third party.

12 Subject Informed Consent

12.1 Informed Consent

The investigator or subinvestigator must ensure that patients are clearly and fully informed about the purpose, potential risks, and other critical issues regarding the clinical study in which they volunteer to participate.

In situations where consent cannot be given by patients, their legally acceptable representatives are clearly and fully informed about the purpose, potential risks, and other critical issues regarding the clinical study in which the subject volunteers to participate.

The sponsor will provide the investigator with an appropriate (i.e., global or local) sample informed consent form, which will include all elements required by ICH-GCP and applicable regulatory

requirements. The sample informed consent form will adhere to the ethical principles that have their origin in the Declaration of Helsinki.

The investigator or subinvestigator should:

1. Provide a copy of the consent form and written information about the study in the language in which the patient is most proficient prior to clinical study participation. The language must be non-technical and easily understood.
2. Allow time necessary for the patient or patient's legally acceptable representative to inquire about the details of the study.
3. Obtain an informed consent form signed and dated by the patient or the patient's legally acceptable representative and by the person who conducted the informed consent discussion.
4. Obtain the IRB/IEC's written approval of the written informed consent form and any other information to be provided to the patients, prior to the beginning of the study and after any revisions are completed for new information.
5. If informed consent is initially given by a subject's legally acceptable representative or legal guardian, and the subject subsequently becomes capable of making and communicating his or her informed consent during the study, consent must additionally be obtained from the subject.
6. Revise the informed consent form whenever important new information becomes available that is relevant to the subject's consent. The investigator, or a person designated by the investigator, should fully inform the subject or the subject's legally acceptable representative or legal guardian of all pertinent aspects of the study and of any new information relevant to the subject's willingness to continue participation in the study. This communication should be documented.

The confidentiality of records that could identify subjects must be protected, respecting the privacy and confidentiality rules applicable to regulatory requirements and the subjects' signed informed consent forms and subject information.

The consent form and subject information must also include a statement that the sponsor and regulatory authorities have direct access to subject records.

The rights, safety, and well-being of the study subjects are the most important considerations and should prevail over the interests of science and society.

12.2 Withdrawal of Consent

Subjects who request to discontinue study drug will remain in the study and must continue to be followed for protocol specified follow-up procedures. The only exception to this is when a subject specifically withdraws consent for any further contact with him/her or persons previously authorized by subject to provide this information. Subjects should notify the investigator of the decision to withdraw consent from future follow-up **in writing**, whenever possible. The withdrawal of consent should be explained in detail in the medical records by the investigator, as to whether the withdrawal is from further treatment with study drug only or also from study procedures and/or post treatment study follow-up, and entered on the appropriate CRF page. In the event that vital status (whether the subject is alive or dead) is being measured, publicly available information should be used to determine vital status only as appropriately directed in accordance with local law.

13 Rules or Criteria for Withdrawing Individual Subjects from the Study or Discontinuing Parts of the Study or Entire Study

13.1 Rules or Criteria for Withdrawing Individual Subjects from the Study

In any of the following instances, the subject will be immediately withdrawn from the treatment phase or the follow-up phase. The investigator or subinvestigator will provide appropriate treatment and also make observations and measurements of clinical symptoms and routine laboratory tests at the time of withdrawal as far as possible. Details of the withdrawal, including the date of withdrawal, reason for withdrawal, treatment after withdrawal, and clinical course will be entered into the eCRF, and the sponsor will be notified promptly. For a subject who do not make a scheduled visit and is thereby discontinued from study treatment, the investigator or subinvestigator will perform follow-up investigations by telephone, letter, or other means to determine the reason for not making the study visit, the subsequent course, and the patient’s current condition, and enter the results into the eCRF.

1. When the subject requests to withdraw from the study
2. When the subject is found not to fulfill the inclusion criteria
3. When the subject is found to meet any of the exclusion criteria
4. When the investigator or subinvestigator has judged it difficult to continue with further study procedures because of the occurrence of an adverse event, whether or not causally related to the investigational product
5. When the investigator or subinvestigator has judged it inappropriate to continue with further study procedures because of progressive disease

6. When the subject has failed to return to the study site and thus no further study procedures can be performed
7. When the investigator or subinvestigator has judged it inappropriate to continue with further study procedures for other reasons

14 Study Management

14.1 Compliance with the Protocol and Protocol Amendments

This study will be conducted as described in this approved protocol. The investigator or subinvestigator should not implement any deviation or change to the protocol without prior review and documented approval/favorable opinion from the IRB/IEC of an amendment, except where necessary to protect study subjects from an immediate hazard.

If a deviation or change to the protocol is implemented to eliminate an immediate hazard prior to obtaining IRB/IEC approval/favorable opinion, the deviation or change will be submitted as soon as possible to:

- IRB/IEC for review and approval/favorable opinion
- Sponsor
- Regulatory Authority(ies), if required by local regulations

Documentation of approval signed by the chairperson or designee of the IRB(s)/IEC(s) must be sent to the sponsor.

If an amendment substantially alters the study design or increases the potential risk to the subject: (1) the consent form and subject information must be revised and submitted to the IRB(s)/IEC(s) for review and approval/favorable opinion; (2) the revised documents must be used to obtain consent from subjects currently enrolled in the study if they are affected by the amendment; and (3) the new documents must be used to obtain consent from new subjects prior to enrollment.

If the revision is made via an administrative letter, investigators must inform their IRB(s)/IEC(s).

15 Completion of Electronic Case Report Forms (eCRFs)

15.1 eCRF

For all subjects who provide written consent, the investigator or designee will enter data into the eCRF promptly after completion of each assessment procedure, as stipulated by the protocol, for each subject. Clinical research coordinators will only be allowed to transcribe data from source documents to the eCRF. All data in the eCRF must be consistent with source documents, and any discrepancy should be explained by the investigator. The investigator will confirm that all entries in the eCRF

(including audit trails and query responses) are accurate and complete and ensure the accuracy and completeness by electronically signing the eCRF. An electronic signature is the legally binding equivalent of the signatory's handwritten signature. Electronically signed entries cannot be modified. Procedures for eCRF data entries, changes and corrections, and electronic signatures will follow a written procedure prepared separately.

The investigator or designee should meet the sponsor's training requirements and access the sponsor's electronic data capture (EDC) system using the unique user account provided by the sponsor. User accounts are not to be shared or reassigned to other individuals.

16 Monitoring

Representatives of the sponsor must be allowed to visit all study site locations periodically to assess the data quality and study integrity. On site they will review study records and directly compare them with source documents, discuss the conduct of the study with the investigator or subinvestigator, and verify that the facilities remain acceptable.

In addition, the study may be evaluated by the sponsor's internal auditors and government inspectors, who must be allowed access to the eCRFs, source documents, other study files, and study facilities. The sponsor's audit reports will be kept confidential.

The investigator shall notify the sponsor promptly of any inspections scheduled by regulatory authorities, and promptly forward copies of inspection reports to the sponsor.

17 Procedures to Ensure Compliance of Subject Medication and Other Stipulations

The investigator or subinvestigator will ensure treatment compliance and specimens such as blood and urine are collected appropriately at the study site.

18 Retention of Essential Documents

18.1 Record Retention

Essential documents should be retained for up to 2 years after the last approval of the marketing application in an ICH region or until there are no pending or contemplated marketing applications in an ICH region, whichever is later, or until 2 years have elapsed since the formal discontinuation of clinical development for the investigational product. However, these essential documents should be retained for the period required by the applicable regulatory requirements or by agreement with the sponsor, whichever is longer. These essential documents must not be destroyed or transferred to

another location without prior written notification of the sponsor. The sponsor is responsible for informing the investigator or institution when these documents no longer need to be retained.

19 Publication of Study Results

Results from the study conducted in accordance with this protocol are the property of the sponsor and study sites. Any publication of the study results requires prior approval of the sponsor. Issues related to publication will be resolved by discussion between the sponsor and the coordinating investigator.

20 Literature references

1) World Health Organization, World Cancer Report 2014
2) Statistics and Information Department, Minister's Secretariat, Ministry of Health, Labour and Welfare. Reports on Vital Statistics, published on September 11, 2014
3) Center for Cancer Control and Information Services, National Cancer Center. Monitoring of Cancer Incidence in Japan: Cancer Incidence and Incidence Rates in Japan in 2011 (MCIJ2011) (2015)
4) Ozawa S, Tachimori Y, Baba H, Fujishiro M, Matsubara H, Numasaki H, et al. Comprehensive Registry of Esophageal Cancer in Japan, 2004. Esophagus. 2012;9:75-98.
5) Center for Cancer Control and Information Services, National Cancer Center. White book on cancer statistics 2012: toward evidence-based care for cancer
6) Health Insurance Review Agency (HIRA). Disease statistics of 2013
7) Department of Health, Executive Yuan. Cancer Registry Annual Report, 2010 TAIWAN
8) Keio Igaku 2009;85 (2) 127-135
9) Steevens J, Schouten LJ, Goldbohm RA, van den Brandt PA. Alcohol consumption, cigarette smoking and risk of subtypes of oesophageal and gastric cancer: a prospective cohort study. Gut. 2010;59:39-48.
10) Sakata K, Hoshiyama Y, Morioka S, Hashimoto T, Takeshita T, Tamakoshi A; JACC Study Group. Smoking, alcohol drinking and esophageal cancer: findings from the JACC Study. J Epidemiol. 2005;15 Suppl 2:S212-9.
11) Ishiguro S, Sasazuki S, Inoue M, Kurahashi N, Iwasaki M, Tsugane S; JPHC Study Group. Effect of alcohol consumption, cigarette smoking and flushing response on esophageal cancer risk: a population-based cohort study (JPHC study). Cancer Lett. 2009;275:240-6.
12) Secretan B, Straif K, Baan R, Grosse Y, El Ghissassi F, Bouvard V, et al. WHO International Agency for Research on Cancer Monograph Working Group. A review of human carcinogens--Part E: tobacco, areca nut, alcohol, coal smoke, and salted fish. Lancet Oncol. 2009;10:1033-4.
13) Japan Esophageal Society. Guidelines for diagnosis and treatment of carcinoma of the esophagus, April 2012
14) Center for Cancer Control and Information Services, National Cancer Center. Monitoring of Cancer Incidence in Japan. Viability report from 2003 to 2005 (2013)
15) Iizuka T, Kakegawa T, Ide H, Ando N, Watanabe H, Tanaka O, et al. Phase II evaluation of cisplatin and 5-fluorouracil in advanced squamous cell carcinoma of the esophagus: a Japanese Esophageal Oncology Group Trial. Jpn J Clin Oncol. 1992;22:172-6.

16) Ajani JA. Contributions of chemotherapy in the treatment of carcinoma of the esophagus: results and commentary. Semin Oncol. 1994;21:474-82.
17) Muro K, Hamaguchi T, Ohtsu A, Boku N, Chin K, Hyodo I, et al. A phase II study of single-agent docetaxel in patients with metastatic esophageal cancer. Ann Oncol. 2004;15:955-9.
18) Kato K, Tahara M, Hironaka S, Muro K, Takiuchi H, Hamamoto Y, et al. A phase II study of paclitaxel by weekly 1-h infusion for advanced or recurrent esophageal cancer in patients who had previously received platinum-based chemotherapy. Cancer Chemother Pharmacol. 2011;67:1265-72.
19) Freeman GJ, Long AJ, Iwai Y, et al. Engagement of the PD 1 immunoinhibitory receptor by a novel B7 family member leads to negative regulation of lymphocyte activation. J Exp Med 2000;192(7):1027-34.
20) Latchman Y, Wood CR, Chernova T, et al. PD-L2 is a second ligand for PD-1 and inhibits T cell activation. Nat Immunol 2001;2(3):261-8.
21) Carter LL, Fouser LA, Jussif J, et al. PD-1:PD-L inhibitory pathway affects both CD4+ and CD8+ T cells and is overcome by IL-2. Eur J Immunol 2002;32(3):634-43.
22) Ohigashi Y, Sho M, Yamada Y, Tsurui Y, Hamada K, Ikeda N, et al. Clinical significance of programmed death-1 ligand-1 and programmed death-1 ligand-2 expression in human esophageal cancer. Clin Cancer Res. 2005;11:2947-53.
23) Kataoka K, Tsushima T, Mizusawa J, Hironaka S, Tsubosa Y, Kii T, et al. A randomized controlled Phase III trial comparing 2-weekly docetaxel combined with cisplatin plus fluorouracil (2-weekly DCF) with cisplatin plus fluorouracil (CF) in patients with metastatic or recurrent esophageal cancer: rationale, design and methods of Japan Clinical Oncology Group study JCOG1314 (MIRACLE study). Jpn J Clin Oncol. 2015; 1-5.
24) Hironaka S, Tsubosa Y, Mizusawa J, Kii T, Kato K, Tsushima T, et al. Phase I/II trial of 2-weekly docetaxel combined with cisplatin plus fluorouracil in metastatic esophageal cancer (JCOG0807). CancerSci. 2014; 105: 1189-95.
25) Mizota A, Shitara K, Kondo C, Nomura M, Yokota T, Takahara D, et al. A Retrospective Comparison of Docetaxel and Paclitaxel for Patients with Advanced or Recurrent Esophageal Cancer Who Previously Received Platinum-Based Chemotherapy. Oncology 2011;81:237-242
26) Shirakawa T, Kato K, Nagashima K, Nishikawa A, Sawada R, Takahashi N, et al. A retrospective study of docetaxel or paclitaxel in patients with advanced or recurrent esophageal squamous cell carcinoma who previously received fluoropyrimidine-and platinum-based chemotherapy. Cancer Chemother Pharmacol (2014) 74:1207-1215

27) Ford HE, Marshall A, Bridgewater JA, Janowitz T, Coxon FY, Wadsley J. Docetaxel versus active symptom control for refractory oesophagogastric adenocarcinoma (COUGAR-02): an open-label, phase 3 randomised controlled trial. Lancet Oncol. 2014;15(1):78-86.

28) R.L Ferris, G. Blumenschein, Jr., J.Fayette, J.Guigay, A.D.Colevas, L.Licitra, et al. Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. N Engl J Med. 2016;375:1856-67.

29) Toshihiro Kudo, Yasuo Hamamoto, Ken Kato, Takashi Ura, Takashi Kojima, Takahiro Tsushima, et al. Nivolumab treatment for oesophageal squamous-cell carcinoma: an open-label, multicentre, phase 2 trial. Lancet Oncol. 2017;18:631-639.

Appendix 1

Criteria and Definitions Used in This Study

In this study, the TNM Classification of Malignant Tumours, Seventh Edition, published by the Union for International Cancer Control (UICC) (hereinafter referred to as "the TNM classification") will be used to determine the lesion sites (anatomical subsites), clinical classification, and disease stage.

1. Lesion site (anatomical subsites: the TNM classification)

1) Cervical esophagus (Ce)

   From the lower border of the cricoid cartilage to the thoracic inlet (superior margin of the sternum), i.e., the range up to approximately 18 cm below the upper incisor teeth

2) Thoracic esophagus (Te)

   (i) Upper thorax (Ut)

   From the thoracic inlet to the tracheal bifurcation, i.e., the range up to approximately 24 cm below the upper incisor teeth.

   (ii) Middle thorax (Mt)

   The proximal half of the two equal portions between the tracheal bifurcation and the esophagogastric junction. The lower border is approximately 32 cm from the upper incisor teeth.

   (iii) Lower thorax (Lt) (including the esophagogastric junction)

   A length of approximately 8 cm including the abdominal esophagus, consisting of the distal half of the two equal portions between the tracheal bifurcation and the esophagogastric junction. The lower border is approximately 40 cm from the upper incisor teeth.

2. Clinical classification (the TNM classification)

T： Primary tumor

TX	Primary tumor cannot be assessed
T0	No evidence of primary tumor
Tis	Intraepithelial carcinoma/ high grade dysplasia
T1	Tumor invades lamina propria, muscularis mucosae, or submucosa
	T1a: Tumor invades lamina propria or muscularis mucosae
	T1b: Tumor invades submucosa
T2	Tumor invades muscularis propria
T3	Tumor invades adventitia

T4 Tumor invades adjacent structures

T4a: Tumor invades pleura, pericardium, or diaphragm

T4b: Tumor invades other adjacent structures, such as the aorta, vertebral body, and trachea

N: Regional lymph nodes

NX Regional lymph node(s) cannot be assessed

N0 No regional lymph node metastasis

N1 Metastasis in 1-2 regional lymph nodes

N2 Metastasis in 3-6 regional lymph nodes

N3 Metastasis in 7 or more regional lymph nodes

M：Distant metastasis

M0 No distant metastasis

M1 Distant metastasis

Residual tumor (R) classification

RX Presence of residual tumor cannot be assessed

R0 No residual tumor

R1 Microscopic residual tumor

R2 Macroscopic residual tumor

Definition of regional lymph nodes:

Regional lymph nodes are those along the lymph flow of the esophagus, including the celiac lymph nodes and the cervical para-esophageal lymph nodes (regardless of the primary site), but do not include the supraclavicular lymph nodes.

3. Disease stage (stage grouping, the TNM classification)

Stage	T	N	M
0	Tis	N0	M0
IA	T1	N0	M0
IB	T2	N0	M0
IIA	T3	N0	M0
IIB	T1, T2	N1	M0
IIIA	T4a	N0	M0
	T3	N1	M0
	T1, T2	N2	M0
IIIB	T3	N2	M0
IIIC	T4a	N1, N2	M0
	T4b	Any N	M0
	Any T	N3	M0
IV	Any T	Any N	M1

4. Definition of terms used in this study

Primary lesion: In the present study, a primary lesion is defined as follows.

- Main lesion in isolated lesion
- Main lesion and all secondary lesions in multiple lesions

Main lesion/secondary lesion: In case of multiple lesions, the lesion with the greatest depth of invasion in clinical diagnosis is regarded as the main lesions, and others as secondary lesions. Both include intraepithelial spread[Note)].

- Main lesion = main focus + intraepithelial spread
- Secondary lesion = secondary focus + intraepithelial spread

[Note)]: In this study, intraepithelial spread refers to the intraepithelial lesion that continuously spreads from the primary focus.

Appendix 2

Tumor Response Assessment Criteria Used in This Study

In this study, the RECIST Guideline Ver. 1.1[Note 1)] will be used for the assessment of tumor response by CT or other imaging examinations.

[Note 1)]: Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST Guideline Ver. 1.1. Eur J Cancer. 2009 Jan;45(2):228-47.

A. Imaging-based assessment according to the RECIST Guideline Ver. 1.1

A-1 Examinations used for the response assessment

- CT of the neck to abdomen (recommended slice thickness, 5 mm) (or MRI if necessary)
- Tumor markers

A-2 Measurability of tumors at baseline

At baseline, tumor lesions/lymph nodes will be categorized as measurable or non-measurable according to criteria A-3 and A-4.

Tumor diameters should be measured in the axial plane of CT. Tumors should not be measured in the sagittal or coronal plane of three-dimensional imaging.

Baseline imaging will be the last imaging performed within 28 days before randomization. However, if other imaging is performed between randomization and the first dose of the investigational product, the later imaging will be used as the baseline.

A-3 Definition of measurable lesions

Non-lymph node lesions (Tumor lesions)

On CT or MRI, a lesion will be considered measurable if the dimension (longest diameter) of the lesion in the plane of measurement is at least twice the slice thickness and at least 10 mm.

Lymph node lesions

To be considered pathologically enlarged and measurable, a lymph node must be ≥15 mm in short axis when assessed by CT scan (with CT slice thickness no greater than 5 mm).

A-4 Definition of non-measurable lesions

All lesions other than measurable lesions will be considered non-measurable. The following lesions will also be considered non-measurable irrespective of the imaging technique and lesion size.

- Esophageal primary lesion
- 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 measurable lesions) (Blastic bone lesions are non-measurable.)
- Cystic lesions ("cystic lesions" thought to represent cystic metastases can be considered as measurable lesions if they meet the definition of measurable lesions; however, if noncystic lesions are present in the same patient, these are preferred for selection as target lesions.)
- Lesions with prior local treatment such as radiotherapy (with the exception of lesions showing progression on imaging after radiotherapy)
- 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

A-5 Target lesions

Among measurable lesions, up to a maximum of the five largest lesions (and a maximum of two lesions per organ) representative of all involved organs should be identified as target lesions. Target lesions should be selected on the basis of their size (longest axis for non-nodal lesions, short axis for nodal lesions) and be representative of all involved organs, but in addition should be those that lend themselves to reproducible repeated measurements. (If the largest lesion does not lend itself to reproducible measurement, the next largest lesion which can be measured reproducibly should be selected.)

The sum of the diameters (longest axis for non-nodal lesions, short axis for nodal lesions) for all target lesions will be calculated and recorded as the baseline sum diameter.

A-6 Non-target lesions

All other lesions (or sites of disease) including pathological lymph nodes should be identified as non-target lesions and should also be recorded at baseline. It is possible to record multiple nontarget lesions involving the same organ as a single item on the case record form (e.g. 'multiple enlarged pelvic lymph nodes' or 'multiple liver metastases').

A-7 Response Criteria

Evaluation of target lesions

Complete Response (CR):

Disappearance of all (non-lymph node) target lesions. Any pathological lymph nodes (whether target or non-target) must have reduction in the short axis to <10 mm.

Partial Response (PR):

At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters.

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.

Stable Disease (SD):

Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameter while on study.

Not Evaluable (NE):

No imaging/measurement is done at all at the time points, or CR, PR, PD and SD are not applicable.

$$\text{Percent decrease in the sum of diameters} = \frac{\text{Baseline sum} - \text{Sum at the evaluation}}{\text{Baseline sum}} \times 100\%$$

$$\text{Percent increase in the sum of diameters} = \frac{\text{Sum at the evaluation} - \text{Smallest sum}}{\text{Smallest sum}} \times 100\%$$

* All lesions recorded at baseline should have their actual measurements recorded at each subsequent evaluation, even when very small (e.g., <5 mm). However, when target lesions are "too small to measure," and if the lesion has likely disappeared, the measurement should be

recorded as 0 mm, irrespective of CT slice thickness. If the lesion is believed to be present and is faintly seen but too small to measure, a default value of 5 mm should be assigned.

* If the percent decrease meets PR and at the same time the percent increase meets PD, then PD should be selected.
* When lesions fragment, the longest diameters of the fragmented portions should be added together to calculate the target lesion sum.
* When lesions coalesce, a plane between them may be maintained that would aid in obtaining maximal diameter measurements of each individual lesion. If the lesions have truly coalesced such that they are no longer separable, the vector of the longest diameter in this instance should be the maximal longest diameter for the coalesced lesion.

Evaluation of non-target lesions

Complete Response (CR):

Disappearance of all non-lymph node, non-target lesions and normalization of tumor marker level (i.e., ≤ upper normal limit of the study site). All lymph nodes must be non-pathological in size (< 10mm short axis).

Non-CR/Non-PD:

Persistence of one or more non-lymph node, non-target lesions and/or maintenance of the tumor marker level above the normal limits of the study site.

Progressive Disease (PD):

"Unequivocal progression" of existing non-target lesions (including recurrence).

To achieve "unequivocal progression", there must be an overall level of substantial worsening in non-target lesions such that the overall tumor burden has increased sufficiently to merit discontinuation of study treatment. In the presence of SD or PR in target lesions, "unequivocal progression" will be selected if there is substantial worsening in non-target lesions beyond the extent of reduction in target lesion burden. If not, "Non-CR/Non-PD" should be selected.

When the patient has no target lesions, "unequivocal progression" will be selected if, as a rough guide, the increase in overall non-target lesion burden is apparently greater than a 20% increase in diameter or a 73% increase in volume.

Not Evaluable (NE):

No imaging/measurement is done at all at the time points, or CR, PR, PD and SD are not applicable.

New lesions

A lesion that was not present at baseline but identified after the start of study treatment is regarded as a “new lesion.”

The finding of a “new lesion” should be unequivocal: i.e. not attributable to differences in scanning technique or changes in imaging modality between baseline and later evaluation, or findings thought to represent something other than a tumor. A lesion identified in an anatomical location that was not scanned at baseline is considered a new lesion.

If a lesion disappears and reappears at a subsequent time point, it should continue to be measured. However, the patient’s response at the time point when the lesion reappears will depend upon the status of his/her other lesions. If the patient’s tumor has reached CR and the lesion reappears, then the patient should be considered PD at the time of reappearance. In contrast, if the tumor status is PD or SD and a lesion which had disappeared then reappears, its maximal diameter should be added to the sum of the remaining lesions for a calculated response: in other words, the reappearance of an apparently “disappeared” single lesion amongst many which remain is not in itself enough to qualify for PD: that requires the sum of all lesions to meet the PD criteria. The rationale for such a categorization is based upon the realization that most lesions do not actually “disappear” but are not visualized because they are beyond the resolving power of the imaging modality employed.

When a lesion might be a new lesion but this is unclear, repeat imaging will be performed after a clinically appropriate interval. If the repeat imaging confirms that there is definitely a new lesion, the date of appearance of the new lesion will be the date of the repeat imaging that confirmed the appearance of a new lesion.

A-8 Overall response

Overall response will be assessed on the basis of the response by target lesions and non-target lesions as well as any appearance of new lesions. Table A below provides the criteria for the overall response status determination at each time point for patients who have target lesions at baseline. If the patient has no non-target lesions, the overall response will be determined based on the response by target lesions and any appearance of new lesions. For patients without target lesions at baseline, the overall response will be determined according to Table B, on the basis of the response by non-target lesions and any appearance of new lesions.

Table A: Overall Response Criteria (in patients with target lesions)

Target lesions	Non-target lesions	New lesions	Overall response
CR	CR	No	CR
CR	Non-CR/Non-PD	No	PR
CR	NE	No	PR
PR	Non-PD or NE	No	PR
SD	Non-PD or NE	No	SD
NE	Non-PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

Table B: Overall Response Criteria (in patients without target lesions)

Non-target lesions	New lesions	Overall response
CR	No	CR
Non-CR/Non-PD	No	Non-CR/Non-PD
NE	No	NE
Unequivocal PD	Yes or No	PD
Any	Yes	PD

A-9 Best overall response

Best overall response will be determined on the basis of the overall response up to the end of the study, but in this study it will require confirmation of CR and PR, and will be assessed according to the criteria in Table C. The following will not be taken into account when assessing the best overall response: overall response after initiation of post-study treatment for esophageal cancer, overall response after the end of the follow-up investigation, and overall response after determination of PD. In this study, the best overall response will be assessed as SD in subjects without an overall response of PD until after Week 6 from the start of Cycle 1 and with SD or a better response at least once after Week 6 from the start of Cycle 1. In order for the best overall response to be rated as CR, 2 or more successive CR ratings should be obtained with intervals of not less than 4 weeks (28 days). In order for the best overall response to be rated as PR, 2 or more successive PR or better ratings (CR of PR) should be obtained with intervals of not less than 4 weeks (28 days). If the imaging assessment prior to the start of study treatment showed no target lesion, best overall response of the subject should be assessed as “NE”.

Table C: Best Overall Response Criteria

Overall response First time point	Overall response Subsequent time point	Best overall response
CR	CR	CR
CR	PR	SD, PD or PR[#]
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

#: If CR is truly obtained at the first time point, then any disease seen at a subsequent time point, even a disease meeting PR criteria relative to baseline, makes the disease PD at that point (since the disease must have reappeared after CR). The best response would depend on whether the minimum duration for SD was met. However, sometimes CR may be claimed when subsequent scans suggest that 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 would be PR.

Appendix 3 Performance Status Score (ECOG)

PS 0	Fully active, able to continue all pre-disease activities without restriction
PS 1	Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g., light housework, office work
PS 2	Ambulatory and capable of all selfcare but unable to carry out any work activities. Up and about more than 50% of waking hours
PS 3	Capable of only limited selfcare, confined to bed or a chair more than 50% of waking hours
PS 4	Completely disabled. Cannot carry out any selfcare. Totally confined to bed or a chair
PS 5	Dead

As published in Am J. Clin. Oncol.: Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, Carbone PP. Toxicity and response criteria of the Eastern Cooperative Oncology Group. Am J Clin Oncol. 1982 Dec;5(6):649-55.

Appendix 4

List of Autoimmune Diseases

Patients will be excluded from the study in principle if they concurrently have any of the autoimmune diseases listed below or have a history of chronic or recurrent episodes thereof. For patients with autoimmune diseases other than those listed below, their eligibility should be carefully assessed. Patients with safety concerns posed by the presence of an autoimmune disease should not be included in the study.

● Acute disseminated encephalomyelitis	● IgA nephropathy	● Addison's disease
● Inflammatory bowel disease	● Alopecia universalis	● Cystitis interstitial
● Ankylosing spondylitis	● Lambert-Eaton myasthenic syndrome	● Antiphospholipid syndrome
● Erythematosus	● Aplastic anemia	● Lyme disease (chronic)
● Asthma	● Meniere's syndrome	● Autoimmune hemolytic anemia
● Mooren ulcer	● Autoimmune hepatitis	● Morphoea
● Autoimmune hypophysitis	● Multiple sclerosis	● Autoimmune hypoparathyroidism
● Myasthenia gravis	● Autoimmune myocarditis	● Neuromyotonia
● Autoimmune oophoritis	● Opsoclonus myoclonus syndrome	● Autoimmune orchitis
● Optic neuritis	● Autoimmune thrombocytopenic purpura	● Ord's thyroiditis
● Behcet's disease	● Pemphigus	● Bullous pemphigoid
● Pernicious anemia	● Celiac disease	● Polyarteritis nodosa
● Chronic fatigue syndrome	● Polyarthritis	● Chronic inflammatory demyelinating polyradiculoneuropathy
● Autoimmune polyglandular syndrome	● Churg Strauss syndrome	● Primary biliary cirrhosis
● Crohn's disease (gastrointestinal ulcers)	● Psoriasis	● Dermatomyositis
● Reiter's syndrome	● Type 1 diabetes mellitus	● Rheumatoid arthritis
● Dysautonomia	● Sarcoidosis	● Eczema
● Scleroderma	● Sjogren's syndrome	● Acquired epidermolysis bullosa
● Stiff-man syndrome	● Pemphigoid gestationis	● Takayasu's arteritis
● Giant cell arteritis	● Ulcerative colitis	● Goodpasture's syndrome
● Graves' disease	● Vogt-Koyanagi-Harada syndrome	● Guillain-Barre syndrome
● Vulvodynia	● Hashimoto's disease	● Wegener's granulomatosis
● Kawasaki's disease		

Person responsible for implementation of the study

Protocol change history (Protocol No. ONO-4538-24/BMS CA209473)

Version 1.0: Prepared on July 27, 2015

Version 2.0: Prepared on August 25, 2015

Version 3.0: Prepared on October 26, 2015

Version 4.0: Prepared on January 22, 2016

Version 5.0: Prepared on May 9, 2016.

Version 6.0: Prepared on August 30, 2016

Version 7.0: Prepared on November 30, 2016

Version 8.0: Prepared on July 12, 2017

Version 9.0: Prepared on November 7, 2017

ONO-4538 PHASE III STUDY

A MULTICENTER, RANDOMIZED, OPEN-LABEL STUDY IN PATIENTS WITH ESOPHAGEAL CANCER REFRACTORY OR INTOLERANT TO COMBINATION THERAPY WITH FLUOROPYRIMIDINE AND PLATINUM-BASED DRUGS

STATISTICAL ANALYSIS PLAN

ONO PHARMACEUTICAL CO., LTD.

Protocol No.: ONO-4538-24/BMS CA209473

Version No.: 3.0

Date of preparation: Dec 17, 2018

TABLE OF CONTENTS

LIST OF ABBREVIATIONS 6

1 DEFINITIONS OF STATISTICAL ANALYSIS PLAN 9

2 SAMPLE SIZE 10

3 DETERMINATION OF SUBJECTS INCLUDED IN ANALYSIS 12

3.1 Definitions of Analysis Sets 12

3.2 Criteria for Handling of Subjects 13

4 SIGNIFICANCE LEVEL TO BE USED 15

5 GENERAL METHODS 16

6 RELIABILITY OF THE STUDY 17

6.1 Analysis Set 17

6.2 Analytical Items and Data Handling 17

6.3 Analytical Methods 18

7 EXTENT OF EXPOSURE AND ADMINISTRATION OF STUDY TREATMENT 19

7.1 Analysis Set 19

7.2 Analytical Items and Data Handlings 19

7.3 Analytical Methods 21

8 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS 22

8.1 Analysis Set 22

8.2 Analytical Items and Data Handlings 22

8.3 Analytical Methods and Classification 23

9 PRIMARY EFFICACY ANALYSIS 29

9.1 Hypothesis 29

9.2 Analysis Sets 29

9.3 Analytical Item and Data Handling 29

9.4 Analytical Methods 29

9.4.1 Primary Analytical Method 29

9.4.2 Sensitivity Analysis for Primary Analytical Method 30

9.4.3 Secondary Analytical Methods 30

9.4.4 Exploratory Analytical Methods 31

10 SECONDARY EFFICACY ANALYSIS 33

10.1 Analysis Set 33

10.2 Analytical Items and Data Handlings 33
10.3 Analytical Methods 38
10.4 Subgroup analyses for the efficacy endpoints....40
10.5 Interaction Analysis 43
10.6 Adjusted Analysis 44
10.7 Relationship Between Endpoints 44
10.8 Other Analysis....44
11 ANALYSIS OF SAFETY 45
11.1 Analysis Sets 45
11.2 Analytical Items and Data Handlings 45
11.3 Analytical methods 51
11.4 Subgroup analyses for the safety endpoints 60
12 PD-L1 EXPRESSION ANALYSIS 62
12.1 Analysis Sets 62
12.2 Analytical Item and Data Handlings 62
12.3 Analytical methods 62
13 ANTI-DRUG ANTIBODY ANALYSIS 64
13.1 Analysis Sets 64
13.2 Analytical Item and Data Handlings 64
13.3 Analytical methods 66
14 PATIENT REPORTED OUTCOMES ANALYSIS 67
14.1 Analysis Sets 67
14.2 Analytical Items and Data Handling 67
14.3 Analytical methods 68
15 INTERIM STATISTICAL ANALYSIS....70
16 DATA REVIEWS 71
16.1 Objectives....71
16.2 Analysis Sets 71
16.3 Analytical Items and Data Handling 71
16.4 Analytical Methods 72
17 TECHNICAL MATTERS 73
17.1 Statistical Software 73
17.2 SAS Programming in Analyses....73

17.3 Data Format 74
18 MAJOR CHANGES IN THE STATISTICAL ANALYSIS PLAN BASED ON THE STUDY-SPECIFIC PROTOCOL 77
19 MAJOR CHANGES IN THE STATISTICAL ANALYSIS PLAN BASED ON THE STATISTICAL ANALYSIS PLAN VERSION 1.0 78
19.1 Reliability of the study 78
19.2 Demographics and other baseline characteristics 78
19.3 Primary efficacy analysis 78
19.3.1 Proportional hazards assumption 78
19.3.2 State of follow-up 79
19.3.3 Exploratory analytical methods 79
19.4 Secondary efficacy analysis 79
19.5 Analysis of safety 79
19.5.1 Adverse event 79
19.6 PD-L1 expression analysis 80
19.7 Patient reported outcomes analysis 80
20 MAJOR CHANGES IN THE STATISTICAL ANALYSIS PLAN BASED ON THE STATISTICAL ANALYSIS PLAN VERSION 2.0 81
20.1 Significance level to be used 81
20.2 Secondary efficacy analysis 81
APPENDIX 1 82
Definition of Select Adverse Events 82
APPENDIX 2 86
Definition of Immune-Mediated Adverse Events 86
APPENDIX 3 87
Definition of Other Adverse Events of Special Interests 87

LIST OF TABLES

Table 10-2-1: Criteria for event/censored assessment in primary definition of PFS 34
Table 10-2-2: Criteria for event/censored assessment in secondary definition of PFS 35
Table 10-2-3: Criteria for event/censored in DOR 36
Table 14-2-1: Time windows for EQ-5D Assessments 68

LIST OF ABBREVIATIONS

ADA	: Anti-Drug Antibody Set
AE	: Adverse Event
ALP	: Alkaline Phosphatase
ALT (GPT)	: Alanine aminotransferase (glutamic pyruvic transaminase)
AST (GOT)	: Aspartate aminotransferase (glutamic oxaloacetic transaminase)
BMI	: Body Mass Index
BOR	: Best Overall Response
BSA	: Body Surface Area
CI	: Confidence Interval
CR	: Complete Response
CTCAE	: Common Terminology Criteria for Adverse Events
DCR	: Disease Control Rate
DOR	: Duration of Response
ECG	: Electrocardiogram
ECOG	: Eastern Cooperative Oncology Group
eCRF	: Electronic Case Report Form
EMR	: Endoscopic Mucosal Resection
ENR	: Enrolled Set
EQ-5D	: EuroQol 5 dimension
EQ-VAS	: EuroQol visual analogue scale
ESD	: Endoscopic Submucosal Dissection
Free T3	: Free Triiodothyronine
Free T4	: Free Thyroxine

GCP : Good Clinical Practice

HR : Heart Rate

IEC : Independent Ethics Committee

IMAE : Immune-Mediated Adverse Event

INF : Informed Consent Set

IRB : Institutional Review Board

ITT : Intention-to-Treat

IWRS : Interactive Web Response System

LLN : Lower Limit of Normal

MedDRA : Medical Dictionary for Regulatory Activities

NCI : National Cancer Institute

NE : Not Evaluable

ORR : Objective Response Rate

OS : Overall Survival

PD : Progressive Disease

PD-L1 : Programmed Cell Death-Ligand 1

PFS : Progression-Free Survival

PP : Persistent Positive

PR : Partial Response

PR Interval : PR interval on an electrocardiogram (atrioventricular conduction time)

PT : Preferred Term

QRS Width : QRS width on an electrocardiogram (ventricular activation time)

QT Interval : QT interval on an electrocardiogram (duration of ventricular electrical systole)

QTcF Interval : Corrected QT interval using the Fridericia's formula

RECIST : Response Evaluation Criteria in Solid Tumors

RES : Response Evaluable Set

RR Interval : RR interval on an electrocardiogram (ventricular rate)

SAF : Safety Set

SAS : Statistical Analysis System software

SD : Stable Disease

SD : Standard Deviation

SOC : System Organ Class

TNM : Tumor Node Metastasis

TSH : Thyroid-Stimulating Hormone

TTR : Time To Response

UK : United Kingdom

ULN : Upper Limit Of Normal

1 DEFINITIONS OF STATISTICAL ANALYSIS PLAN

This statistical analysis plan refers to the definition of terminologies, valid analytical methods and the detailed contents of statistical analysis for ONO-4538 Phase III study.

The statistical analysis plan will be developed in accordance with the protocol for ONO-4538 Phase III study. The statistical analysis plan will be finalized before randomization for the first subject. The analysis principles are outlined in Section 10 “Statistical Analysis” in the study-specific protocol.

If additional analysis not stated in the statistical analysis plan was performed, the result will be reported in the clinical study report with the purpose of analysis.

2 SAMPLE SIZE

195 subjects in each group, totaling 390 subjects

Rationale for the Sample Size

This study is intended to verify the superiority of ONO-4538 group over the control groups (docetaxel or paclitaxel) in terms of OS, the primary endpoint, in patients with esophageal cancer refractory or intolerant to fluoropyrimidine and platinum-based agents, the standard chemotherapy regimen for the treatment of esophageal cancer. Subjects will be randomized to either the ONO-4538 group or the control group in a 1:1 ratio and stratified by region (Japan vs. rest of world) and number of organs with metastases (≤1 vs. ≥2) and PD-L1 expression (≥1% vs. <1% or indeterminate).

As for OS, assuming an exponential distribution for the control group (7.2 months median OS) and a piecewise mixture model for the ONO-4538 group with 5.0% long term survival rate (the ratio that long term survival is expected due to the effect of ONO-4538), the hazard ratio of the ONO-4538 group relative to the control group was assumed to be 1.0 for the first 3 months and 0.65 for non-long term survival population thereafter (10.0 months overall median OS for the ONO-4538 group), the expected average hazard ratio of the ONO-4538 group to the control groups was assumed to be 0.70 in this study. The number of events required to detect superiority of the ONO-4538 group over the control groups with two-sided significance level of 5% and 90% or more power by the log-rank test was calculated to be 331.

Assuming the enrollment period to be 16 months and the follow-up period after the last patient's enrollment to be 18 months, the number of subjects required to ensure the required 331 events was estimated to be 390. For the calculation of the required events and sample size at the time of planning the study, the statistical analysis software SAS (version 9.3) was used.

Rationale for assuming the median OS in the control group to be 7.2 months

The phase II single-arm study of docetaxel and the retrospective study of paclitaxel as 2nd-line treatment for patients with esophageal cancer reported that the median OS of docetaxel and

paclitaxel was 5.5-8.1 months and 6.1-10.4 months, respectively. Based on this information, the median OS in the control groups of this study is assumed to be 7.2 months.

Rationale for assuming long term survival rate (the ratio that long term survival is expected due to the effect of ONO-4538) as 5.0% and the hazard ratio of the ONO-4538 group relative to the control groups was assumed to be 1.0 for the first 3 months

We set long term survival rate and the hazard ratio referring to the results of a randomized phase 3 clinical trial of nivolumab vs investigator's choice in recurrent or metastatic platinum-refractory squamous cell carcinoma of the head and neck (CA209141) and a phase 2 uncontrolled study of nivolumab in esophageal cancer refractory or intolerant to standard therapy (ONO-4538-07).

Rationale for setting the hazard ratio of the ONO-4538 group relative to the control groups was assumed to be 0.65 after first 3 months in non-long term survival population

In the global phase III study conducted in patients with stage IIIB/IV or recurrent squamous non-small cell lung cancer who received a prior platinum-containing chemotherapy regimen (CA209017) to verify the superiority of ONO-4538 over docetaxel, one of the controls in this study, the hazard ratio of ONO-4538 relative to docetaxel was 0.59 [96.85% confidence interval: 0.43, 0.81].

If the same histological type of cancer and control are used in the present study, the efficacy on ONO-4538 is expected to be similar. However, this study includes paclitaxel as well as docetaxel as controls. Efficacy data on ONO-4538 vs. paclitaxel has not been obtained previously. Based on these reasons, the expected average hazard ratio of the ONO-4538 group to the control groups was assumed to be 0.70 in this study.

In this study, since we assumed that 5.0% of long term survival rate and the hazard ratio of the ONO-4538 group relative to the control group was to be 1.0 for the first 3 months, the hazard ratio of the ONO-4538 group relative to the control groups after first 3 months in non-long term survival population was calculated as 0.65 to become the expected average hazard ratio of the ONO-4538 group relative to the control groups was 0.70.

3 DETERMINATION OF SUBJECTS INCLUDED IN ANALYSIS

For efficacy endpoints, the Intention-to-Treat (ITT) population will be the analysis set. The Response Evaluable Set (RES) will also be used for some secondary endpoints.

For safety endpoints, the Safety Set (SAF) will be the analysis set.

For anti-drug antibody endpoint, the Anti-Drug Antibody Set (ADA) will be the analysis set.

Individual sets of subjects are defined below.

3.1 Definitions of Analysis Sets

1) Informed Consent Set (INF)

The INF will consist of all subjects who signed an informed consent.

2) Enrolled Set (ENR)

The Enrolled Set (ENR) will consist of all subjects enrolled (enrollment) in this study via IWRS.

3) ITT

ITT will consist of all randomized (randomization) subjects.

4) SAF

The SAF will consist of all subjects given at least one dose of the investigational product.

5) RES

The RES will consist of all subjects which meet all of the following items in the ITT population.

(1) It is not GCP noncompliant subject defined in 3.2 Criteria for Handling of Subjects.

(2) It has target lesion measurements at baseline.

6) ADA

The ADA will consist of all subjects which meet all of the following items in the SAF population.

(1) It is not GCP noncompliant subject defined in 3.2 Criteria for Handling of Subjects.

(2) It has measurements of anti-ONO-4538-antibody (the sample which screening result is potential positive but confirmatory result is not available will be excluded) at both baseline and at least one post-baseline anti-drug antibody assessment.

3.2 Criteria for Handling of Subjects

Criteria for handling of subjects are specified below.

1) Disqualified Subjects

Disqualified subjects are defined as those who fail to fulfill 4.2 Inclusion Criteria or conflict with 4.3 Exclusion Criteria in the study-specific protocol.

2) GCP Noncompliant Subjects

GCP noncompliant subjects are defined as those who were enrolled based on materials not reviewed by the IRB/IEC, who were enrolled at a medical institution where no study contract had been signed, or who did not provide informed consent appropriately.

3) Untreated Subjects

Untreated subjects are defined as those who received no dose of the investigational product.

4) Incomplete Target Lesion Measurements Subjects

Incomplete target lesion measurements subjects are defined as those who were without baseline target lesion measurements.

5) Non-enrolled Subjects

Non-enrolled subjects are defined as those who signed an informed consent but not enrolled.

6) Non-randomized Subjects

Non-randomized subjects are defined as those who enrolled but not randomized.

7) Incomplete Anti-Drug Antibody Subjects

Incomplete anti-drug antibody subjects are defined as those who were without baseline anti ONO-4538 antibody measurements, and/or all post-baseline anti ONO-4538 antibody measurements.

For any other subject with unexpected problems, the handling in the analysis will be determined by the sponsor before database lock, on the basis of discussion with the medical officer.

4 SIGNIFICANCE LEVEL TO BE USED

1) Efficacy

Two-sided test will be performed with 5% significance level. If superiority in OS is determined, a hierarchical hypothesis testing approach for the key secondary endpoints will be used to preserve a study-wise type I error rate at 5%. The key secondary endpoints will be tested in the following hierarchical order:

1. ORR

2. PFS

The ITT will be the analysis set for OS and PFS, and the RES will be the analysis set for ORR.

Interactions will be tested with two-sided 15% significance level.

2) Safety

Not applicable because no statistical tests will be performed on safety.

5 GENERAL METHODS

Unless otherwise noted, analysis will be conducted based on treatment group randomized by IWRS.

In the following sections, treatment group will refer to ONO-4538 group and control group. Control regimen will refer to docetaxel and paclitaxel. Investigational product will refer to ONO-4538, docetaxel and paclitaxel drugs.

The treatment group “As Randomized” will be retrieved from IWRS. The treatment group “As Treated” will be the same as the arm randomized by IWRS. However, if a subject received the incorrect treatment for the entire period of treatment, the subject’s treatment group will be defined as the incorrect treatment the subject actually received.

Efficacy analysis will be conducted by “As Randomized” while safety analysis will be conducted by “As Treated”.

A term of ONO-4538 will be replace with Nivolumab in tables, figures and listings.

6 RELIABILITY OF THE STUDY

6.1 Analysis Set

The INF will be the analysis set for the analytical Item (2).

The ENR will be the analysis set for the analytical Item (3).

The ITT will be the analysis set for the analytical Item (1), (4) and (6).

The SAF will be the analysis set for the analytical Item (5).

6.2 Analytical Items and Data Handling

1) Analytical Item

(1) Relevant protocol deviations

(2) Reason excluded from each analysis set

(3) Reason for subjects enrolled but not randomized

(4) Reason for subjects randomized but not treated

(5) Reason of discontinuation of study treatment

(6) The number of randomized subjects per study center

2) Handling of data

The following deviations will be considered as relevant protocol deviations.

Eligibility:

- Subjects who failed to fulfill inclusion criteria #3
- Subjects who failed to fulfill inclusion criteria #4
- Subjects who failed to fulfill inclusion criteria #6

On-study:

- Subjects receiving any concurrent anti-cancer therapy (ie. chemotherapy, hormonal therapy, immunotherapy, surgery, or radiation therapy) while on study therapy

- Subjects treated differently as randomized (subjects who received the wrong treatment excluding the never treated.)

6.3 Analytical Methods

1) The number of subjects with relevant protocol deviations will be summarized by treatment group and control regimen.
2) For INF and ENR, the frequency of inclusion and exclusion for each analysis set as well as the frequency of the reason for exclusion will be summarized. For the others, those will be summarized by treatment group and control regimen.
3) The number of subjects enrolled but not randomized as well as the frequency by its reason will be summarized.
4) The number of subjects randomized but not treated as well as the frequency by its reason will be summarized.
5) The number of subjects who continued the treatment period and discontinued study treatment as well as the frequency of the reason for discontinuation of study treatment will be summarized by treatment group and control regimen.
6) For the number of subjects randomized per study center, frequency distribution and summary statistics will be summarized by treatment group and control regimen.

7 EXTENT OF EXPOSURE AND ADMINISTRATION OF STUDY TREATMENT

7.1 Analysis Set

The SAF will be the analysis set.

7.2 Analytical Items and Data Handlings

1) Analytical Items

(1) Number of dose received

(2) Duration of Treatment

(3) Number of cycle

(4) Cumulative dose

(5) Relative dose intensity

2) Handlings of data

(1) Duration of Treatment

Duration of Treatment (Months)

= (“Date of the last dose” − “Date of the first dose” + 1) / 30.4375

(2) Number of cycle

The number of cycle will be calculated for the cycle proceeding to the next cycle. The discontinued cycle or the cycle receiving no investigational product also will be calculated.

(3) Cumulative dose

Cumulative dose of ONO-4538 (mg) is the sum of the doses (mg) administrated to a subject during the treatment period.

Cumulative dose of Docetaxel (mg/m^2) is the sum of the doses (mg/m^2) administrated to a subject during the treatment period.

Dose of Docetaxel (mg/m^2) will be calculated by the following equation.

Dose (mg/m^2) = "Total Dose administrated of Docetaxel (mg)" / "Most recent BSA [NOTE1)]"

Cumulative dose of Paclitaxel (mg/m^2) is the sum of the doses (mg/m^2) administrated to a subject during the treatment period.

Dose of Paclitaxel (mg/m^2) will be calculated by the following equation.

Dose (mg/ m^2) = "Total Dose administrated of Paclitaxel (mg)" / "Most recent BSA[NOTE1)]"

NOTE1: Body Surface Area (BSA) (DuBois Formula)

BSA="[Weight (kg)]$^{0.425}$" x "[Height (cm)]$^{0.725}$ "x 0.007184

(4) Relative dose intensity

Relative dose intensity of ONO-4538 (%) will be calculated by the following equation.

Relative dose intensity (%) = "Cumulative dose (mg)" / "[Date of the last dose – Date of the first dose + 14] (days) x240(mg) / 14 (days)" x 100

Relative dose intensity of Docetaxel (%) will be calculated by the following equation.

Relative dose intensity (%) = "Cumulative dose (mg/m^2)" / "[Date of the last dose – Date of the first dose + 21] (days) x 75 (mg/m^2) / 21 (days)" x 100

Relative dose intensity of Paclitaxel (%) will be calculated by the following equation.

Relative dose intensity (%) = "Cumulative dose (mg/m^2)" / "The sum of defined dose in each cycle (mg/m^2)" x 100

For a completed cycle, defined dose is defined as 6 x 100 (mg/m^2) . For discontinued cycle, defined dose is defined as "{[Date of last dose – Start date of discontinued cycle + 7] (days) x 100 (mg/m^2)}/7 (days)".

7.3 Analytical Methods

1) For number of dose received, summary statistics will be calculated by investigational product.
2) For duration of treatment, summary statistics will be calculated by treatment group and control regimen. Additionally, frequency distribution of subjects on duration of treatment for > 6 months and > 12 months will be calculated by treatment group and control regimen.
3) For number of cycle, frequency distributions and summary statistics will be calculated by treatment group and control regimen.
4) For cumulative dose, summary statistics will be calculated by investigational product.
5) For relative dose intensity, frequency distributions and summary statistics will be calculated by treatment group and control regimen.

8 DEMOGRAPHICS AND OTHER BASELINE CHARACTERISTICS

8.1 Analysis Set

The ITT will be the analysis set.

8.2 Analytical Items and Data Handlings

1) Analytical Items

(1) Demographic Variables

Sex, age, race, ethnicity, height, weight, BMI, BSA

(2) Baseline Characteristics

Performance status score (ECOG), time from the date of diagnosis of the primary disease, lesion sites (anatomical subsites according to the TNM classification), clinical classification (TNM clinical classification), disease stage (stage grouping: TNM classification), histological classification, presence of multiple lesions (secondary lesions), sites of metastases, recurrent, time from the date of surgery for the primary lesion (including EMR or ESD) to randomization (for subjects with recurrent esophageal cancer), residual tumor (R) classification (for subjects with recurrent esophageal cancer), site of recurrence (for subjects with recurrent esophageal cancer), Medical history, concurrent diseases, past treatments for esophageal cancer (e.g., surgery, radiotherapy, pharmacotherapy), history of alcohol consumption, history of smoking, geographic location (IWRS and eCRF source), number of organs with metastases (IWRS and eCRF source), PD-L1 expression (IWRS and test result source)

(3) Baseline Values of Observation

Sum of reference diameters of target lesions

(4) Concomitant drugs and therapies

Analgesics, Immune modulating concomitant medication

(5) Subsequent anti-cancer therapy

Surgery, Radiotherapy, Pharmacotherapy by generic name

2) Handlings of data

(1) BMI

BMI (kg/m^2) = “Weight (kg)” / “[Height (m)]^2”

(2) Body Surface Area (BSA) (DuBois Formula)

BSA=“[Weight (kg)]$^{0.425}$” x “[Height (cm)]$^{0.725}$ ”x 0.007184

(3) Time from the date of diagnosis of the primary disease to the date of randomization

Time from the date of diagnosis of the primary disease to the date of randomization (Months)

= (“Date of randomization” − “The most recent date of diagnosis of the primary disease” + 1) / 30.4375

(4) Time from the date of surgery for the primary lesion (including EMR or ESD) to the date of randomization

Time from the date of surgery for the primary lesion to the date of randomization (Months)

= (“Date of randomization” − “The most recent date of surgery for the primary lesion” + 1) / 30.4375

(5) Immune modulating concomitant medication

When calculaing the number of subjects who received immune modulating concomitant medication for management of adverse event and drug-related select adverse event, adverse events and select adverse events to be analyzed will be the events occurring between the start date of the first administration of the investigational product and 28 days after the last dose or the start date of subsequence anti-cancer therapy after the last dose whichever comes first.

When caclulaing the number of subjects who received immune modulating concomitant medication for management of IMAE, IMAE to be analyzed will be the events occurring between the start date of the first administration of the investigational product and 100 days after the last dose.

8.3 Analytical Methods and Classification

Each analytical item displayed in 8.2 Analytical Items and Data Handling will be summarized according to the following classification by treatment group and control regimen.

<Method of classification >

1) Classification for the ordinal scale or the continuous value will be determined based on clinical symptom or equipartition of the number of subjects (i.e., 1/3, 1/4, 1/5).
Each classification of analytical item will be displayed in the following table.

Analytical Item	Stratum	Analytical Method
Sex	Female, Male	Frequency distributions
Age[Note1)]	<65, 65 - <75, >=75, >=65	Summary statistics, Frequency distributions
Race	American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White, Other	Frequency distributions
Ethnicity	Hispanic or Latino, Not Hispanic or Latino, Not Reported, Unknown	Frequency distributions
Height	＜Will be revised under the planned Data Reviews＞	Summary statistics, Frequency distributions
Weight	＜Will be revised under the planned Data Reviews＞	Summary statistics, Frequency distributions
BMI	＜Will be revised under the planned Data Reviews＞	Summary statistics, Frequency

		distributions
BSA	＜Will be revised under the planned Data Reviews＞	Summary statistics Frequency distributions
Performance Status score (ECOG)	0, 1	Frequency distributions
Time from the date of diagnosis of the primary disease to randomization	<=12, >12 - 24, >24	Summary statistics, Frequency distributions
Lesion sites (TNM classification)	Cervical Esophagus, Thoracic Esophagus (Upper Thorax, Middle Thorax, Lower Thorax), Unknown	Frequency distributions
Clinical classification (TNM classification) Note2)	T: Primary tumor TX, T0, Tis, T1a, T1b, T2, T3, T4a, T4b, Unknown N: Regional lymph nodes NX, N0, N1, N2, N3, Unknown M: Distant metastasis M0, M1, Unknown	Frequency distributions
Disease stage (TNM classification) Note2)	0, IA, IB, IIA, IIB, IIIA, IIIB, IIIC, IV, Unknown	Frequency distributions
Histological classification	Squamous Cell Carcinoma, Adenosquamous Cell Carcinoma, Other, Unknown	Frequency distributions

Presence of multiple lesions (secondary lesions)	No, Yes	Frequency distributions
Sites of metastases	Lymph Node, Peritoneum, Liver, Lung, Pleural Tissue, Adrenal Gland, Brain, Bone, Bone Marrow, Skin, Stomach, Other	Frequency distributions
Recurrent	No, Yes	Frequency distributions
Time from the date of surgery for the primary lesion (including EMR or ESD) to randomization[Note3)]	<=12, >12 - 24, >24	Summary statistics, Frequency distributions
Residual tumor (R) classification[Note3)]	RX, R0, R1, R2, Unknown	Frequency distributions
Site of recurrence[Note3)]	Esophagogastric, Cervical Esophagogastric, Thoracic Esophagus (Upper Thorax, Middle Thorax, Lower Thorax, Unknown), Lymph Node, Stomach, Peritoneum, Liver, Lung, Pleural Tissue, Adrenal Gland, Brain, Bone, Bone Marrow, Skin, Other	Frequency distributions
Medical history	No, Yes	Frequency distributions

Concurrent diseases	No, Yes	Frequency distributions
Past treatments for cancer (surgery)	No, Yes	Frequency distributions
Past treatments for cancer (radiotherapy)	No, Yes	Frequency distributions
Past treatments for cancer (pharmacotherapy)	No, Yes	Frequency distributions
History of alcohol consumption	Never, Former, Current	Frequency distributions
History of smoking	Never, Former, Current	Frequency distributions
Sum of reference diameters of target lesions		Summary statistics
Analgesics	No, Yes	Frequency distributions
Geographic location (IWRS source)	Japan, Rest of the world	Frequency distributions
Geographic location (eCRF source)	Japan, Rest of the world	Frequency distributions
Number of organs with metastases (IWRS source)	<=1, >=2	Summary statistics, Frequency distributions

PD-L1 (IWRS source)	>=1%, <1% or Indeterminate	Frequency distributions
PD-L1 (test results source)	(1) >=1%, <1% (2) >=5%, <5% (3) >=10%, <10% (4) Not Quantifiable	Frequency distributions

Note1): Additional categories 75-<85 and >=85 will be reported if there are subjects >=85 years old

Note2): Summarized for subjects with non-recurrent esophageal cancer.

Note3): Summarized for subjects with recurrent esophageal cancer.

1) Past pharmacotherapy for cancer will be summarized for each treatment group and control regimen by medication class and generic name.
2) Details of subsequence anti-cancer therapy (surgery, radiotherapy, pharmacotherapy) will be summarized by treatment group and control regimen.
3) For immune modulating concomitant medication, the number of subjects who received immune modulating concomitant medication for management of adverse event, management of concurrent disease, management of drug-related select adverse event (any grade, grade 3-5), management of IMAEs (any grade, grade 3-5), prevention, other use, and any use will be calculated separately for each treatment group by medication class and generic name.
4) Cross-tabulation between stratification factors (IWRS source) and stratification factors (eCRF source) will be provided by treatment group and control regimen.

(1) Number of organs with metastases (IWRS source) vs Number of organs with metastases (eCRF source)

(2) PD-L1 expression (IWRS source) vs PD-L1 expression (test result source)

9 PRIMARY EFFICACY ANALYSIS

9.1 Hypothesis

The ONO-4538 group is superior to the control group in terms of OS.

9.2 Analysis Sets

The ITT will be the primary analysis set. The primary analytical method will be conducted for the SAF, if necessary.

9.3 Analytical Item and Data Handling

1) Analytical Item

OS

2) Handling of data

OS will be calculated from the following equation:

OS (Months) = (“Date of death due to any cause” − “Date of randomization” + 1) / 30.4375

For subjects lost to follow-up and subjects who are alive at the time of data cutoff date, data will be censored at the time the subject was last confirmed to be alive.

9.4 Analytical Methods

9.4.1 Primary Analytical Method

Data will be compared between the two treatment groups using the stratified log-rank test with the location (Japan vs. Rest of the world), the number of organs with metastases (1 organ or less vs. 2 organs or more), and PD-L1 expression (≥1% vs. <1% or indeterminate) which are the randomization factors from IWRS as the stratification factors.

9.4.2 Sensitivity Analysis for Primary Analytical Method

1) The distribution of OS will be compared between the two treatment groups by using the unstratified log-rank test.
2) The distribution of OS will be compared between the two treatment groups by using the stratified log-rank test adjusted by the three stratification factors (location and the number of organs with metastases from eCRF source and PD-L1 expression from test result source), if there exist more than 10% discrepancy between stratification factors (IWRS source) and stratification factors (eCRF source).

9.4.3 Secondary Analytical Methods

1) The hazard ratio and the corresponding two-sided 95% confidence interval (CI) for the ONO-4538 group relative to the control group, docetaxel group and paclitaxel group will be estimated using a stratified Cox proportional hazards model with the randomization factors as the stratification factors.
2) To examine the assumption of proportional hazards in the stratified Cox proportional hazards model described in 9.4.3 1), in addition to treatment, a time-dependent variable defined by treatment by time interaction will be added into the model. A two-sided Wald Chi-square p-value of less than 0.1 may indicate a potential nonconstant treatment effect. In that case, additional exploratory analyses may be performed.
3) The Kaplan-Meier curve will be plotted for each treatment group and control regimen. Using the Kaplan-Meier method, the median OS and the corresponding two-sided 95% CI will be estimated for each treatment group and control regimen. CI for median OS will be calculated based on a log-log transformed CI for the survivor function.
4) For each treatment group and control regimen, the survival rates at months 6, 9, 12, 15 and 18 will be derived from the Kaplan-Meier method and corresponding two-sided 95% confidence interval will be calculated based on Greenwood's formula for variance derivation and on log-log transformation applied on the survivor function.
5) The number of events and censors will be tabulated by treatment group and control regimen. The status of subjects who are censored in the Kaplan-Meier analysis will also be tabulated by treatment groups and control regimens using the following categories: (1) on-study, (2) off-study.

6) Current status of follow-up (time from date of last contact to cutoff date) for OS will be summarized for each treatment group and control regimen by using following categories: 0, >0 - 3 months, >3 - 6 months, >6 - 9 months, >9 - 12 months, and >12 months. Subjects who died, or whose last known alive date is on or after data cutoff date, are classified as having current follow-up 0.
7) Summary statistics (median, minimum and maximum) for duration of follow-up (time between randomization date and last known date alive (for subjects who are alive) or date of death) will be summarized by treatment group and control regimen.

9.4.4 Exploratory Analytical Methods

1) Intended investigator's choice therapy subset analysis

Per IWRS procedure, investigators are to indicate their choice of therapy (docetaxel or paclitaxel) for each patient prior to randomization. To further examine the efficacy of ONO-4538 compare to docetaxel and paclitaxel, following analysis will be performed on OS.

(1) The Kaplan-Meier curve will be plotted for each intended investigator's choice therapy subset. Using the Kaplan-Meier method, the median OS and the corresponding two-sided 95% CI will be estimated for each subset. CI for median OS will be calculated based on a log-log transformed CI for the survivor function.
(2) The hazard ratio and the corresponding two-sided 95% CI for the ONO-4538 group relative to the docetaxel group and paclitaxel group in each intended investigator's choice subset will be estimated using a stratified Cox proportional hazards model with the randomization factors as the stratification factors.

2) Examine influence of subsequence anti-cancer therapy

To examine the influence of subsequence anti-cancer therapy on OS from various perspectives, following analysis will be performed on OS, if necessary.

- Rank-preserving structural failure time model
- Inverse probability of censoring weighted methods

- Stratified Cox proportional hazards model with the randomization factors from IWRS as the stratification factors and subsequence anti-cancer therapy as the time-dependent covariate

10 SECONDARY EFFICACY ANALYSIS

10.1 Analysis Set

The ITT will be the analysis set. With regard to the analytical item (7), the RES will be the analysis set. With regard to the analytical item (1), (2), (4), (5) and (6), the RES also will be the analysis set as well as ITT and the primary population for the analytical item (1), (2), (4), (5) and (6) will be RES.

10.2 Analytical Items and Data Handlings

1) Analytical Item

(1) Objective response rate (ORR)

(2) Disease control rate (DCR)

(3) Progression-free survival (PFS)

(4) Duration of response (DOR)

(5) Time to response (TTR)

(6) Best overall response (BOR)

(7) Maximum percent change from baseline in the sum of diameters of target lesions

2) Handlings of data

Overall response and best overall response will be determined solely by imaging assessment according to the RECIST Guideline Version 1.1, and will not take into account any clinical/symptomatic progression. Evaluable imaging data will be those without an overall response of "NE."

For each analytical item, the handling of data will be as follows:

(1) ORR

ORR is defined as the percentage of subjects whose best overall response (BOR) is assessed as either confirmed CR or confirmed PR.

(2) DCR

DCR is defined as the percentage of subjects whose BOR is assessed as confirmed CR, confirmed PR or SD.

(3) PFS

PFS will be calculated from the following equation:

PFS (Months) = ("Earlier date on which either the overall response was assessed as PD or the patient died due to any cause" − "Date of randomization" + 1) / 30.4375

Event / censored assessment in primary and secondary definition of PFS will be performed according to Table 10-2-1 and Table 10-2-2.

In primary definition of the PFS, if the date on which the overall response is assessed as PD for the first time and the date of the start of subsequence anti-cancer therapy are the same, it will be handled as event on the date on which the overall response is assessed as PD. Additionally, if the date of tumor assessment which overall response is not PD or NE and the date of the start of subsequence anti-cancer therapy are the same, PFS will be censored at that date.

When a subject corresponds to more than one situation, the shortest PFS will be handled as the PFS of the subject.

Table 10-2-1: Criteria for event/censored assessment in primary definition of PFS

Situation	Date of Progression or Censoring	Outcome
No baseline tumor assessments	Date of randomization	Censored
No death without an overall response of PD	Date of the last evaluable tumor assessment	Censored
No death without an evaluable tumor assessment	Date of randomization	Censored

Death without an evaluable tumor assessment	Date of death	Event
Subsequence anti-cancer therapy before an overall response of PD or death	Date of the last evaluable tumor assessment before initiation of subsequence anti-cancer therapy	Censored
End of investigating subsequence anti-cancer therapy before an overall response of PD or death	Date of the last evaluable tumor assessment before the end of investigating subsequence anti-cancer therapy	Censored
With an overall response of PD	Date of the first overall response of PD	Event
Death	Date of death	Event

Table 10-2-2: Criteria for event/censored assessment in secondary definition of PFS

Situation	Date of Progression or Censoring	Outcome
No baseline tumor assessments	Date of randomization	Censored
No death without an overall response of PD	Date of the last evaluable tumor assessment	Censored
No death without evaluable tumor assessment	Date of randomization	Censored
Death without an evaluable tumor assessment	Date of death	Event

With an overall response of PD	Date of the first overall response of PD	Event
Death	Date of death	Event

(4) DOR

DOR will be calculated from the following equation.

DOR (Months) = ("Earlier date on which either the overall response was assessed as PD for the first time after confirmed response or the patient died due to any cause" − "Date of first assessment of confirmed CR or PR" + 1) / 30.4375

Event / censored assessment in analysis of DOR will be performed according to Table 10-2-3. If the date on which the overall response is assessed as PD for the first time and the date of the start of subsequence anti-cancer therapy are the same, it will be handled as event on the date on which the overall response is assessed as PD.

When a subject corresponds to more than one situation, the shortest DOR will be handled as the DRO of the subject.

DOR will be calculated in subjects whose BOR was assessed as either confirmed CR or confirmed PR.

Table 10-2-3: Criteria for event/censored in DOR

Situation	Date of Progression or Censoring	Outcome
Death without an evaluable diagnostic imaging after the confirmation of response	Date of death	Event
Death before an overall response of PD	Date of death	Event

Overall response of PD	Date of the first overall response of PD	Event
No death without an overall response of PD	Date of the last evaluable diagnostic imaging	Censored
Subsequence anti-cancer therapy before an overall response of PD or death	Date of the last evaluable diagnostic imaging before initiation of subsequence anti-cancer therapy	Censored
End of investigating subsequence anti-cancer therapy before an overall response of PD or death	Date of the last evaluable diagnostic imaging before end of investigating subsequence anti-cancer therapy	Censored

(5) TTR

TTR will be calculated from the following equation:

TTR (Months) = ("Date of first assessment of confirmed CR or confirmed PR" − "Date of randomization" + 1) / 30.4375

For subjects without a response in each treatment group, the data on the time to response will be censored as at "the longest time to response among subjects with response in the treatment group + 1 (day)".

(6) Maximum percent change from baseline in the sum of diameters of target lesions

In subjects with target lesions, on the basis of the diameters of target lesions as measured according to the RECIST Guideline Version 1.1, the maximum percent change from baseline in the sum of diameters of target lesions will be calculated from the following equation. However, this calculation will exclude the diameter data obtained after an overall

response of PD, after start of subsequence anti-cancer therapy and after end of investigating subsequence anti-cancer therapy.

$$\text{Maximum percent change from baseline in the sum of diameters of target lesions (\%)} = \left(\frac{\text{"Smallest sum of diameters of target lesions"}}{\text{"Baseline sum of diameters of target lesions}} - 1 \right) \times 100$$

10.3 Analytical Methods

The following analyses will be performed on the analytical items specified in 10.2 Analytical Items and Data Handling.

1) Analytical items for ORR

(1) Data will be compared between the two treatment groups using the Cochran-Mantel-Haenszel (CMH) test with the randomization factors as the stratification factors (IWRS). The associated odds ratio and the corresponding two-sided 95% CI and the estimate of the difference and corresponding two-sided 95% CI for ONO-4538 group relative to control group, docetaxel group and paclitaxel group will be calculated using CMH methodology and adjusted by the same stratification factors.

(2) The proportions and the corresponding two-sided 95% CI will be estimated using the Clopper-Pearson method for each treatment group and control regimen.

2) Analytical items for DCR

(1) Data will be compared between the two treatment groups using the CMH test with the randomization factors as the stratification factors. The associated odds ratio and the corresponding two-sided 95% CI, and the estimate of the difference and corresponding two-sided 95% CI for ONO-4538 group relative to control group, docetaxel group and paclitaxel group will be calculated using CMH methodology and adjusted by the same stratification factors.

(2) The proportions and the corresponding two-sided 95% CI will be estimated using the Clopper-Pearson method for each treatment group.

3) Analytical item for PFS

(1) Data will be compared between the two treatment groups using the stratified log-rank test with the randomization factors as the stratification factors. The hazard ratio and the corresponding two-sided 95% CI for the ONO-4538 group relative to the control group and docetaxel group and paclitaxel group will be estimated using a stratified Cox proportional hazards model with the randomization factors as the stratification factors.

(2) The Kaplan-Meier curve will be plotted for each treatment group and control regimen. Using the Kaplan-Meier method, the median PFS and the corresponding two-sided 95% CI will be estimated for each treatment group and control regimen. CI for median PFS will be calculated using the Brookmeyer and Crowley method based on a log-log transformed CI for the survivor function.

(3) For each treatment group and control regimen, the progression free survival rates at months 3, 6, 9, 12, 15 and 18 will be derived from the Kaplan-Meier estimates and the corresponding two-sided 95% confidence intervals will be derived based on Greenwood's formula for variance derivation and on log-log transformation applied on the survivor function.

(4) The status of subjects who are censored in the Kaplan-Meier analysis will be tabulated for each treatment group and control regimen by using following categories: (a) on-study, (b) off-study, (c) received subsequence anti-cancer therapy, (d) never treated.

4) Analytical item for DOR

(1) The Kaplan-Meier curve will be plotted for each treatment group and control regimen. Using the Kaplan-Meier method, the median DOR and the corresponding two-sided 95% CI will be estimated for each treatment group and control regimen. CI for median DOR will be calculated using the Brookmeyer and Crowley method based on a log-log transformed CI for the survivor function. Additionally, summary statistics (minimum and maximum) will also be estimated for each treatment group and control regimen. This analysis will be performed on ITT/RES subjects whose best overall response is confirmed CR or confirmed PR

(2) The number of events and censors will be tabulated by treatment group and control regimen. The status of subjects who are censored in the Kaplan-Meier analysis will also be tabulated by treatment group and control regimen using following categories: (1) on-study, (2) off-study, (3) received subsequence anti-cancer therapy.

5) Analytical item for TTR

(1) Summary statistics will be calculated for ITT/RES subjects whose best overall response is confirmed CR or confirmed PR by treatment group and control regimen.

(2) The Kaplan-Meier curve will be plotted for each treatment group and control regimen on RES subjects.

6) Analytical item for BOR

(1) The percentage of confirmed CR, confirmed PR, SD, PD and NE will be calculated for each treatment group and control regimen. For the percentages of CR, PR and SD, the corresponding 95% CI will be estimated using the Clopper-Pearson method for each treatment group.

7) Analytical item for maximum percent change from baseline in the sum of diameters of target lesions

(1) Waterfall plot will be displayed by treatment group and control regimen.

10.4 Subgroup analyses for the efficacy endpoints

Subgroup analyses will be performed on following analytical items by the subgroups presented below. With regard to the analytical item (1) and (3), the ITT will be the analysis set. With regard to the analytical item (2), (4) and (5), the RES also will be the analysis set as well as ITT.

1) Analytical Items

(1) OS

(2) ORR

(3) PFS (Primary definition)

(4) DOR

(5) BOR

2) Handling of data

(1) Stratification factor and Method of classification

The stratification factors will be classified as the following table. Unless otherwise noted, the following two factors will only be extracted from IWRS; location and the number of organs with metastases. The others except for PD-L1 expression will be extracted from eCRF. PD-L1 expression will be extracted from both IWRS and test result.

Factor	Stratum
PD-L1 Expression (IWRS)	(1) >=1% (2) <1% or Indeterminate
PD-L1 Expression (test results)	(1) >=1%, <1% (2) >=5%, <5% (3) >=10%, <10% (4) Not Quantifiable
Location (IWRS)	Japan, Rest of world
Age[Note1)]	(1) <65, 65 - <75, >=75 (2) <65, >=65
Sex	Male, Female
Race	American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White,

	Other
ECOG Performance Status score at baseline	0, 1
Recurrent	No, Yes
Lesion sites (TNM classification)	Cervical Esophagus, Thoracic Esophagus (Upper Thorax, Middle Thorax, Lower Thorax), Unknown
Histological classification	Squamous Cell Carcinoma, Adenosquamous Cell Carcinoma, Other, Unknown
Number of organs with metastases (IWRS)	=<1, >=2
Lymph Node metastasis	No, Yes
Liver metastasis	No, Yes
Lung metastasis	No, Yes
Bone metastasis	No, Yes
Target lesion	No, Yes
Past treatments for cancer (surgery)	No, Yes
Past treatments for cancer (radiotherapy)	No, Yes
History of smoking	Never, Former, Current

Note1): Additional categories 75-<85 and >=85 will be reported if there are subjects >=85 years old

3) Analytical Methods

(1) For OS, PFS and DOR, the hazard ratio and the corresponding 95% CIs for the ONO-4538 group relative to the treatment group will be calculated using the unstratified Cox proportional-hazards model in each subgroup.

(2) For OS, PFS and DOR, the median for each endpoint and the corresponding two-sided 95% CI will be estimated for each treatment group and subgroup. CIs for each median will be calculated using the Brookmeyer and Crowley method based on a log-log transformed CI for the survivor function.

(3) Forest plots of OS, PFS and DOR will be created using (1) and (2) to present the hazard ratio and the corresponding 95% CIs and the median and the corresponding 95% CIs.

(4) For OS and PFS, the Kaplan-Meier curve will be plotted for each treatment group by PD-L1 Expression (test results).

(5) For ORR, the associated odds ratio and the corresponding two-sided 95% CI for ONO-4538 group relative to control group will be calculated using the logistic regression model with the treatment group as the single covariate in each subgroup. Forest plots will also be created to present the odds ratio and the corresponding 95% confidence interval.

(6) For ORR, the proportions and the corresponding two-sided 95% CI will be estimated using the Clopper-Pearson method for each treatment group and control regimen by PD-L1 Expression (test results).

(7) For BOR, the percentage of confirmed CR, confirmed PR, SD, PD and NE will be calculated for each treatment group by PD-L1 Expression (test results). For the percentages of CR, PR and SD, the corresponding 95% CI will be estimated using the Clopper-Pearson method for each treatment group.

10.5 Interaction Analysis

For the primary efficacy endpoint (i.e. OS), the interaction between treatment group and the demographic factors presented in 10.4 Subgroup analyses for the efficacy endpoints will be assessed using a Cox proportional-hazards model. For the analysis, each demographic factor,

treatment group and interaction between each demographic variable and treatment group will be used for the factors.

10.6 Adjusted Analysis

The adjusted analysis for primary endpoint (i.e. OS) will be performed using a demographic factors presented in 10.4 Subgroup analyses for the efficacy endpoints except for location, the number of organs with metastases and PD-L1 expression. The hazard ratio and the corresponding two-sided 95% CI of the ONO-4538 group relative to the control group will be calculated using the stratified Cox proportional hazard model with the stratification factors (IWRS source) presented in 9.4 Analytical Methods as the stratification factors, treatment groups and each demographic factor as the covariance factors.

10.7 Relationship Between Endpoints

Not planned in this study.

10.8 Other Analysis

1) Swimmer plot will be displayed by investigational product and BOR. Other exploratory analyses may be performed additionally as necessary.
2) Summary statistics of observation period will be calculated by treatment group and control regimen for the following analysis populations.
   - ITT
   - ITT and subjects continuing the study (ongoing treatment period or post observation period)

   Observation period for subject who died and who alive or lost to follow-up will be calculated by the following equations.

   Observation period for subjects who died (Months) = (“Date of death due to any cause” − “Date of randomization” + 1) / 30.4375

   Observation period for subjects who alive or lost to follow-up (Months) = (“Date of last survive confirmed” − “Date of randomization” + 1) / 30.4375

11 ANALYSIS OF SAFETY

11.1 Analysis Sets

The SAF will be the analysis set.

11.2 Analytical Items and Data Handlings

1) Analytical Items

(1) Adverse events, drug-related adverse events and deaths

(2) Select adverse events and drug-related select adverse events

(3) Immune-mediated adverse events (IMAEs)

(4) Other adverse events of special interests (AEOSIs)

(5) Laboratory tests

a) Hematology Tests

Hemoglobin, White blood cell count, Neutrophils count, Lymphocytes count, Platelet count

b) Biochemistry Tests

ALP, AST (GOT), ALT (GPT), Total bilirubin, Creatinine, Na, K, Ca

(6) Hormone Tests

Thyroid-stimulating hormone (TSH), Free triiodothyronine (free T3), Free thyroxine (free T4)

(7) 12-lead ECG

Heart Rate (HR), PR Interval, RR Interval, QRS Width, QT Interval, QTcF Interval

2) Handling of data

Unless otherwise noted, the items occurring between the start date of the first administration of the investigational product and 28 days after the last dose or the start date of subsequence anti-cancer therapy after the last dose whichever comes first will be tabulated by treatment group.

(1) Adverse events, drug-related adverse events and deaths

An adverse event (AE) is any unfavorable or unintended sign (including an abnormal laboratory finding), symptom, or disease temporally associated with the use of the investigational product, regardless of whether the event is causally related to the investigational product. AEs occurring between the start date of the first administration of the investigational product and 28 days after the last dose or the start date of subsequence anti-cancer therapy after the last dose whichever comes first will be tabulated. In addition, the entire AEs occurring in the treatment period as well as the post-treatment observation period will be reported by listing.

Drug-related AEs will be defined as any AEs for which a causal relationship to the investigational product is “Related” or missing. If the causal relationship to the investigational product is missing, the AE will be considered as “Related”.

AEs will be coded by using the Medical Dictionary for Regulatory Activities (MedDRA), and classified by System Organ Class (SOC) and Preferred Term (PT). In the summary table, a subject will be counted as the only one subject in case of the multiple occurrence of the same AE within the same subject. AEs will be graded by using the NCI Common Terminology Criteria for Adverse Event (CTCAE), version 4.0, when preparing the tabulation by Grade. The worst CTCAE grade for each subject will be summarized in case of the multiple occurrence of the same AE within the same subject, when preparing the tabulation by CTCAE grade.

When preparing the tabulation by period, the earliest AE occurrence time in each period will be counted as one time in case of the multiple occurrence of the same AE within the same subject. Of note, the analysis set will include the only subjects who will be monitored during each period.

Deaths will be counted regarding the deceased subjects from the start date of the first administration of the investigational product through the earlier date on which either 28 days after the last dose or the start date of subsequence anti-cancer therapy after the last dose.

(2) Multiple events of adverse events

Analyses that take into account the multiple occurrences of a given adverse event will be conducted. This data will be presented as the rate per 100 person-years of exposure. The analyses will take into account all on-treatment events (allowing more than 1 event per subject) and the total exposure time. The person-year exposure will be computed as the sum over the subjects' exposure expressed in years where the exposure time is defined as

"Min (28 days after last dose of study treatment or Start date of subsequence anti-cancer therapy after the last dose of study treatment or Date of death or Last known date alive or cut off date – date of first dose of study treatment) / 365.25 "

(3) Select AEs and drug-related select AEs

Select AEs will be defined as any AEs which are described in Appendix 1. Select AEs occurring between the start date of the first administration of the investigational product and 28 days after the last dose or the start date of subsequence anti-cancer therapy after the last dose whichever comes first will be tabulated.

Time to onset of select AEs (any grade) in each category (Weeks) = ("The onset date of the earliest select AE in each category" – "The day of first dose of study treatment" + 1) / 7

The time to onset of select AEs of a subject who does not have select AEs in each category will be censored at the maximum time of the period of the time to onset of select AEs in each category and each treatment group.

Time to onset of select AEs (CTCAE grade 3-5) for a specific category is defined similarly but restricted to CTCAE grade 3-5 select AEs.

Time to onset of drug-related select AEs (CTCAE grade 3-5 or any grade) for a specific category is defined similarly but restricted to drug-related select AEs.

Time to onset for a specific subcategory is defined similarly but restricted to event of this subcategory.

The time to resolution of select AEs in each category in each subject is the maximum time to resolution of select AEs in each subject calculated by the following equation. The time to

resolution of select AEs in each category will be defined only for a subject who has at least one select AE included in each category.

Time to resolution of select AEs (any grade) in each category (Weeks) = ("The last complete resolution date of a select AEs appearing consecutively in each category" - " The first onset date of a select AEs appearing consecutively in each category" + 1) / 7

If a select AE is considered as unresolved, the resolution date will be censored on the last assessment date of outcome.

The time to resolution of select AEs (CTCAE grade 3-5) for a specific category is defined similarly with an onset date corresponding to a CTCAE grade 3-5 select AE.

Time to resolution of drug-related select AEs (CTCAE grade 3-5 or any grade) for a specific category is defined similarly but restricted to drug-related select AEs.

(4) Immune-mediated AEs

Immune-mediated AEs (IMAEs) will be defined as any AEs which are described in Appendix 2. IMAEs occurring between the start date of the first administration of the investigational product and 100 days after the last dose will be tabulated. All IMAEs analyses will be performed using the 100 days window. Analyses for IMAEs will be limited to subjects who included in SAF and received immunosuppressive medication for treatment of the event, with the exception of endocrine events, which will be included in the analysis regardless of treatment since endocrine events are often managed without immunosuppression.

Time to onset of IMAEs (any grade) in each category (Weeks) = ("The onset date of the earliest IMAE in each category" – "The day of first dose of study treatment" + 1) / 7

The time to onset of IMAEs of a subject who does not have IMAEs in each category will be censored at the maximum time of the period of the time to onset of IMAEs in each category and each treatment group.

Time to onset of IMAEs (CTCAE grade 3-5) for a specific category is defined similarly but restricted to CTCAE grade 3-5 IMAEs.

The time to resolution of IMAEs (any grade) in each category in each subject is the maximum time to resolution of IMAEs in each subject calculated by the following equation. The time to

resolution of IMAEs in each category will be defined only for a subject who has at least one IMAE included in each category.

Time to resolution of IMAEs (any grade) in each category (Weeks) = ("The last complete resolution date of IMAEs appearing consecutively in each category" – " The first onset date of IMAEs appearing consecutively in each category" + 1) / 7

Events which worsened into CTCAE grade 5 events (death) are considered unresolved. If IMAE is considered as unresolved, the resolution date will be censored to the last assessment date of outcome.

The time to resolution of IMAEs (CTCAE grade 3-5) for a specific category is defined similarly with an onset date corresponding to a CTCAE grade 3-5 IMAEs.

(5) Other adverse events of special interests (AEOSIs)

AEOSIs will be defined as any AEs which are described in Appendix 3. AEOSIs occurring between the start date of the first administration of the investigational product and 28 days after the last dose or the start date of subsequence anti-cancer therapy after the last dose whichever comes first will be tabulated.

Time to onset of AEOSIs (any grade) in each category (Weeks) = ("The onset date of the earliest AEOSIs in each category" – "The day of first dose of study treatment" + 1) / 7

The time to onset of AEOSIs of a subject who does not have AEOSIs in each category will be censored at the maximum time of the period of the time to onset of AEOSIs in each category and each treatment group.

Time to onset of AEOSIs (CTCAE grade 3-5) for a specific category is defined similarly but restricted to CTCAE grade 3-5 AEOSIs.

Time to onset of drug-related AEOSIs (CTCAE grade 3-5 or any grade) for a specific category is defined similarly but restricted to drug-related AEOSIs.

The time to resolution of AEOSIs (any grade) in each category in each subject is the maximum time to resolution of AEOSIs in each subject calculated by the following equation. The time to

resolution of AEOSIs in each category will be defined only for a subject who has at least one AEOSI included in each category.

Time to resolution of AEOSIs (any grade) in each category (Weeks) = ("The last complete resolution date of AEOSIs appearing consecutively in each category" - " The first onset date of AEOSIs appearing consecutively in each category" + 1) / 7

Events which worsened into CTCAE grade 5 events (death) are considered unresolved. If AEOSI is considered as unresolved, the resolution date will be censored to the last assessment date of outcome. This measure is defined only for subjects who experienced at least one IMAE in the specific category.

The time to resolution of AEOSIs (CTCAE grade 3-5) for a specific category is defined similarly with an onset date corresponding to a CTCAE grade 3-5 AEOSIs.

Time to resolution of drug-related AEOSIs (CTCAE grade 3-5 or any grade) for a specific category is defined similarly but restricted to drug-related AEOSIs.

(6) Laboratory tests

Laboratory tests between the start date of the first administration of the investigational product and 28 days after the last dose or the start date of subsequence anti-cancer therapy after the last dose whichever comes first will be tabulated.

Baseline data will be defined as the data in the screening period. If more than one data are collected in screening period then latest one is defined as the baseline data.

If the measurement is lower than XX, upper than or equal to YY, and so on, the numerical value (i.e. XX, YY) will be used.

Laboratory tests will be graded by using CTCAE version 4.0, when preparing the tabulation by Grade.

Clinical laboratory data will be first analyzed using International System of Units (SI). Analyses will be repeated using US conventional units.

(7) Hormone tests

Hormone tests between the start date of the first administration of the investigational product and 28 days after the last dose or the start date of subsequence anti-cancer therapy after the last dose whichever comes first will be tabulated.

Baseline data will be defined as the data in the screening period. If more than one data are collected in screening period then latest one is defined as the baseline data.

If the measurement is lower than XX, upper than or equal to YY, and so on, the numerical value (i.e. XX, YY) will be used.

Hormone tests data will be first analyzed using International System of Units (SI). Analyses will be repeated using US conventional units.

(8) 12-lead ECG

12-lead ECG between the start date of the first administration of the investigational product and 28 days after the last dose or the start date of subsequence anti-cancer therapy after the last dose whichever comes first will be tabulated.

Baseline data will be defined as the data at Day 1 pre-dose in Cycle 1. Of note, in case of the missing data at Day 1 pre-dose in Cycle 1, the alternative baseline data will be defined as the data in the screening period.

11.3 Analytical methods

1) Adverse events, drug-related adverse events and deaths

(1) Numbers of subjects with AEs and drug-related AEs, serious AEs (SAEs) and drug-related SAEs, AEs and drug-related AEs leading to discontinuation of study treatment, AEs and drug-related AEs leading to dose delay of study treatment, AEs and drug-related AEs leading to dose reduction of study treatment will be summarized by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group and control regimen.

(2) Number of subjects who died within 28 days / 100 days after the last dose of the study treatment as well as the frequency by reason of death will be summarized by treatment group and control regimen.

(3) Number of subjects who died within 28 days after the last dose of study treatment or the start date of subsequence anti-cancer therapy after the last dose of study treatment whichever comes first as well as the frequency by reason of death will be summarized by treatment group and control regimen.

(4) Number of subjects who died as well as the frequency by reason of death will be summarized by treatment group and control regimen.

(5) AEs and drug-related AEs with the incidence rate greater than or equal to 5% in any treatment group and control regimen will be summarized by PT as well as by period.

(6) AEs and drug-related AEs by the worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group and control regimen will be summarized by SOC and PT.

(7) AEs and drug-related AEs by the worst CTCAE grade (grade 1, 2, 3, 4, and 5) with the incidence rate greater than or equal to 5% in any treatment group and control regimen will be summarized by SOC and PT.

(8) AEs and drug-related AEs by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group and control regimen will be summarized by SOC and PT.

(9) AEs and drug-related AEs by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group and control regimen with the incidence rate greater than or equal to 5% in any treatment group and control regimen will be summarized by SOC and PT.

(10) AEs and drug-related AEs leading to discontinuation of study treatment by the worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group and control regimen will be summarized by SOC and PT.

(11) AEs and drug-related AEs leading to discontinuation of study treatment by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group and control regimen will be summarized by SOC and PT.

(12) AEs and drug-related AEs leading to dose delay of study treatment by the worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group and control regimen will be summarized by SOC and PT.

(13) AEs and drug-related AEs leading to dose delay of study treatment by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group and control regimen will be summarized by SOC and PT.

(14) AEs and drug-related AEs leading to dose reduction of study treatment by the worst CTCAE grade (grade 1, 2, 3, 4, and 5) in control group and control regimen will be summarized by SOC and PT.

(15) AEs and drug-related AEs leading to dose reduction of study treatment by the worst CTCAE grade (any grade, grade 3-4, grade 5) in control group and control regimen will be summarized by SOC and PT.

(16) AEs and drug-related AEs with any CTCAE grade 3 or 4 in each treatment group and control regimen will be summarized by SOC and PT.

(17) SAEs and drug-related SAEs by the worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group and control regimen will be summarized by SOC and PT. The same analysis will be performed for SAEs and drug-related SAEs occurring between the start date of the first administration of the investigational product and 100 days after the last dose.

(18) SAEs and drug-related SAEs by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group and control regimen will be summarized by SOC and PT. The same analysis will be performed for SAEs and drug-related SAEs occurring between the start date of the first administration of the investigational product and 100 days after the last dose.

(19) AEs and drug-related AEs leading to death in each treatment group and control regimen will be summarized by SOC and PT.

2) Multiple events of adverse events

(1) The total number and rate (exposure adjusted) of occurrences for all AEs and drug-related AEs in each treatment group and control regimen will be summarized by SOC and PT.

(2) The number of subjects experiencing a select AEs once or multiple times in each treatment group and control regimen will be summarized by specific category, subcategory and PT.

3) Select AEs and drug-related select AEs

(1) Numbers of subjects with select AEs and drug-related select AEs, serious select AEs and drug-related serious select AEs, select AEs and drug-related select AEs leading to discontinuation of study treatment, select AEs and drug-related select AEs leading to dose delay of study treatment will be summarized by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group and specific category and subcategory.

(2) Select AEs and drug-related select AEs by the worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group will be summarized by specific category, subcategory and PT.

(3) Select AEs and drug-related AEs by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group will be summarized by specific category, subcategory and PT.

(4) Select AEs and drug-related select AEs leading to discontinuation of study treatment by the worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group will be summarized by specific category, subcategory and PT.

(5) Select AEs and drug-related select AEs leading to discontinuation of study treatment by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group will be summarized by specific category, subcategory and PT.

(6) Select AEs and drug-related select AEs leading to dose delay of study treatment by the worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group will be summarized by specific category, subcategory and PT.

(7) Select AEs and drug-related AEs leading to dose delay of study treatment by worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group will be summarized by specific category, subcategory and PT.

(8) Serious select AEs and drug related serious select AEs by the worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group will be summarized by specific category, subcategory and PT. The same analysis will be performed for serious select AEs and drug-related serious

select AEs occurring between the start date of the first administration of the investigational product and 100 days after the last dose.

(9) Serious select AEs and drug related serious select AEs by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group will be summarized by specific category, subcategory and PT. The same analysis will be performed for serious select AEs and drug-related serious select AEs occurring between the start date of the first administration of the investigational product and 100 days after the last dose.

(10) Select AEs and drug-related select AEs leading to death in each treatment group will be summarized by specific category, subcategory and PT.

4) IMAEs

(1) Numbers of subjects with IMAEs, serious IMAEs, IMAEs leading to discontinuation of study treatment, IMAEs leading to dose delay of study treatment will be summarized by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group and specific category.

(2) IMAEs by the worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group will be summarized by specific category and PT.

(3) IMAEs by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group will be summarized by specific category and PT.

(4) IMAEs leading to discontinuation of study treatment by worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group will be summarized by specific category and PT.

(5) IMAEs leading to discontinuation of study treatment by worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group will be summarized by specific category and PT.

(6) IMAEs leading to dose delay of study treatment by worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group will be summarized by specific category and PT.

(7) IMAEs leading to dose delay of study treatment by worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group will be summarized by specific category and PT.

(8) Serious IMAEs by worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group will be summarized by specific category and PT.

(9) Serious IMAEs by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group will be summarized by specific category and PT.

(10) IMAEs leading to death in each treatment group will be summarized by specific category and PT.

(11) AEs which are qualified for IMAE's PT list by worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group will be summarized by specific category and PT.

5) Total medication treatment duration (excluding overlaps) of immune-modulating concomitant medication

(1) Percentage of subjects who received immune modulating concomitant medication for management of drug-related select AEs (any grade, grade 3-5) among subjects who experienced at least one drug-related select AEs in the category/subcategory will be calculated by treatment group.

(2) Percentage of subjects who received immune modulating concomitant medication for management of IMAEs (any grade, grade 3-5) among subjects who experienced at least one drug-related select AEs in the category/subcategory will be calculated by treatment group.

(3) For each category/subcategory of drug-related select AEs (any grade, grade 3-5) and IMAEs (any grade, grade 3-5), summary statistics (median, minimum and maximum) of total medication treatment duration (excluding overlaps) of immune-modulating concomitant medication will be calculated by treatment group.

6) AEOSIs

(1) AEOSIs and drug-related AEOSIs by the worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group will be summarized by specific category and PT.

(2) AEOSIs and drug-related AEOSIs by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group will be summarized by specific category and PT.

(3) Serious AEOSIs and drug-related serious AEOSIs of special interests by the worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group will be summarized by specific category and PT.

(4) AEOSIs and drug-related AEOSIs where immune modulating medication was initiated by the worst CTCAE grade (grade 1, 2, 3, 4, and 5) in each treatment group will be summarized by specific category and PT.

7) Time to onset of select AEs, IMAEs and AEOSIs

Summary statistics (median, minimum and maximum) of time to onset of the following specific events will be calculated by treatment group.

(1) Time to onset of any grade IMAEs by treatment group

(2) Time to onset of CTCAE grade 3 to 5 IMAEs by treatment group

(3) Time to onset of any grade select AEs and drug-related select AEs by treatment group

(4) Time to onset of CTCAE grade 3 to 5 select AEs and drug-related select AEs by treatment group

(5) Time to onset of any grade AEOSIs and drug-related AEOSIs by treatment group

(6) Time to onset of CTCAE grade 3 to 5 AEOSIs and drug-related AEOSIs by treatment group

(7) Time to onset of any grade AEOSIs and drug-related AEOSIs where immune-modulating medication was initiated

(8) Time to onset of CTCAE grade 3 to 5 AEOSIs and drug-related AEOSIs where immune modulating concomitant medication was initiated

8) Time to resolution of select AEs, IMAEs and AEOSIs

Using the Kaplan-Meier method, the median and the corresponding two-sided 95% CI for time to resolution and range of the following specific events will be estimated for each category/subcategory of select AEs, IMAEs (only for category) and AEOSIs (only for category). The CI will be calculated using the Brookmeyer and Crowley method based on a log-log transformed CI for the survivor function.

(1) Time to resolution of any grade IMAEs by treatment group

(2) Time to resolution of CTCAE grade 3 to 5 IMAEs by treatment group

(3) Time to resolution of any grade select AEs and drug-related select AEs by treatment group

(4) Time to resolution of CTCAE grade 3 to 5 select AEs and drug-related select AEs by treatment group

(5) Time to resolution of any grade select AEs and drug-related select AEs where immune modulating medication was initiated, by treatment group

(6) Time to resolution of CTCAE grade 3 to 5 select AEs and drug-related select AEs where immune modulating concomitant medication was initiated, by treatment group

(7) Time to resolution of any grade AEOSIs and drug-related AEOSIs by treatment group

(8) Time to resolution of CTCAE grade 3 to 5 AEOSIs and drug-related AEOSIs by treatment group

(9) Time to resolution of any grade AEOSIs and drug-related AEOSIs where immune modulating concomitant medication was initiated

(10) Time to resolution of CTCAE grade 3 to 5 AEOSIs and drug-related AEOSIs where immune modulating concomitant medication was initiated

9) Laboratory tests

(1) The worst CTCAE Grade compared to the baseline grade will be evaluated by treatment group and control regimen in the shift table during the treatment period as to the following laboratory tests: Hematology (Hemoglobin, White blood cell count, Platelet count, Lymphocytes count, Neutrophils count), Liver (ALP, AST (GOT), ALT (GPT), Total bilirubin), Creatinine and Electrolyte (Na, K, Ca).

(2) The worst CTCAE Grade (Grade 1-4 and Grade 3-4) in laboratory test that worsened relative to baseline will be evaluated by treatment group as to the following laboratory tests: Hematology (Hemoglobin, White blood cell count, Platelet count, Lymphocytes count, Neutrophils count), Liver (ALP, AST (GOT), ALT (GPT), Total bilirubin), Creatinine and Electrolyte (Na, K, Ca).

(3) Focusing on the abnormal hepatic function, the frequency distributions of the appropriate subjects will be specified according to the following criteria by treatment group.

a) ALT or AST > 3xULN (Upper Limit of Normal), 5xULN, 10xULN, 20xULN

b) Total bilirubin > 2xULN

c) ALT or AST > 3xULN as well as total bilirubin collected 1 day before and after > 2xULN

d) ALT or AST > 3xULN as well as total bilirubin collected 30 days before and after > 2xULN

e) ALT or AST > 3xULN as well as total bilirubin collected 1 day before and after ≥ 2xULN, ALP < 2xULN

10) Hormone tests

(1) Focusing on the thyroid dysfunction, the frequency distributions of the appropriate subjects will be specified according to the following criteria by treatment group. Both free T3 and free T4 are targeted within a 2-week window after abnormal TSH test date.

a) TSH > ULN (Upper limit of Normal)

b) TSH > ULN as well as TSH ≤ ULN at the baseline

c) TSH > ULN as well as either free T3 or free T4 < LLN (Lower Limit of Normal) at least one assessment time

d) TSH > ULN as well as both free T3 and free T4 ≥ LLN at all assessment time

e) TSH > ULN as well as either free T3 or free T4 is missing value.

f) TSH < LLN

g) TSH < LLN as well as TSH ≥ LLN at the baseline

h) TSH < LLN as well as either free T3 or free T4 > ULN at least one assessment time

i) TSH < LLN as well as both free T3 and free T4 ≤ ULN at all assessment time

j) TSH < LLN as well as either free T3 or free T4 is missing value.

11) 12-lead ECG

(1)The maximal data of QRS Width, PR Interval, RR Interval, QT Interval, QTcF Interval, the change of QTcF Interval from baseline and HR will be summarized by treatment group using frequency distributions according to the following category.

ECG parameters (unit)	Classification
RR (ms) (actual value)	<=600, >600 - 1200, >1200
PR (ms) (actual value)	<=120, >120 - 200, >200
QRS (ms) (actual value)	<=60, >60 - 109, >109
HR (beats/min) (actual value)	<=50, >50 - 100, >100
QT (ms) (actual value)	<=450, >450 - 480, >480 - 500, >500
QTcF (ms) (actual value)	<=450, >450 - 480, >480 - 500, >500
QTcF (ms) (the change from baseline)	<=30, >30 - 60, >60

Focusing on the high incidence of adverse events, exploratory analyses may be performed additionally relating to the safety of ONO-4538 as necessary.

11.4 Subgroup analyses for the safety endpoints

AEs and drug-related AEs, serious AEs and drug-related serious AEs by the worst CTCAE grade (any grade, grade 3-4, grade 5) in each treatment group will be summarized by SOC and PT for the following factors.

Factor	Stratum
PD-L1 Expression (test results)	(1) >=1%, <1% (2) >=5%, <5% (3) >=10%, <10% (4) Not Quantifiable
Location (IWRS)	Japan, Rest of world

Age	<65, >=65
Sex	Male, Female
Race	Asian, Other
ECOG Performance Status score at baseline	0, 1
Recurrent	No, Yes
Past treatments for cancer (surgery)	No, Yes
Past treatments for cancer (radiotherapy)	No, Yes
History of smoking	Never, Former, Current

12 PD-L1 EXPRESSION ANALYSIS

12.1 Analysis Sets

The ITT will be the analysis set

12.2 Analytical Item and Data Handlings

1) Analytical Item

PD-L1 expression

2) Handlings of data

(1) Quantifiable

Subjects with an available tissue specimens and number of available cancerous cells are ≥ 100 and percentage of tumor PD-L1 expression is ≥ 0%.

(2) PD-L1 expression at baseline

If more than one tissue specimens are available, baseline PD-L1 expression will be determined from the most recently collected specimen (prior to first dose of study treatment or prior to randomization date for subjects randomized but never treated) with a quantifiable result.

(3)PD-L1 status at baseline

PD-L1 status at baseline is a categorical valuable by X% cutoff for quantifiable PD-L1 baseline expression:

- PD-L1 Positive at X% cutoff value: ≥ X% PD-L1 expression
- PD-L1 Negative at X% cutoff value: < X% PD-L1 expression and ≥ 0%
- PD-L1 Not Quantifiable: subjects with no quantifiable PD-L1 expression at baseline (includes subjects without tissue samples for PD-L1)

where X denotes the PD-L1 expression cutoff of 1%, 5%, or 10%.

12.3 Analytical methods

1) Frequency of PD-L1 status at baseline will be summarized by treatment group and control regimen.

2) Cumulative distribution plot of baseline PD-L1 expression versus population percentile will be created by treatment group and control regimen on subjects with quantifiable PD-L1 expression at baseline.

13 ANTI-DRUG ANTIBODY ANALYSIS

13.1 Analysis Sets

The ADA will be the analysis set.

13.2 Analytical Item and Data Handlings

1) Analytical Item

Anti-ONO-4538 antibody up-regulated expression

2) Handlings of data

(1) Handling of each sample

Each sample will be handled in the following manner;

a) Baseline-positive sample

Baseline-positive sample will be defined as the sample detected any anti-ONO-4538 antibody in the last sample prior to the first administration of the investigational product.

b) Baseline-negative sample

Baseline-negative sample will be defined as the sample not detected any anti-ONO-4538 antibody in the last sample prior to the first administration of the investigational product.

c) Anti-ONO-4538 antibody-positive sample

Anti-ONO-4538 antibody-positive sample will be defined as the sample collected after the administration of the investigational product, meeting any conditions as follows;

i) The sample detected no anti-ONO-4538 antibody prior to the first administration of the investigational product as well as detected any anti-ONO-4538 antibody after the first administration of the product.

ii) For the sample detected any anti-ONO-4538 antibody in the last sample prior to the first administration of the investigational product, the sample detected greater than or equal to quadruple expression of the anti-ONO-4538 antibody relative to the baseline-sample expression after the first administration of the investigational product.

d) Anti-ONO-4538 antibody-negative sample

Anti-ONO-4538 antibody-negative sample will be defined as the sample not positive relative to baseline after the first administration of the investigational product, excluding the missing data.

(2) Handling of each subject

Each subject will be handled in the following manner;

a) Baseline-positive subject, Baseline neutralizing positive subject

Baseline-positive subject will be defined as the subject who has the baseline-positive sample. Among the baseline-positive subjects, the subject who has greater than or equal to one baseline-positive sample detected the neutralizing antibody will be defined as the baseline neutralizing positive.

b) Anti-ONO-4538 antibody-positive subject

Anti-ONO-4538 antibody-positive subject will be defined as the subject who has at least one anti-ONO-4538 antibody-positive sample after the first administration of the investigational product. In addition, Anti-ONO-4538 antibody-positive subject will be classified according to the following definitions;

i) Persistent Positive (PP) :

The subject who has the anti-ONO-4538 antibody-positive samples on greater than or equal to two consecutive sampling time points after the first administration of the investigational product, provided that the time interval should leave at least 16 weeks (including 1 week as the defined allowable window).

ii) Not PP – Last Sample Positive :

The subject who is not PP with the anti-ONO-4538 antibody-positive sample at the last sampling time point.

iii) Other positive :

The subject who is not PP but some the anti-ONO-4538 antibody-positive samples after the first administration with last sample being negative.

iv) Neutralizing positive :

The subject who has greater than or equal to one anti-ONO-4538 antibody-positive sample detected the neutralizing antibody after the first administration of investigational product.

c) Anti-ONO-4538 antibody-negative subject

Anti-ONO-4538 antibody-negative subject will be defined as the subject who has no anti-ONO-4538 antibody-positive sample.

13.3 Analytical methods

The percentage of the baseline-positive subject, baseline neutralizing positive subject, anti-ONO-4538 antibody-positive subject and anti-ONO-4538 antibody-negative subject will be calculated for each treatment group. For the anti-ONO-4538 antibody-positive subject, the percentage of the persistent positive (PP), not PP - last sample positive and the other positive will also be calculated. The percentage of the baseline neutralizing positive as well as neutralizing positive will be calculated as necessary.

14 PATIENT REPORTED OUTCOMES ANALYSIS

14.1 Analysis Sets

The ITT will be the analysis set.

14.2 Analytical Items and Data Handling

1) Analytical Items

(1) Patient with questionnaire completion

(2) EuroQol 5 Dimension (EQ-5D)

(3) EQ-5D index scores (UK based scoring)

(4) Change from baseline of EQ-5D index scores (UK based scoring)

(5) EQ-5D index scores (Japan based scoring)

(6) Change from baseline of EQ-5D index scores (Japan based scoring)

(7) EuroQol visual analogue scale (EQ-VAS) scores

(8) Change from baseline of EQ-VAS scores

2) Handling of data

(1) EQ-5D index scores (UK and Japan based scoring)

EQ-5D questionnaire asks the subjects to describe their current health state on 5 dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) with each dimension having 3 response choices (1 = no problem, 2 = some problem, and 3 = extreme problem). For example, if mobility = 2, self-care = 2, usual activities = 1, pain/discomfort = 1, and anxiety/depression = 1, then “Health State” would be “22111”.

(At least one 3 (N3), M2 (Mobility = 2), M3 (Mobility = 3), SC2 (Self Care = 2), SC3 (Self Care = 3), UA2 (Usual Activities = 2), UA3 (Usual Activities = 3), PD2 (Pain/Discomfort = 2), PD3 (Pain/Discomfort = 3), AD2 (Anxiety/Depression = 2), AD3 (Anxiety/Depression = 3))

If Full health “11111”, the value set is equal to 1. In other cases, the below mathematical representation is used.

UK Value Set: 1 - 0.081 - 0.269N3 - 0.069M2 - 0.314M3 - 0.104SC2 - 0.214SC3 - 0.036UA2 - 0.094UA3 - 0.123PD2 - 0.386PD3 - 0.071AD2 - 0.236AD3

Japan Value Set: 1 - 0.152 - 0.075MO2 - 0.418MO3 - 0.054SC2 - 0.102SC3 - 0.044UA2 - 0.133UA3 - 0.080PD2 - 0.194PD3 - 0.063AD2 - 0.112AD3

(2) Time window for EQ-5D assessments

All questionnaires completed at baseline and on-study will be assigned to a time-point according to the windowing criteria in Table 14-2-1 and included in the analysis. In case a subject has two on-study assessments within the same window, the assessment closest to the time-point will be used. And, in the case of two assessments at a similar distance to the time-point, the latest one will be chosen. In the event where the subject has no assessment at all in a specific window, the observation will be treated as missing for that time-point.

Table 14-2-1: Time windows for EQ-5D Assessments

Nominal Time-Point	Time Window
Screening Phase (Baseline)	< Date of the first dose of study treatment
Treatment phase (until last dose date)	
Week 6 (Day 43)	Day 1 - <=Day 64
Week 12 (Day 85)	Day 65 - <=Day 106
Week 18 (Day 127)	Day 107 - <= Day 148
…Every 6 weeks thereafter while on treatment	Nominal Day (+21 days / -20 days, inclusive)
Follow-up Phase (from last dose date)	
Follow-up 1 (Week 12 (Day 85) from last dose date)	Last dose date to Last dose date + 127 days
Follow up 2 (Week 24 (Day 169) from last dose date)	Last dose date + 128 days to Last dose date + 211 days
…Every 12 weeks thereafter while on study	Nominal Day(Range of previous follow-up +84 days)

14.3 Analytical methods

1) Patient with questionnaire completion

The percentage of patients with questionnaire completion in each investigational product will be summarized at each time point. Percentage is defined as the proportion of questionnaires actually

received out of the expected number (i.e., the number of subjects still on treatment or in follow up at each time point).

2) EQ-5D

The percentages of patients with no health problems, moderate health problems and extreme health problems will be calculated for the each EQ-5D dimension at each time point by investigational product. Percentages will be calculated based on the number of patients assessed at each assessment time point for the each EQ-5D dimension.

3) EQ-5D index scores (UK based scoring) and EQ-5D index score (Japan based scoring)

(1) For EQ-5D index scores at each time point in each investigational product will be summarized using summary statistics. 95% CIs for mean scores of EQ-5D index scores at each time point in each investigational product will also be calculated.

(2) For change from baseline of EQ-5D index scores at each time point in each investigational product will be summarized using summary statistics. 95% CIs for mean scores of change from baseline of EQ-5D index scores at each time point in each investigational product will also be calculated.

4) EQ-VAS scores

(1) For EQ-VAS scores at each time point in each investigational product will be summarized using summary statistics. 95% CIs for mean scores of EQ-VAS scores at each time point in each investigational product will also be calculated.

(2) For change from baseline of EQ-VAS scores at each time point in each investigational product will be summarized using summary statistics. 95% CIs for mean scores of change from baseline of EQ-5D index scores at each time point in each investigational product will also be calculated.

15 INTERIM STATISTICAL ANALYSIS

Formal interim analysis is not planned in this study.

16 DATA REVIEWS

16.1 Objectives

After data collection from this study, using the data before fixed, a preliminary analysis will be performed in terms of the distribution of background factors etc., to determine the appropriateness of planned analytical methods.

16.2 Analysis Sets

The ITT will be the analysis set.

16.3 Analytical Items and Data Handling

1) Analytical Items

(1) Patient background factors

(2) Subsequence anti-cancer therapy (surgery, radiotherapy, pharmacotherapy)

(3) Efficacy endpoints

a) OS

b) PFS (Primary definition)

c) ORR

(4) Adverse events and drug-related adverse events

2) Handling of data

Each item will be analyzed according to the method defined above.

16.4 Analytical Methods

1) Frequency distributions and summary statistics will be calculated for patient background factors, and the classification in 8.3 Analytical Methods and Classification will be determined.
2) Details of subsequence anti-cancer therapy (surgery, radiotherapy, pharmacotherapy) will be summarized.
3) The number of events and censors will be tabulated. The status of subjects who are censored in the Kaplan-Meier analysis will also be tabulated using the following categories: (1) on-study, (2) off-study.
4) The median OS and the corresponding 95% confidence interval by using Kaplan-Meier method will be estimated. 95% confidence interval for the median OS will be calculated using the Brookmeyer and Crowley method based on a log-log transformed confidence interval for the survivor function.
5) To examine the necessity of sensitive analysis for OS, cross-tabulation between stratification factors (IWRS source) and stratification factors (eCRF source) will be provided.
6) The number of events and censors of PFS and reason for censored will be tabulated. Additionally, the median PFS and the corresponding 95% confidence interval by using Kaplan-Meier method will be estimated. 95% confidence interval for the median PFS will be calculated using the Brookmeyer and Crowley method based on a log-log transformed confidence interval for the survivor function.
7) To verify the validity of the analysis method, ORR and the corresponding exact (Clopper-Pearson method) 95% confidence intervals will be calculated by using Clopper-Pearson method. As necessary, the missing condition will be verified.
8) Numbers of patients with AEs and drug-related AEs will be summarized as a whole, or by System Organ Class (SOC), Preferred Term (PT) and Grade.

17 TECHNICAL MATTERS

17.1 Statistical Software

1) All statistical analyses will be carried out by using SAS (Statistical Analysis System software, SAS Institute, North Carolina, USA) version 9.3 or newer unless otherwise noted.
2) Listings, tables and figures will be generated by using Microsoft Word or Microsoft Excel.

   Of note, Software version will be updated according to the situation.

17.2 SAS Programming in Analyses

The statistical analyses will be conducted in accordance with this statistical analysis plan. The contents of the SAS procedure for statistical analyses will be presented in this section.

1) 95% confidence interval (Clopper-Pearson method)

   95% confidence interval (Clopper-Peason method) will be performed by using FREQ procedure in SAS programming. The following sample SAS programming will be applied for the calculation of 95% confidence interval (Clopper-Pearson method).

```
PROC FREQ DATA = [Analysis data set];
    TABLES [binary variable];
    EXACT BINOMIAL;
    WEIGHT [frequency] / ZEROS;
RUN;
```

2) The median OS and corresponding two-sided 95% CI from the Kaplan-Meier estimate, stratified log-rank test

   The median OS and the corresponding two-sided 95% CI derived from the Kaplan-Meier estimate and stratified log-rank test will be performed by using LIFETEST procedure in SAS programming.

PROC LIFETEST DATA = [Analysis data set] ALPHA = 0.05 METHOD=KM CONFTYPE=LOGLOG;

TIME [the survival time variable]* [the censoring variable](the value indicating censoring) ;

STRATA [the stratification factor] / GROUP = [treatment group] test = logrank ;

RUN;

3) The hazard ratio and corresponding two-sided 95% CI from the stratified Cox proportional-hazards model

The hazard ratio and corresponding two-sided 95% CI derived from the stratified Cox proportional-hazards model will be performed by using PHREG procedure in SAS programming.

PROC PHREG DATA = [Analysis data set] ALPHA = 0.05;

CLASS [treatment group];

STRATA [the stratification factor];

MODEL [the variables that define the survival time] * [the censoring variable](the value indicating censoring) = [treatment group] / RL TIES=EXACT;

RUN;

17.3 Data Format

1) p-value

The p-value will have fixed 4 decimal places, rounding off to 5 decimal places. Of note, p-value less than 0.0001 will be indicated as “p<0.0001”.

2) Convert days to months or years

1 month = 30.4375 days and 1 year = 365.25 days.

3) Calculation of data

In SAS programming, data will be evaluated complying with the double-precision operation in SAS system. Under the same condition, the calculation error will be removed from both the analysis data and the analysis result.

4) Definition of descriptive statistics and number of significant figures

Unless otherwise noted, the descriptive statistics (N, mean, median, SD, maximum, and minimum) will be calculated in case the number of subjects is more than two. Each number of significant figures for individual data will be indicated below. Of note, the number of significant figures for the specific descriptive statistics (mean, median, and SD) will be indicated as the individual number of significant figures +1(i.e., in case the individual number of significant figures is first decimal place, the specific descriptive statistics (mean, median and SD) will be indicated as second decimal place).

In addition, maximum and minimum will be indicated as the same decimal place as individual number of significant figures.

・Demographic variables and efficacy endpoints

Integral number (No decimal places)	Age(years) Number of organs with metastases (IWRS source and eCRF source)
First decimal place	Height(cm) Weight(kg) BMI(kg/m^2) Time from the date of diagnosis of the primary disease to randomization (months) Time from the date of surgery (including EMR or ESD) for the primary lesion to randomization (months)
Second decimal place	Diameters of target lesions (mm)
Third decimal place	BSA (m^2)

18 MAJOR CHANGES IN THE STATISTICAL ANALYSIS PLAN BASED ON THE STUDY-SPECIFIC PROTOCOL

None

19 MAJOR CHANGES IN THE STATISTICAL ANALYSIS PLAN BASED ON THE STATISTICAL ANALYSIS PLAN VERSION 1.0

Major analysis plan and data handlings wich were changed, added or deleted from statistical analysis plan version 1.0 are shown below.

19.1 Reliability of the study

<Additional items>

Relevant protocol deviations, reason for subjects enrolled but not randomized and reason for subjects but not treated were added.

<Deleted item>

Reason of withdrawals and dropouts from the study was deleted.

<Reason>

To assess reliability of the study more appropriately

19.2 Demographics and other baseline characteristics

<Additional items>

Analysis for past pharmacotherapy for cancer and details of immune modulating concomitant medication were added.

<Reason>

To assess imbalance of factors which may impact evaluation of efficacy and safety between the treatment groups

19.3 Primay efficacy analysis

19.3.1 Proportional hazards assumption

<Additional item>

Test for examining the proportional hazards assumption for OS was added.

<Reason>

To assess the proportional hazards assumption for OS

19.3.2 State of follow-up

<Additional items>

Analysis for current status of follow-up and duration of follow-up for OS were added.

<Reason>

To assess current status of follow-up for OS

19.3.3 Exploratory analytical methods

<Additional items>

Analyses using intented investigator’s choice therapy subset were added.

<Reason>

To evaluate an efficacy of ONO-4538 relative to control regimen from multilateral perspective

19.4 Secondary efficacy analysis

<Additional item>

Secondary definition of PFS was added.

<Reason>

To evaluate robustness of analysis result using PFS

19.5 Analysis of safety

19.5.1 Adverse event

<Additional items>

Multiple events of adverse events and other adverse events of special interests (AEOSIs) were added.

<Reason>

To evaluate safety of ONO-4538 relative to control group from multilateral perspective

19.6 PD-L1 expression analysis

<Additional item>

Section of PD-L1 expression analysis was added.

<Reason>

To evaluate expression level of PD-L1 in each treatment group

19.7 Patient reported outcomes analysis

<Additional items>

Change from baseline of EQ-5D and EQ-VAS scores were added.

<Additional items>

To evaluate quality of life of ONO-4538 relative to control regimen from multilateral perspective

20 MAJOR CHANGES IN THE STATISTICAL ANALYSIS PLAN BASED ON THE STATISTICAL ANALYSIS PLAN VERSION 2.0

Major analysis plan and data handlings which were changed, added or deleted from statistical analysis plan version 2.0 are shown below.

20.1 Significance level to be used

<Additional item>

Analysis population to be used for OS, ORR and PFS when performing the test using the Hierarchical hypothesis testing approach was added.

<Reason>

In this study, two analysis populations of ITT and RES (Response Evaluable Set) are defined as the analysis population for efficacy, and in order to clarify which analysis population will be tested for each endpoint.

20.2 Secondary efficacy analysis

<Additional item>

An analysis was added to calculate summary statistics of observation period of subjects for each treatment group.

<Reason>

To examine the observation period of subjects included in this study in more detail.

APPENDIX 1

Definition of Select Adverse Events

Category	Subcategory	PT
Endocrine Adverse Event	Adrenal Disorder	Adrenal insufficiency
		Adrenal suppression
		Adrenocortical insufficiency acute
		Blood corticotrophin decreased
		Blood corticotrophin increased
		Hypothalamic pituitary adrenal axis suppression
		Primary adrenal insufficiency
		Secondary adrenocortical insufficiency
	Diabetes	Diabetes mellitus
		Diabetic ketoacidosis
		Diabetic ketosis
		Fulminant type 1 diabetes mellitus
		Latent autoimmune diabetes in adults
		Type 1 diabetes mellitus
	Pituitary Disorder	Hypophysitis
		Hypopituitarism
		Lymphocytic hypophysitis
	Thyroid Disorder	Atrophic thyroiditis
		Autoimmune hypothyroidism
		Autoimmune thyroid disorder
		Autoimmune thyroiditis
		Basedow's disease
		Blood thyroid stimulating hormone decreased
		Blood thyroid stimulating hormone increased
		Hyperthyroidism
		Hypothyroidism
		Primary hyperthyroidism
		Primary hypothyroidism
		Silent thyroiditis
		Thyroid function test abnormal
		Thyroid hormones decreased
		Thyroid hormones increased
		Thyroiditis
		Thyroiditis acute
		Thyroxine decreased
		Thyroxine free decreased
		Thyroxine free increased

Category	Subcategory	PT
		Thyroxine increased
		Tri-iodothyronine uptake increased
Gastrointestinal Adverse Event		Autoimmune colitis
		Autoimmune enteropathy
		Colitis
		Colitis ulcerative
		Diarrhoea
		Duodenal perforation
		Enteritis
		Enterocolitis
		Enterocolitis haemorrhagic
		Frequent bowel movements
		Gastrointestinal perforation
		Lower gastrointestinal perforation
		Upper gastrointestinal perforation
Hepatic Adverse Event		Acute hepatic failure
		Acute on chronic liver failure
		Alanine aminotransferase increased
		Aspartate aminotransferase increased
		Autoimmune hepatitis
		Bilirubin conjugated increased
		Blood alkaline phosphatase increased
		Blood bilirubin increased
		Drug-induced liver injury
		Gamma-glutamyltransferase increased
		Hepatic enzyme increased
		Hepatic failure
		Hepatitis
		Hepatitis acute
		Hepatotoxicity
		Hyperbilirubinaemia
		Immune-mediated hepatitis
		Liver disorder
		Liver function test abnormal
		Liver function test increased
		Liver injury
		Transaminases increased

Category	Subcategory	PT
Hypersensitivity/Infusion Reaction		Anaphylactic reaction
		Anaphylactic shock
		Bronchospasm
		Hypersensitivity
		Infusion related reaction
Pulmonary Adverse Event		Acute respiratory distress syndrome
		Acute respiratory failure
		Idiopathic interstitial pneumonia
		Interstitial lung disease
		Lung infiltration
		Pneumonitis
Renal Adverse Event		Acute kidney injury
		Autoimmune nephritis
		Blood creatinine increased
		Blood urea increased
		Creatinine renal clearance decreased
		Hypercreatininaemia
		Nephritis
		Nephritis allergic
		Paraneoplastic glomerulonephritis
		Renal failure
		Renal tubular necrosis
		Tubulointerstitial nephritis
		Urine output decreased

Category	Subcategory	PT
Skin Adverse Event		Autoimmune dermatitis
		Blister
		Dermatitis
		Dermatitis allergic
		Dermatitis exfoliative
		Drug eruption
		Eczema
		Erythema
		Erythema multiforme
		Exfoliative rash
		Fixed eruption
		Nodular rash
		Palmar-plantar erythrodysaesthesia syndrome
		Pemphigoid
		Pemphigus
		Photosensitivity reaction
		Pruritus
		Pruritus allergic
		Pruritus generalised
		Psoriasis
		Rash
		Rash erythematous
		Rash generalised
		Rash macular
		Rash maculo-papular
		Rash morbilliform
		Rash papular
		Rash pruritic
		Rash vesicular
		Skin exfoliation
		Skin hypopigmentation
		Skin irritation
		Stevens-Johnson syndrome
		Toxic epidermal necrolysis
		Toxic skin eruption
		Urticaria
		Vitiligo

MedDRA v21.1

APPENDIX 2

Definition of Immune-Mediated Adverse Events

Category	Subcategory	PT
Endocrine	Adrenal insufficiency	Adrenal insufficiency
	Diabetes mellitus	Diabetes mellitus
		Diabetic ketoacidosis
	Hyperthyroidism	Hyperthyroidism
	Hypophysitis	Hypophysitis
	Hypothyroidism/Thyroiditis	Autoimmune thyroiditis
		Hypothyroidism
		Thyroiditis
		Thyroiditis acute
Diarrhea/Colitis		Colitis
		Diarrhoea
		Enterocolitis
Hepatitis		Alanine aminotransferase increased
		Aspartate aminotransferase increased
		Autoimmune hepatitis
		Blood alkaline phosphatase increased
		Blood bilirubin increased
		Hepatitis
		Hepatitis acute
		Hepatotoxicity
		Hyperbilirubinaemia
Nephritis and renal dysfunction		Nephritis
		Nephritis allergic
		Tubulointerstitial nephritis
		Acute kidney injury
		Renal failure
		Blood creatinine increased
Pneumonitis		Pneumonitis
		Interstitial lung disease
Rash		Rash
		Rash maculo-papular

APPENDIX 3

Definition of Other Adverse Events of Special Interests

Category	PT
Autoimmune Neuropathy	Autoimmune neuropathy
Demyelination Event	Anti-myelin-associated glycoprotein associated polyneuropathy
	Autoimmune demyelinating disease
	Demyelination
Encephalitis Event	Acute encephalitis with refractory, repetitive partial seizures
	Bickerstaff's encephalitis
	Encephalitis
	Encephalitis allergic
	Encephalitis autoimmune
	Encephalitis brain stem
	Encephalitis haemorrhagic
	Encephalitis lethargica
	Encephalitis toxic
	Limbic encephalitis
	Lupus encephalitis
	Noninfective encephalitis
	Panencephalitis
	Rasmussen encephalitis
	Subacute sclerosing panencephalitis
Graft Versus Host Disease	Acute graft versus host disease
	Acute graft versus host disease in intestine
	Acute graft versus host disease in liver
	Acute graft versus host disease in skin
	Chronic graft versus host disease
	Chronic graft versus host disease in intestine
	Chronic graft versus host disease in liver
	Chronic graft versus host disease in skin
	Graft versus host disease
	Graft versus host disease in eye
	Graft versus host disease in gastrointestinal tract
	Graft versus host disease in liver
	Graft versus host disease in lung
	Graft versus host disease in skin
Guillain-Barre Syndrome	Guillain-Barre syndrome
	Miller Fisher syndrome
Myasthenic Syndrome	Myasthenia gravis
	Myasthenia gravis crisis

Category	PT
	Myasthenic syndrome
	Ocular myasthenia
Myocarditis Event	Autoimmune myocarditis
	Eosinophilic myocarditis
	Hypersensitivity myocarditis
	Myocarditis
Myositis Event	Dermatomyositis
	Inclusion body myositis
	Myositis
	Necrotising myositis
	Paraneoplastic dermatomyositis
	Polymyositis
Pancreatitis Event	Autoimmune pancreatitis
	Haemorrhagic necrotic pancreatitis
	Pancreatitis
	Pancreatitis acute
	Pancreatitis necrotising
Rhabdomyolysis Event	Rhabdomyolysis
Uveitis Event	Autoimmune uveitis
	Chorioretinitis
	Cyclitis
	Iridocyclitis
	Iritis
	Keratouveitis
	Uveitis

MedDRA v21.1

SAP Amendments

There were 2 SAP amendments during the study.

Major Changes from the Statistical Analysis Plan 1.0

Major analytical and data handling plans that were changed, added, or deleted from the statistical analysis plan 1.0 are shown below.

Reliability of the Study

<Additional items>

Relevant protocol deviations, reason for subjects enrolled but not randomized and reason for subjects randomized but not treated were added.

<Deleted item>

Reason for withdrawals and dropouts from the study was deleted.

<Reason>

To assess the reliability of the study more appropriately.

Demographic and Other Baseline Characteristics

<Additional items>

Analysis for past pharmacotherapy for cancer and details of immune-modulating concomitant medication were added.

<Reason>

To assess the imbalance of factors which could impact on the evaluation of efficacy and safety between treatment groups.

Primary Efficacy Analysis

Proportional hazards assumption

<Additional items>

Test for examining the proportional hazards assumption for OS was added.

<Reason>

To assess the proportional hazards assumption for OS.

State of Follow-up

<Additional items>

Analyses for current status of follow-up and duration of follow-up for OS were added.

<Reason>

To assess the current status of follow-up for OS.

Exploratory Analytical Methods

<Additional items>

Analyses using the intended investigator's choice therapy subset were added.

<Reason>

To evaluate the efficacy of nivolumab relative to the control regimen from a multilateral perspective.

Secondary Efficacy Analysis

<Additional items>

A secondary definition of PFS was added.

<Reason>

To evaluate the robustness of analysis result using PFS.

Analysis of Safety

Adverse Events

<Additional items>

Multiple occurrences of AEs and other AEOSIs were added.

<Reason>

To evaluate the safety of nivolumab relative to the control group from a multilateral perspective.

PD-L1 Expression Analysis

<Additional item>

Section of PD-L1 expression analysis was added.

<Reason>

To evaluate the expression level of PD-L1 in each treatment group.

Patient Reported Outcomes Analysis

<Additional items>

Change in EQ-5D and EQ-VAS scores from baseline were added.

<Reason >

To evaluate quality of life with nivolumab relative to the control regimen from a multilateral perspective.

Major Changes from the Statistical Analysis Plan 2.0

Major analytical and data handling plans that were changed, added, or deleted from the statistical analysis plan 2.0 are shown below.

Significant Level to be Used

<Additional items>

Analysis population to be used for OS, ORR and PFS when performing the test using the Hierarchical hypothesis testing approach was added.

<Reason >

In this study, two analysis populations of ITT and RES were defined as the analysis population for efficacy, and in order to clarify which analysis population would be tested for each endpoint.

Secondary Efficacy Analysis

<Additional items>

An analysis was added to calculate summary statistics of observation period of subjects for each treatment group.

<Reason >

To examine the observation period of subjects included in this study in more detail.