Statistical Analysis Plan

Study ID: 213403

Official Title of Study: A Randomized, Phase 2, Double-blind Study to Evaluate the Efficacy of Dostarlimab Plus Chemotherapy versus Pembrolizumab Plus Chemotherapy in Metastatic Non-Squamous Non-Small Cell Lung Cancer.

NCT number: NCT04581824

Date of Document: 27-JUL-2022

Division	:	Worldwide Development
Information Type	:	Statistical Analysis Plan (SAP)

Title Page

Protocol Title: A Randomized, Phase 2, Double-blind Study to Evaluate the Efficacy of Dostarlimab Plus Chemotherapy versus Pembrolizumab Plus Chemotherapy in Metastatic Non-Squamous Non-Small Cell Lung Cancer

Protocol Number: 213403

Compound Number: GSK4057190 (Dostarlimab)

Short Title: Efficacy Comparison of Dostarlimab Plus Chemotherapy vs Pembrolizumab Plus Chemotherapy in Participants with Metastatic Non-squamous Non-small Cell Lung Cancer

Acronym: PERLA

Sponsor Name: GlaxoSmithKline Research & Development Limited

Regulatory Agency Identifier Number(s)

Registry	**ID**
IND Number	150997
EudraCT Number	2020-002327-11

SAP Author(s):

Author
PPD Principal Statistician, Dev Biostats Onc
PPD Senior Statistician, Veramed

SAP Team Review Confirmations (E-mail Confirmation):

Reviewer
PPD PPD, Oncology Clinical Development
PPD Clinical Scientist, Oncology Clinical Development
PPD PPD, Clinical Pharmacology (Oncology)
PPD Clinical Development Manager
PPD Clinical Data Manager
PPD PPD, Signal Evaluation and Risk Management

SAP Biostatistics Line Approval (Pharma TMF eSignature):

Approver
PPD PPD, Dev Biostats Onc

TABLE OF CONTENTS

PAGE

Version history

Table 1 SAP Version History Summary

SAP Version	Document Date	Protocol Version (Date) on which SAP is Based	Change	Rationale
1	15-MAR-2022	Version 04 30-NOV-2021	Not Applicable	Original version
2	27-JUL-2022	Version 04 30-NOV-2021	1) Response-evaluable population updated to use BICR-assessment (section 3) 2) Added unconfirmed ORR analysis (sections 4.2.3, 4.5.1.1.1) 3) Specified to use profile-likelihood confidence limits and Efron's method of handling ties (section 4.3.2.1) 4) Added details of Cox Proportional Hazards analysis to PFS (section 4.3.2.2) 5) Dose modifications section updated to reflect the data we have available (section 4.5.4.1) 6) Updated ALT thresholds, to remove 10xULN and add 8xULN (section 4.5.5.1) 7) Added PD-L1 subgroup analysis for PFS and DOR (section 4.6.1)	Updated based on changes and clarifications required following the Early Dry Run delivery

1. INTRODUCTION

The purpose of this SAP is to describe the planned analyses to be included in the Clinical Study Report for Study 213403 (PERLA). Details of the final analyses are provided.

Descriptive study population analyses such as summary of demography and baseline characteristics and additional detail with regards to data handling conventions and the specification of data displays will be provided in the Output and Programming Specification (OPS) document.

1.1. Objectives, Estimands and Endpoints

1.1.1. Objectives and Endpoints

Objectives	Endpoints
Primary	
• To compare the ORR of PD-1 inhibitor dostarlimab vs pembrolizumab administered in combination with chemotherapy as evaluated using RECIST v1.1 based on BICR in participants with metastatic non-squamous NSCLC, without a known EGFR, ALK, ROS-1, or BRAF V600E mutation or other genomic aberration for which an approved targeted therapy is available, who have received no prior treatment of metastatic disease	• The primary efficacy endpoint ORR will be evaluated by RECIST v1.1 based on BICR and will be defined as the proportion of participants with BOR of CR or PR in the analysis population.
Secondary	
• To evaluate the following measures of clinical benefit of PD-1 inhibitor administered in combination with chemotherapy: – OS – PFS evaluated using RECIST v1.1 based on Investigator assessment	• OS will be defined as the time from the date of randomization to the date of death by any cause. • PFS will be evaluated using RECIST v1.1 based on Investigator assessment and will be defined as the time from the date of randomization to the date of PD or death by any cause, whichever occurs first.
• To evaluate the safety of PD-1 inhibitor in combination with chemotherapy	• Assess the incidence of TEAEs, SAEs, irAEs, TEAEs leading to death, and AEs leading to discontinuation occurring while participants are on treatment or up to 90 days after the last dose of study treatment. Clinical laboratory parameters (hematology, chemistry, thyroid function, urinalysis), vital signs, ECOG performance status, ECG parameters, physical examinations, and usage of concomitant medications will be collected.

Objectives	Endpoints
Exploratory	
• To evaluate DOR using RECIST v1.1 based on BICR	• DOR will be evaluated using RECIST v1.1 based on BICR and will be defined as the time from first documented CR or PR until subsequently documented PD, or death, whichever occurs first
• To evaluate ORR using RECIST v1.1 based on Investigator assessment	• ORR will be evaluated using RECIST v1.1 based on Investigator assessment and will be defined as the proportion of participants with BOR of CR or PR in the analysis population
• To evaluate the correlation between PD-L1 expression and efficacy outcomes	• Tumor tissue will be evaluated for PD-L1 expression using IHC and may be correlated with ORR and potentially other clinical endpoints to treatment
• To assess the PK and immunogenicity of dostarlimab and pembrolizumab	• Dostarlimab serum PK analysis will assess C_{min}, C_{max}, $C_{min,ss}$, and $C_{max,ss}$. ADAs will be analyzed in a tiered approach (i.e., Screening, confirmation, titer, and neutralizing antibody assay) using electrochemiluminescence, if appropriate • PK and Immunogenicity for pembrolizumab will be assessed only as needed • Blood cells may be assessed for PDy receptor occupancy at multiple time points to characterize the PK/PDy profile of therapeutic agents
• To evaluate circulating biomarkers in blood that may be predictive of response to PD-1 inhibition in combination with chemotherapy	• ctDNA may be extracted from plasma and analyzed for specific genomic aberrations, including assessment of TMB and mutations in genes related to NSCLC biology, as well as sensitivity or resistance to PD-1 inhibitors • Results from blood-based ctDNA biomarker analyses may be compared with corresponding analyses on tumor tissue samples for concordance, whenever applicable. They may also be correlated with efficacy outcomes.

Objectives	Endpoints
Exploratory (continued)	
• To assess genomic and protein biomarkers in tumor tissue that may be predictive of response to PD-1 inhibition in combination with chemotherapy	• Protein expression may be analyzed by IHC or other techniques to evaluate tumor immune contexture, status of immune checkpoint proteins such as TIM-3 and LAG-3, or other biomarkers associated with efficacy of PD-1 inhibitors • DNA may be extracted from the tumor tissue sample and analyzed for specific genomic aberrations, including assessment of TMB and mutations in genes related to NSCLC biology, as well as other oncogenic lesions associated with sensitivity or resistance to PD-1 inhibitors • RNA may be extracted from the tumor tissue samples and analyzed for gene expression signatures associated with sensitivity or resistance to PD-1 inhibitors
• To evaluate disease- and treatment-related lung cancer symptoms and severity and HRQoL, including TTD in lung cancer symptoms and change from baseline	• TTD in lung cancer symptoms, defined as time from randomization to meaningful deterioration on a composite endpoint of dyspnea, chest pain, and cough, assessed by the EORTC-QLQ-LC13 • Change from baseline as assessed by the EORTC-QLQ-C30 and EORTC-QLQ-LC13 total and domain scores, PGIS, and PGIC and frequency and severity of participant-reported AEs based on PRO-CTCAE and FACT-GP5

Abbreviations: ADA=anti-drug antibody; AE=adverse event; ALK=anaplastic lymphoma kinase; BICR=blinded independent central review; BOR=best overall response; BRAF= proto-oncogene B-raf; C_{max}=maximum concentration; $C_{max,ss}$=C_{max} at steady state; C_{min}=minimum concentration; $C_{min,ss}$=C_{min} at steady state; CR=complete response; ctDNA=circulating tumor DNA; DOR=duration of response; ECG=electrocardiogram; ECOG=Eastern Cooperative Oncology Group; EGFR=epidermal growth factor receptor; EORTC-QLQ-C30=European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire 30-item Core Module; EORTC-QLQ-LC13=European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire 13-item Lung Cancer Module; EOT=End-of-Treatment; FACT-GP5=Functional Assessment of Cancer Therapy-General Population; HRQoL=health-related quality of life; IHC=immunohistochemistry; irAE=immune-related adverse event; LAG-3=lymphocyte-activation gene 3; NSCLC=non-small cell lung cancer; ORR=objective response rate; OS=overall survival; PD=progressive disease; PD-1=programmed cell death protein 1; PD-L1=programmed death-ligand 1; PFS=progression-free survival; PDy=pharmacodynamic; PGIC=Patient Global Impression of Change; PGIS=Patient Global Impression of Severity; PK=pharmacokinetics; PR=partial response; PRO-CTCAE=Patient-reported Outcomes Version of the Common Terminology Criteria for Adverse Events; RECIST=Response Evaluation Criteria in Solid Tumors; ROS-1=receptor tyrosine kinase-1; SAE=serious

adverse event; TEAE=treatment-emergent adverse event; TIM-3=T-cell immunoglobulin and mucin-domain containing-3; TMB=tumor mutational burden; TTD=time to deterioration.

The analyses of pharmacodynamic PK/PDy profile of therapeutic agents and the endpoints related to the circulating biomarkers, genomic and protein biomarkers are not in scope of this SAP. These analyses will be defined in a separate SAP and will be reported outside CSR.

All PRO related analyses (disease- and treatment-related lung cancer symptoms and severity and HRQoL, including TTD in lung cancer symptoms and change from baseline) are not in scope of this SAP. These PRO endpoints will be analysed in the scope of CSR appendix by the MMA group as defined in a separate SAP.

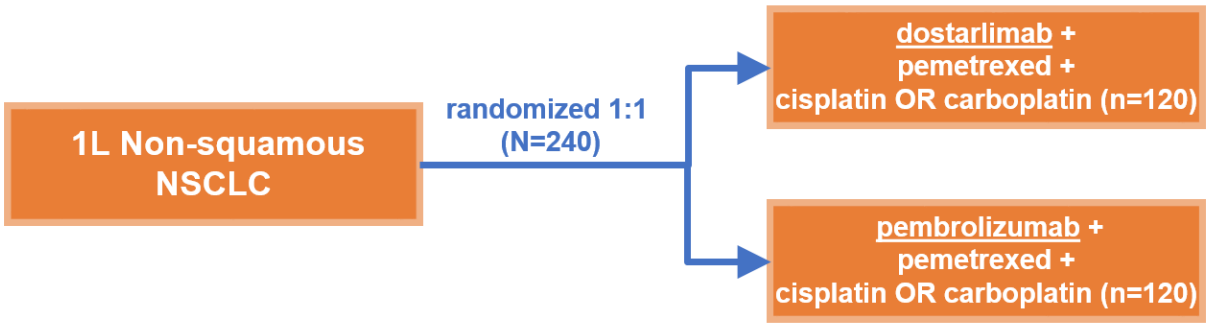
1.1.2. Estimands

Table 2 Estimands

Objective	Estimand Category	Estimand			
		Population	Variable/ Endpoint	Intercurrent Event Strategy	Population Level Summary Measure
Primary Objective: To compare the ORR of dostarlimab vs pembrolizumab administered in combination with chemotherapy as evaluated using RECIST v1.1 based on BICR in participants with metastatic non-squamous NSCLC, without a known EGFR, ALK, ROS-1, or BRAF V600E mutation or other genomic aberration for which an approved targeted therapy is available, who have received no prior treatment of metastatic disease	Primary	Participants with metastatic non-squamous NSCLC, without a known EGFR, ALK, ROS-1, or BRAF V600E mutation or other genomic aberration for which an approved targeted therapy is available, who have received no prior treatment of metastatic disease	ORR	Start of new anticancer therapy: while on treatment strategy Treatment discontinuation: treatment policy strategy	Risk difference between dostarlimab + chemotherapy vs pembrolizumab +chemotherapy

1.2. Study Design

Overview of Study Design and Key Features



Abbreviations: 1L=first-line; NSCLC=non-small cell lung cancer.
Note: Both arms will be stratified by PD-L1 status (TPS <1% vs 1% to 49% vs ≥50%) and smoking status (never vs former/current).

Design Features	• This is a randomized, Phase 2, double-blind, 2-arm study to compare the efficacy and safety of PD-1 inhibitors dostarlimab and pembrolizumab, when administered in combination with chemotherapy, in male and female participants 18 years and older with non-squamous NSCLC without a known sensitizing epidermal growth factor receptor (EGFR), anaplastic lymphoma kinase (ALK), receptor tyrosine kinase-1 (ROS-1), or proto-oncogene B-raf (BRAF) V600E mutation or other genomic aberration for which an approved targeted therapy is available who have not received previous systemic anticancer therapy for metastatic disease.
	• Approximately 240 participants will be randomized in a 1:1 ratio into the dostarlimab plus chemotherapy arm or the pembrolizumab plus chemotherapy arm, such that approximately 120 evaluable participants in each of the 2 arms complete the study.
	• Randomization will be stratified by:
	  ○ PD-L1 status of the tumor (TPS <1% versus 1% to 49% versus ≥50%);
	  ○ Smoking status (never versus former/current).
	• The study consists of a Screening Period (Day -28 to Day -1) for completion of all Screening assessments and subsequent randomization, a Treatment Period, an EOT Visit (within 7 days of the decision to discontinue treatment for any reason), a Safety Follow-up Period with a visit at 30 (+7) and 90 (+7) days after the last dose of study treatment, and a Post-treatment Follow-up Period with assessments occurring 180 days after the last dose of study treatment and every 90 (±14) days thereafter, continuing until death, withdrawal of consent, or the end of study data collection.
	• The study will last approximately 5 years. During the Treatment Period participants will take part in up to 35 cycles with a visit frequency of Q3W (cycles are 21 [±3] days).

Overview of Study Design and Key Features	
Study intervention	• Dostarlimab will be administered through a 30-minute infusion at a dose of 500 mg IV Q3W up to a maximum of 35 cycles total (approximately 24 months). • Pembrolizumab will be administered through a 30-minute infusion at a dose of 200 mg Q3W up to a maximum of 35 cycles total (approximately 24 months). • Chemotherapy (standard of care) will be administered to participants in both treatment arms: o Pemetrexed will be administered at 500 mg/m2 IV through a 10-minute IV infusion Q3W, up to a maximum of 35 cycles total (approximately 24 months). o Platinum chemotherapy (Cisplatin / Carboplatin) will be administered for the first 4 cycles only, following Pemetrexed administration. ▪ If selected by the Investigator, Cisplatin (75 mg/m2) will be administered via IV infusion (approximately 30 minutes after pemetrexed infusion) Q3W for the first 4 cycles. ▪ If selected by the Investigator, Carboplatin will be administered at area under the concentration-time curve 5 mg/mL/min (AUC 5 mg/mL/min) Q3W immediately following the pemetrexed infusion for 4 cycles.
Study intervention Assignment	• Participants who meet eligibility criteria will be randomized in a 1:1 ratio into the dostarlimab plus chemotherapy arm or the pembrolizumab plus chemotherapy arm, such that approximately 120 evaluable participants in each of the 2 arms complete the study. The randomization will be stratified by PD-L1 status of the tumour (TPS <1% vs 1% to 49% vs ≥50%) and smoking status (never vs former/current).
Interim Analysis	• No formal interim analysis is planned for this study. An ad-hoc interim analysis may be performed to inform internal business development decisions whilst keeping the study team blinded to the interim results internal development plan.

2. STATISTICAL HYPOTHESES

The primary efficacy endpoint, ORR using RECIST v1.1 based on BICR, of dostarlimab plus chemotherapy is similar to that of pembrolizumab plus chemotherapy in participants with metastatic non-squamous NSCLC without a known EGFR, ALK, ROS-1, or BRAF V600E mutation or other genomic aberration for which an approved targeted therapy is available and who have not received prior treatment of metastatic disease.

2.1. Multiplicity Adjustment

No adjustments for multiplicity have been made.

3. ANALYSIS SETS

The analysis sets are presented in Table 4.

Table 3 Analysis Sets

Analysis Set	Definition / Criteria	Analyses Evaluated
Screened	• All participants who were screened for eligibility.	Study population
Enrolled	• All participants who entered the study. • Note screening failures (who never passed screening even if rescreened) and participants screened but never enrolled into the study (Reserve, Not Used) are excluded from the Enrolled analysis set as they did not enter the study.	Study population
Intent-to-treat (ITT)	• All participants who were randomized. • Participants will be analyzed according to the treatment assigned at randomization even if no study treatment was received. • Participants who were incorrectly stratified during randomization will be analyzed and presented under the stratum assigned during randomization.	Study Population Efficacy
Response Evaluable	• A subset of the ITT population that will only include participants who meet all of the following criteria. Participants who meet all of the following criteria will be included in the response evaluable population: – Participants receive at least 1 dose of the randomized study treatment. – Participants have measurable disease at baseline per BICR Assessment. – Participants have any post-baseline tumor assessment (unless PD or death is observed before that time, in which case the participant will not be excluded from the response evaluable population). • Participants will be analyzed according to the treatment assigned during randomization.	Selected Efficacy

Analysis Set	Definition / Criteria	Analyses Evaluated
Per protocol (PP)	• All participants in the ITT population who do not have protocol deviations that may significantly impact the interpretation of efficacy results. • Specific criteria that would exclude participants from the PP population are provided in Section 6.2. A detailed specification of the per-protocol population will be provided prior to database lock. • Patients who received the incorrect randomised therapy throughout the study or if patients were administered both randomised treatment at different times at any time of the study, then the patients would be excluded from PP analyses set (the Appendix 2: Exclusions from Per Protocol Population).	Selected Efficacy The PP set will not be analysed if PP population comprises at least 90% of the ITT population.
Safety	• All participants who received at least 1 dose of study treatment. • Participants will be analyzed according to the actual treatment (i.e., the treatment assigned at randomization unless the incorrect treatment(s) was/were received, in which case participants will be analyzed according to the treatment they received for >50% of the time).	Safety
Dostarlimab PK[1]	• All participants who received any amount of dostarlimab and have at least 1 measurable dostarlimab concentration. • Non-quantifiable [NQ] values will be considered as non-missing values.	PK
Pembrolizumab PK[1]	• If Pembrolizumab PK analysis is performed, this will include all participants who received any amount of pembrolizumab and have at least 1 measurable pembrolizumab concentration. • NQ values will be considered as non-missing values.	PK
Dostarlimab immunogenicity	• All participants who received at least 1 dose of dostarlimab and who have at least 1 ADA sample with an assay result.	Immunogenicity
Pembrolizumab immunogenicity	• If Pembrolizumab immunogenicity analysis is performed, this will include all participants who received at least 1 dose of pembrolizumab and who have at least 1 ADA sample with an assay result.	Immunogenicity
COVID-19	• All participants in the Safety set who had a confirmed, probable or suspected COVID-19 case diagnosis.	Selected Demography, Safety

1. Note: PK samples that may be affected by protocol deviations will be reviewed by the study team to determine whether or not the sample will be excluded.

4. STATISTICAL ANALYSES

4.1. General Considerations

4.1.1. General Methodology

The ITT Set will be used for all Study Population analyses and Efficacy analyses and the Safety Analysis Set will be used for all safety analyses, unless otherwise stated.

Analysis and summaries performed on the ITT set will be presented by treatment assigned to the participant during randomisation unless otherwise specified. This will be identified from the randomisation list.

Unless otherwise specified, continuous data will be summarized using descriptive statistics: n, mean, standard deviation (std), median, minimum and maximum. Categorical data will be summarized as the number and percentage of participants in each category.

For the primary endpoint of ORR, the difference in ORR and its 80% and 95% CIs will be reported. For other endpoints, confidence intervals will use 95% confidence levels unless otherwise specified.

4.1.2. Baseline Definition

For all endpoints, unless otherwise specified, the baseline value will be the latest pre-dose assessment with a non-missing value, including those from unscheduled visits. If time is not collected, Day 1 assessments are assumed to be taken prior to first dose and used as baseline. If baseline data is missing no derivation will be performed and baseline will be set to missing

4.1.3. Multicenter Studies

In this multicenter global study, enrolment will be presented by country and site.
Data from all participating centers will be integrated and no controlling for center-effect will be considered in the statistical analyses. It is anticipated that participant accrual will be spread thinly across centers and summaries of data by center would unlikely be informative and will not be provided.

4.1.4. Visit Windows

It is expected that all visits should occur according to the protocol schedule. Unless specified otherwise, by-visit summaries and analyses will be by nominal visit (all data will be tabulated per the evaluation visit as recorded on the eCRF even if the assessment is outside of the visit window for analysis).

4.1.5. Study Population Summaries

The OPS document will provide additional details regarding the participant disposition and study population summaries, including, but not limited to the following:

- Participant status and disposition
- Treatment status and reasons for discontinuation of study treatment
- Screen status and reasons for screen failures
- Participants by country and site ID
- Important protocol deviations
- Study populations
- Demographics and baseline characteristics (including age, gender, ethnicity, race, and stratification factors: smoking status and PD-L1 status)
- Medical history / physical exam
- Advanced/Metastatic NSCLC history
- Prior anti-cancer therapy, prior radiotherapy and prior surgery
- Prior and concomitant medications
- Concomitant procedures

4.2. Primary Endpoint Analyses

The primary efficacy endpoint objective response rate (ORR) will be evaluated by RECIST v1.1 based on blinded independent central review (BICR). The primary analysis of ORR will be based on the ITT population, and a supportive sensitivity analysis will be performed on the response evaluable population and per-protocol population.

The primary analysis of ORR will occur after all enrolled participants have completed the third on-study tumor assessment (approximately 6 months) or have been discontinued from the study, whichever occurs first.

4.2.1. Definition of endpoint

ORR per RECIST v1.1 is defined as the proportion of patients who have a complete response (CR) or partial response (PR) as their best overall response (BOR) based on BICR.

Best Overall Response (BOR)
The best overall response is the best response recorded from the date of randomization until disease progression or initiation of new anti-cancer therapy, whichever is earlier, as assessed by the BICR per RECIST v1.1, using all scans regardless of whether they were scheduled or not. The order from best to worst of the available responses is CR, PR, stable disease (SD), progressive disease (PD) and not evaluable (NE). In order to assign a BOR of CR or PR, a participant initial response

must be confirmed by repeat assessment performed no less than 4 weeks after the initial criteria for response were met.

To assign a status of SD as BOR, the minimum criteria for SD duration of 28 days must be met. If the minimum time for SD is not met, best response will depend on the subsequent assessments. For example, if an assessment of PD follows the assessment of SD and SD does not meet the minimum time requirement, the best response will be PD. Alternatively, participants lost to follow-up after an SD assessment not meeting the minimum time criteria will be considered not evaluable.

4.2.2. Main analytical approach

All analyses will be performed using the ITT population unless otherwise stated. The number and percentage of participants with the BOR in the following response categories will be summarized by treatment arm: CR, PR, overall response (CR+PR), SD, PD and NE. The corresponding 95% CI and point estimate for ORR using the Clopper-Pearson method will also be provided. Participants with unknown or missing responses will be treated as non-responders, i.e., these participants will be included in the denominator when calculating percentages of response.

ORR will also be compared between treatment arms using the Mantel and Haenszel method with Sato's variance estimator and will be stratified by the following factors: PD-L1 status (TPS <1% vs 1% to 49% vs ≥50%) and smoking status (never vs former/current). The difference in ORR and its 80% and 95% CIs will be reported.

The analyses will be performed based on the strata data collected in Interactive Response Technology (IRT) at randomization, even if it is subsequently discovered that these values were incorrect.

4.2.3. Sensitivity analyses

The primary analysis of confirmed ORR evaluated by RECIST v1.1 based on BICR will be repeated using the per-protocol population and the response evaluable population.

Unconfirmed ORR based on BICR assessment will also be analysed using the ITT, per-protocol and the response evaluable population. The endpoint is the same as the confirmed BOR endpoint, with the exception being that CR and PR do not need to be confirmed by repeat assessments.

4.3. Secondary Endpoints Analyses

The secondary efficacy endpoints overall survival (OS) and progression free survival (PFS) using RECIST v1.1 based on Investigator assessment will be evaluated. They will be summarized using the ITT population.

4.3.1. Definition of endpoints

4.3.1.1. Overall Survival (OS)

Overall Survival (OS), defined as the interval of time (in months) from randomization to the date of death due to any cause, regardless of whether the patient withdraws from randomised therapy or receives another anti-cancer therapy. Participants who are alive will be censored at the date of last contact. Any patient not known to have died at the time of analysis will be censored based on the last recorded date on which the patient was known to be alive.

4.3.1.2. Progression-Free Survival per RECIST v1.1 (PFS)

Progression-free survival (PFS) per RECIST v1.1 based on investigator assessment, defined as the interval of time (in months) between the date of randomization and the earlier of the date of disease progression (PD) as assessed by the investigator per RECIST 1.1 criteria and the date of death due to any cause.

Since PFS is interval censored, extended loss to follow-up prior to PD or death increases the uncertainty when the event occurs. As such, PFS will be analyzed censoring for extended time without an adequate assessment to account for missed response assessments prior to disease progression or death. Specifically, if there are two or more assessments which are missing followed by an assessment of PD or death, PFS will be censored at the last adequate assessment prior to PD or death.

The PFS time will always be derived based on scan/assessment dates and not visit dates. The date of progression will be determined based on the earliest of the dates of the component that triggered the progression.

Determination of dates of PFS events and dates for censoring are described in Table 4.

Table 4 Censoring Rules per RECIST v1.1

Situation	Primary Analysis
No or incomplete baseline disease assessments and the participant has not died	Censored at the date of randomization
No adequate[1] post-baseline disease assessments (prior to subsequent anticancer treatment, if initiated) and the participant has not died[2]	Censored at the date of randomization
With adequate post-baseline disease assessments, subsequent anticancer treatment is not initiated, and no documented PD or death	Censored at the date of last adequate radiological disease assessment
With adequate post-baseline disease assessments before the start of subsequent anticancer treatment, and subsequent anticancer treatment is initiated (prior to documented PD or death)[3]	Censored at the date of last adequate radiological disease assessment on or prior to starting new anticancer treatment
PD or death documented after ≤1 missed disease assessment[4]	Progressed at the date of documented PD[5] or death, whichever occurs first
PD or death documented after ≥2 missed disease assessments[4,6]	Censored at the date of last adequate radiological disease assessment prior to the ≥2 missed disease assessment[7]

Abbreviations: CR=complete response; PD=progressive disease; PFS=progression-free survival; PR=partial response; RECIST=response evaluation criteria in solid tumors; SD=stable disease

1. An adequate assessment is defined as an assessment where the investigator assessed (or BICR assessed for BICR endpoints) response is CR, PR, or SD
2. In case participant has documented PD but no other adequate assessments, see scenarios below.
3. If PD and subsequent anticancer treatment occur on the same day, it is assumed that the progression was documented first (i.e., outcome is progression; the date is the date of the assessment for progression).
4. The case where PD or death documented after the initiation of subsequent anticancer treatment is described above and is not included in this situation.
5. The earliest of (i) Date of radiological assessment showing new lesion (if progression is based on new lesion); or (ii) Date of radiological assessment showing unequivocal progression in non-target lesions, or (iii) Date of last radiological assessment of measured lesions (if progression is based on increase in sum of measured lesions).
6. Refer to Section 6.3 for details in extended time without an adequate assessment.
7. The date of randomization will be used if there are no adequate post-baseline assessments.

4.3.2. Main analytical approach

4.3.2.1. Overall Survival (OS)

The distribution of OS for each treatment arm will be estimated using the Kaplan-Meier method (PROC LIFETEST). The median, 25th and 75th percentiles of OS will be estimated and corresponding 95% confidence intervals will be estimated using the Brookmeyer-Crowley method (1982). OS rate at regular time intervals and the corresponding 95% CI will also be estimated from the Kaplan-Meier analysis (time intervals to be defined in the OPS). Kaplan-Meier plots of OS will be presented by treatment arm. Summaries of the number and percentage of patients who have died, those still in survival follow-up, those lost to follow-up and those who have withdrawn consent will be provided along with the median OS for each treatment.

A stratified Cox proportional hazard model will be used to assess the magnitude of the treatment difference (i.e., the hazard ratio). The hazard ratio (HR) and its corresponding 95% confidence interval will be estimated using the Cox's proportional hazard model stratified by PD-L1 status

and smoking status (as entered IRT at randomization) with treatment arm as the sole explanatory variable (PROC PHREG). The 95% confidence interval will be based on the profile-likelihood confidence limits, and Efron's method will be used to handle ties in failure time.

4.3.2.2. Progression-Free Survival per RECIST v1.1 (PFS)

The distribution of PFS for each treatment arm will be estimated using the Kaplan-Meier method (PROC LIFETEST). The median, 25th and 75th percentiles of PFS will be estimated and corresponding 95% confidence intervals will be estimated using the Brookmeyer-Crowley method (1982). PFS rate at regular time intervals and the corresponding 95% CI will also be estimated from the Kaplan-Meier analysis (time intervals to be defined in the OPS).

The following summaries will be also provided for each treatment arm:

- number (%) of patients who were on treatment at the time of progression
- the number (%) of patients who discontinued study treatment prior to progression
- the number (%) of patients who have not progressed and were on treatment (this will be summarised as part of the disposition table)

Kaplan-Meier plots of PFS will be presented by treatment arm.

A stratified Cox proportional hazard model will be used to assess the magnitude of the treatment difference (i.e., the hazard ratio). The hazard ratio (HR) and its corresponding 95% confidence interval will be estimated using the Cox's proportional hazard model stratified by PD-L1 status and smoking status (as entered IRT at randomization) with treatment arm as the sole explanatory variable (PROC PHREG). The 95% confidence interval will be based on the profile-likelihood confidence limits, and Efron's method will be used to handle ties in failure time.

4.4. Supportive Secondary Endpoint(s)

4.4.1. Adverse Events/Serious Adverse Events

The safety analyses will be based on the Safety Analysis Set, unless otherwise specified.

Adverse events analyses including the analysis of adverse events (AEs), Serious AEs (SAEs), AESIs (if applicable) and other significant AEs will be based on GSK Core Data Standards.

For reporting, adverse events will be coded using the standard MedDRA coding dictionary and grouped by system organ class (SOC) and preferred term (PT), or by PT only. Adverse events will be graded by the investigator according to the NCI-CTCAE, Version 5.0. A Standardised MedDRA Query (SMQ) will be used to identify all COVID-19 AEs.

Adverse event analyses will be based on GSK Core Data Standards. A high level overview of AEs will presented in a summary table and a forest plot showing AE incidence rates and relative risks will be provided. Table 5 gives details of further AE summary tables that will be produced.

Table 5 Adverse event summary tables

AE table[1]	By SOC and PT	By SOC, PT and Maximum toxicity grade	By PT
All AEs	Y		
Common (≥5%) AEs			Y
Common (≥5%) non-serious AEs	Y		
Study drug-related AEs	Y		
TEAEs	Y	Y	Y
Study drug related TEAEs (overall and by treatment)	Y	Y	Y
TEAEs leading to treatment discontinuation (overall and by treatment component)	Y		
SAEs	Y	Y	Y
Treatment emergent SAEs	Y		
Study drug-related SAEs (overall and by treatment component)	Y	Y	Y
Grade 3-5 TEAEs		Y	
Study drug-related Grade 3-5 TEAEs (overall and by treatment component)		Y	
AEs leading to death	Y		
Serious fatal and non-fatal drug-related AEs			Y
Study drug-related AEs leading to death (overall and by treatment component)	Y		
Non-serious drug-related AEs			Y
AEs leading to infusion interruption (overall and by treatment component)	Y		
AEs leading to infusion delay (overall and by treatment component component)	Y		
AEs leading to dose reduction (overall and by chemotherapy treatment component)	Y		
AEs leading to treatment discontinuation (overall and by treatment component)	Y		
Immune-related SAEs	Y	Y	Y
Immune-related non-serious AEs	Y	Y	Y
Immune-related AEs leading to treatment discontinuation (overall and by treatment)	Y	Y	
Immune-related Grade 3-5 AEs	Y		
Study drug-related immune-related SAEs (overall and by treatment component)	Y	Y	Y
Infusion-Related Reactions		Y	

1. For all tables the frequency and percentage of experiencing adverse events will be summarized by treatment group.

Additional AE summary tables, listings and figures may be provided for the purposes of disclosure and will be described in the OPS.

A treatment-emergent adverse event (TEAE) is any adverse event that was not present prior to the initiation of study treatment or any adverse event already present that worsens in either severity or frequency following exposure to study treatment until the study treatment end date plus a washout time limit as specified in OPS .
AEs are defined as study drug-related if the investigator classifies the possible relationship any study treatment (Dostarlimab/Pembrolizumab infusion or Pemetrexed infusion or Platinum chemotherapy) as 'Yes'. A worst-case scenario approach will be taken to handle missing relatedness data, i.e. if the relationship is missing then the AE will be assumed to be related to Dostarlimab/Pembrolizumab Infusion and all chemotherapy background therapies (Pemetrexed infusion or Platinum chemotherapy) that the subject is exposed to at the onset date of the AE.

For summaries presenting AEs by maximum grade, the following algorithms for counting participants will be used:

- **Preferred term row:** Participants experiencing the same AE preferred term several times with different grades will only be counted once with the maximum grade.
- **SOC term row:** Participants experiencing the same AE SOC several times with different grades will only be counted once with the maximum grade.
- **Any event row:** Each participant with at least one adverse event will be counted only once at the maximum grade no matter how many events they have.

4.4.1.1. Adverse Events of Special Interest

An adverse event of special interest (AESI) is any AE (serious or nonserious) that is of scientific and medical concern specific to the study treatment, for which ongoing monitoring and rapid communication by the Investigator to the Sponsor are appropriate.

During clinical development, the following AESI categories were defined at the compound level based on the potential of being drug related as seen with other compounds of the same or similar class. These AESI categories are outlined in Table 6. Analyses of AESIs are based on pre-defined MedDRA PTs.

Table 6 Adverse Events of Special Interest

AESI	Further details and definitions[1]
Immune-related AEs	Any >=Grade 2 AEs identified by PTs based on a pre-specified search list of PTs, documented in a version-controlled repository, and finalized for analysis of the current study data prior to database freeze.
Infusion-related reactions	Any dostarlimab related AEs which occurred on or up to 1 day after dostarlimab infusion and identified by PTs as documented in OPS.

1. Both lists will be finalized prior to the database freeze based on the most recent MedDRA version at the time of the data cutoff and will be documented in OPS.

4.5. Exploratory Endpoints Analyses

4.5.1. Exploratory Efficacy endpoints

4.5.1.1. Definition of endpoint

4.5.1.1.1. ORR

Confirmed ORR per RECIST v1.1 based on investigator assessment is defined as the proportion of patients who have a complete response (CR) or partial response (PR) as their best overall response (BOR) based on investigator assessment. In order to assign a BOR of CR or PR, a participant initial response must be confirmed by repeat assessment performed by the investigator no less than 4 weeks after the initial criteria for response were met.

Unconfirmed ORR based on investigator assessment will also be analysed.

4.5.1.1.2. DOR

Duration of confirmed response (DoR) per RECIST v1.1 based on BICR, defined as the interval of time (in months) from first documented evidence of PR or better to the time when disease progression (PD) is documented as assessed by the BICR per RECIST 1.1, or death due to any cause among participants with a PR or better as the BOR. Censoring rule will follow those for PFS as specified in Table 4.

Confirmed DOR based on investigator assessment will also be analysed in the same way. Unconfirmed DOR based on BICR as well as unconfirmed DOR based on investigator assessment will also be analysed.

4.5.1.1.3. Maximum Percent Reduction from Baseline in Tumor Measurement

The Maximum Percent Reduction from Baseline in Tumor Measurement based on BICR is defined as the biggest percent reduction from the baseline tumour measurements in the sum of the longest diameters of the target lesions, assessed by the BICR.

4.5.1.2. Main analytical approach

4.5.1.2.1. ORR

ORR per RECIST v1.1 based on investigator assessment will be summarized using the ITT and per-protocol analysis set in the same way as the primary analysis in section 4.2.2, for both confirmed and unconfirmed BOR.

4.5.1.2.2. DOR

Both confirmed and unconfirmed DoR per RECIST v1.1 based on BICR and investigator assessment will be summarized using the ITT and per-protocol analysis set in the same way as the secondary PFS analysis in section 4.3.2. Median DoR calculated from the KM curve will be summarised. Only patients who have a confirmed response will be included in the summary table for confirmed DOR.

Kaplan-Meier plots of DOR will be presented by treatment arm.

4.5.1.2.3. Maximum Percent Reduction from Baseline in Tumor Measurement

A waterfall plot will present the maximum percent reduction from Baseline in BICR-assessed Tumor Measurement for each subject by treatment. The subject's BoR based on the BICR will be indicated on the plot.

4.5.2. Pharmacokinetic Analyses

Dostarlimab and pembrolizumab (if analyzed) concentration-time data will be listed for each participant and summarized by descriptive statistics at PK time points in the study. This includes pre-dose PK samples at cycles 1, 2, 5, 11, and every 6 cycles thereafter (up to 35 cycles total), end of infusion samples on day 1 of cycles 1, 2, and 5, PK samples collected at EOT, 90-day Safety Follow-up Visit, and at the 180-day Post-treatment Follow-up Visit. Concentrations that are below the limit of quantitation (BLQ) will be treated as zero for the computation of descriptive statistics and for graphical presentations.

The following PK parameters will be determined, if data permit:

- minimum concentration (Cmin)
- maximum concentration (Cmax)
- Cmin at steady state
- Cmax at steady state

Serum concentrations and PK parameter estimates will be presented using mean, standard deviation, coefficient of variation (CV), geometric mean, geometric mean CV, median, minimum, and maximum.

Exclusions from PK analyses set will be documented and reported.

4.5.2.1. Drug Concentration Measures

Pharmacokinetic Concentration Data	
Descriptive Summary Statistics, Graphical Displays and Listings	Refer to the GSK Standard PK Display Standard. Refer to the GSK Standard Statistical Display Principle 6.06.1. Note: Concentration values will be imputed as per GUI_51487 for descriptive summary statistics/analysis and summarized graphical displays only.
Pharmacokinetic Parameter Data	
Descriptive Summary Statistics, Graphical Displays and Listings	Refer to GSK Standard PK Display Standard.

4.5.2.2. Summary Measure

Serum concentration will be summarized per PK time point per analyte (Dostarlimab, and Pembrolizumab if analyzed). Unscheduled measurements will not be included in summary statistics. Serum concentrations may be flagged for exclusion and excluded from summary measure if deemed necessary due to data errors, however all serum concentrations will be included in listings along with the reason for exclusion if applicable. Exclusion flags will be generated based on reasons including but not limited to the following:

i) non-BQL PK sample collected prior first dose administered

ii) pre-dose sample collected at/post start of infusion

iii) pre-dose sample not collected within 24 hour prior to the start of infusion

iv) post-dose sample collected before, during, or at start of infusion

v) duplicate sampling time points

vi) dose reduction>20% from the corresponding dose for the PK sample,

vii) dosing errors/interruptions/missed doses from the corresponding dose for the PK sample

In the case of a retest of a scheduled assessment, the earliest available measurement for that scheduled time (i.e., the original assessment) will be used for summaries unless flagged as invalid.

Listings will include all scheduled, unscheduled, retest, and early discontinuation data.

4.5.2.3. Population of Interest

The primary pharmacokinetic analyses will be based on the Dostarlimab PK population and Pembrolizumab PK population (if analyzed), unless otherwise specified.

4.5.2.4. Strategy for Intercurrent (Post-Randomization) Events

Missing data will not be imputed, regardless of the reasons.

4.5.2.5. Statistical Analyses / Methods

Details of the planned displays are provided in the OPS and will be based on GSK Data Standards and statistical principles.

Unless otherwise specified, endpoints / variables defined in Section 4.5.3.2 will be summarized using descriptive statistics, graphically presented (where appropriate) and listed.

4.5.2.6. Population Pharmacokinetic (POPPK) Analysis

Data from this study may be combined with data from other studies for a population PK, analysis, which will be performed if needed, under a separate SAP, and will be reported separately.

4.5.3. Immunogenicity Analyses

For each subject, the anti-dostarlimab (drug) antibody (ADA) results, titers, and neutralizing antibody assay results will be listed for each assessment timepoint. Descriptive summary statistics for the immunogenicity results will be provided.. Laboratory evaluations will be based on GSK Core Data Standards. The same will be done for Pembrolizumab if analyzed.

The primary immunogenicity analyses will be based on the Dostarlimab immunogenicity population and Pembrolizumab immunogenicity population, if analyzed, unless otherwise specified.

4.5.4. Other Safety Analyses

4.5.4.1. Extent of Exposure

Extent of exposure will be summarized according to GSK Oncology Data Standards. Exposure to Dostarlimab/Pembrolizumab, Pemetrexed Infusion and Platinum Chemotherapy will be summarized separately. The summaries will include:

- Number of treatment cycles
- Treatment duration (months)
- Dose intensity

- Actual dose intensity
- Intended dose intensity (mg/weeks)
  - Dostarlimab 500mg/Q3W
  - Pembrolizumab 200mg /Q3W
- Relative dose intensity
- Cumulative actual dose

Missed doses will be summarized by number of missed doses and reasons for the missed doses.

Dose reductions will be summarized by number of dose reductions, both overall and by planned time.

Dose delays will be summarized by number of delays.

Infusion Interruptions will be summarized by number of infusion interruptions and reasons for the interruptions.

Duration of exposure to the overall study treatment will be summarized.

A swimmer plot of Duration of treatment will be presented at an individual patient level, paged by treatment arm.

4.5.4.2. Cardiovascular Events

AEs and SAEs that are considered as cardiovascular events are detailed in Appendix 8 of the protocol. Cardiovascular event data captured in the specific cardiovascular event section of the eCRF will be listed in the form of patient profiles. A summary of family history of cardiovascular events will be produced.

4.5.4.3. Impact of COVID-19 Pandemic on Adverse Event Reporting

The impact of the COVID-19 pandemic on the safety results will be assessed. As this study was initiated after pandemic measures began in all participating countries, summaries split by pre- and post-pandemic measures will not be produced.

Summaries of COVID-19 assessments and symptoms will be produced based on GSK Core Data Standards.

If greater than 24 participants have a suspected, probable or confirmed COVID 19 infection then the following displays will be produced:

- Summary of COVID-19 Assessments for Subjects with COVID-19 Adverse Events

- Summary of COVID-19 Additional Assessments for Subjects with COVID-19 Adverse Events
- Summary of COVID-19 symptoms for subjects with COVID-19 Adverse Events
- Summary of Current (and/or Past) Medical Conditions for Subjects with COVID-19 Adverse Events
- Summary of Baseline Characteristics for Subjects with COVID-19 Adverse Events
- Summary of Exposure to Study Treatment for Subjects with COVID-19 Adverse Events

If >10% participants report ≥1 COVID-19 AE, then the following data displays will be produced:

- Summary of Onset and Duration of the First Occurrence of COVID-19 AEs (produced overall, and for each preferred term mapping to the coronavirus infections high-level term)
- Summary of Characteristics of COVID-19 AEs (produced overall, and for each preferred term mapping to the coronavirus infections high-level term)
- Summary of Worst Case Chemistry Results by Maximum Grade Increase Post-Baseline Relative to Baseline for Subjects with Suspected, Probable or Confirmed COVID-19 case diagnosis
- Summary of Worst Case Haematology Results by Maximum Grade Increase Post-Baseline Relative to Baseline for Subjects with Suspected, Probable or Confirmed COVID-19 case diagnosis

Additionally, for SAEs of COVID-19, the following data displays will be produced:

- Summary of Onset and Duration of the First Occurrence of a Coronavirus Infection SAE (produced overall, and for each preferred term mapping to the coronavirus infections high-level term)
- Summary of Characteristics of Coronavirus Infection SAEs (produced overall, and for each preferred term mapping to the coronavirus infections high-level term)

4.5.5. Additional Safety Assessments

4.5.5.1. Laboratory Data

Laboratory evaluations of chemistry, hematology and urinalysis will be summarized based on GSK Core Data Standards and will be further described in the OPS document. A continuous summary of thyroid panel laboratory assessments will also be produced.

Summaries of worst case grade increase from baseline grade will be provided for all the lab tests that are gradable by CTCAE v5.0. These summaries will display the number and percentage of subjects with a maximum post-baseline grade increasing from their baseline grade. Any increase in grade from baseline will be summarized along with any increase to a maximum grade of 3 and any increase to a maximum grade of 4. Missing baseline grade will be assumed as grade 0. For laboratory tests that are graded for both low and high values, summaries will be done separately and labelled by direction, e.g., sodium will be summarized as hyponatremia and hypernatremia separately.

For lab tests that are not gradable by CTCAE v5.0, summaries of worst case changes from baseline with respect to normal range will be generated. Decreases to low, changes to normal or no changes from baseline, and increases to high will be summarized for the worst case post-baseline. If a subject has a decrease to low and an increase to high during the same time interval, then the subject is counted in both the "Decrease to Low" categories and the "Increase to High" categories.

For hematology, chemistry and thyroid function, summaries of change from baseline will be produced.

Separate summary tables for hematology and chemistry laboratory tests will be produced. Liver function laboratory tests will be included with chemistry lab tests. The determination of the worst-case post baseline lab values will consider both planned and unscheduled assessments. Participants with missing baseline values are assumed to have a normal/Grade 0 baseline value. The percentages are based on the number of participants in the treatment group with baseline and post-baseline data.

Summaries of hepatobiliary laboratory events including possible Hy's law cases will be provided in addition to what has been described above. The number of participants fulfilling the following criteria will be summarized:

- ALT ≥ 3×ULN, ALT ≥ 5×ULN, ALT ≥ 8×ULN, ALT ≥ 20×ULN
- Total bilirubin ≥ 2×ULN
- Concurrent ALT ≥ 3×ULN and total bilirubin ≥ 2×ULN
  Concurrent ALT ≥ 3×ULN and INR > 1.5
- Concurrent ALT ≥ 3×ULN and total bilirubin ≥ 2×ULN and ALP < 2×ULN
- Time from First Dose to First ALT Elevation >3xULN (days)
- Hepatocellular injury defined as ((ALT/ALT ULN)/(ALP/ALP ULN)) >5 and ALT >3xULN.

The summary will be produced for worst case post baseline only.

A summary of liver monitoring/stopping event reporting will be provided.

An e-DISH plot of maximum post baseline total bilirubin versus maximum post baseline ALT will be created. A similar plot will also be produced for maximum post baseline ALT versus baseline ALT.

4.5.5.2. Vital Signs

Vital signs evaluations will be summarized based on GSK Core Data Standards and will be further described in the OPS document. Pulse Oximetry evaluations will be included in the vital signs displays.

Summaries of vital signs results by potential clinical importance (PCI) criteria for the worst-case post-baseline relative to baseline will be produced. Participants will be counted twice in this summary if they have post-baseline values to have changed 'To High' and 'To Low' relative to baseline. PCI criteria for vital signs are defined in the OPS.

Summaries of grade increase in systolic blood pressure (SBP) and diastolic blood pressure (DBP) will be provided separately. These summaries will display the number and percentage of participants with any grade increase, increase to Grade 2 and increase to Grade 3 for worst case post-baseline only. The grade definition for SBP is: Grade 0 (<120), Grade 1 (120-139), Grade 2 (140-159), Grade 3 (>=160). The grade definition for DBP is: Grade 0 (<80), Grade 1 (80-89), Grade 2 (90-99), Grade 3 (>=100). The summaries will be produced for worst case post baseline only.

Summaries of Heart rate (HR) will also be provided and categorized into the clinical concern ranges which are specific to HR: ≤ 50 bpm and decrease from baseline ≥ 20 bpm or ≥ 120 bpm and increase from baseline ≥ 20 bpm. A summary of absolute value and change in HR will display the number and percentage of subjects with a change within each range for the worst case post-baseline only. Subjects with missing baseline values will be excluded from this summary.

4.5.5.3. ECG

ECG evaluations will be summarized based on GSK Core Data Standards. Summaries of ECG findings and change from baseline in ECG values will be produced.

The QTc data analysis will use the collected values based on Fridericia's formula. Any QTc values collected based on Bazett's formula will be converted prior to analysis. The QTc values based on Fridericia formula will be rounded to the integer and the values will be categorized into the following CTCAE grade and ranges: Grade 0 (<450 msec), Grade 1 (450-480 msec), Grade 2 (481-500 msec), and Grade 3 (≥501 msec). Summaries of grade increase will be provided. These summaries will display the number and percentage of subjects with any grade increase, increase to grade 2 and increase to grade 3 for the worst case post-baseline only. Missing baseline grade will be assumed as grade 0.

The changes in QTc values will be categorized into the clinical concern ranges which are specific to changes in QTc: 31-60 and >60 msec. A summary of change in QTc value will display the number and percentage of subjects with a change within each range for the worst case post-baseline only. Subjects with missing baseline values will be excluded from this summary.

Summaries of PR interval will also be provided and categorized into the clinical concern ranges which are specific to PR interval: ≥ 220 ms and increase from baseline ≥ 20 ms. A summary of absolute value and change in PR Interval will display the number and percentage of subjects with a change within each range for the worst case post-baseline only. Subjects with missing baseline values will be excluded from this summary.

Summaries of QRS will also be provided and categorized into the clinical concern ranges which are specific to PR interval: ≥ 120 ms. A summary of absolute value will display the number and percentage of subjects with a value within the range for the worst case post-baseline only. Subjects with missing baseline values will be excluded from this summary.

4.5.5.4. ECOG Performance Status

ECOG performance status will be summarized by visit. In addition, the ECOG shift from baseline to highest post-baseline score (including all post-baseline scheduled and unscheduled visits) will be summarized.

4.5.5.5. Deaths

Deaths will be summarized and listed based on GSK Core Data Standards and will be further described in the OPS document.

4.6. Other Analyses

4.6.1. Subgroup analyses

The following subgroup analyses will be performed comparing the primary estimand of the primary endpoint (Confirmed ORR per RECIST v1.1 based on BICR, ITT population) between treatments if data permits.

Subgroup	Categories[1]
Age at randomization	<65, ≥65 years of age
Sex	Male, Female
ECOG PS at Baseline	0, 1
Region of Enrollment	Asia, Rest of the World
Smoking status[2]	(Current or Former), Never
PD-L1 TPS Status[2]	TPS <1%, TPS 1-49%, TPS ≥50% and TPS ≥1%
Brain Metastases at Baseline	Yes, No
Platinum agent	Cisplatin, Carboplatin

1. If the percentage of participants is small within a particular subgroup, then the subgroup categories may be refined prior to unblinding the trial.
2. As per the strata data collected in Interactive Response Technology (IRT) at randomization, even if it is subsequently discovered that these values were incorrect.

The purpose of the subgroup analyses is to assess the consistency of treatment effect across expected prognostic factors.

Subgroup analyses by PD-L1 TPS status will also be performed for the following endpoints, based on the ITT population:

- PFS Investigator-Assessed per RECIST v1.1
- Confirmed DOR BICR-Assessed per RECIST v1.1

4.7. Interim Analyses

No formal interim analysis is planned for this study. An ad-hoc data look may be performed to inform internal development plan and the corresponding analysis plan will be documented in a separate OPS document.

4.8. Changes to Protocol Defined Analyses

Changes from the originally planned statistical analysis specified in the protocol amendment 3 (Dated: 17-JUN-2021) are outlined below.

Protocol	Statistical Analysis Plan	Rationale for Changes
• No formal interim analysis will be done for this study.	• No formal interim analysis is planned for this study. An ad-hoc data look may be performed to inform internal development plan and the corresponding analysis plan will be documented in a separate OPS.	• To incorporate the potential data look to inform internal development plan.
• There is no protocol defined adverse events of special interest (AESI) for Dostarlimab.	• Analyses of AESIs are based on pre-defined MedDRA PTs.	• During clinical development, AESI categories were defined at the compound level based on the potential of being drug related as seen with other compounds of the same or similar class.

5. SAMPLE SIZE DETERMINATION

Approximately 240 participants will be randomly assigned to 2 study treatments, such that approximately 120 evaluable participants in each of the 2 arms complete the study.

Participants will be randomized in a 1:1 ratio into either the dostarlimab plus chemotherapy arm or the pembrolizumab plus chemotherapy arm.

With 240 participants (120 in each treatment arm), the study has 85% power to detect a 15% difference in the ORR between the 2 arms at the 10% one-sided type I error rate when the true ORR is 45% for both treatment groups.

6. SUPPORTING DOCUMENTATION

6.1. Appendix 1 Abbreviations and Trademarks

6.1.1. List of Abbreviations

Abbreviation	Description
ADA	Anti-Drug Antibody
AE	Adverse Event
AESI	Adverse Event of Special Interest
ALT	Alanine Aminotransferase
BICR	Blinded Independent Central Review
BoR	Best Overall Response
CI	Confidence Interval
CPMS	Clinical Pharmacology Modelling and Simulation
CR	Complete Response
CTCAE	Common Terminology Criteria for Adverse Events
ctDNA	Circulating Tumor
DCR	Disease Control Rate
DNA	Deoxyribonucleic Acid
DoR	Duration of Response
ECG	Electrocardiography
ECOG	Electrocardiography
eCRF	Electronic Case Record Form
EGFR	Epidermal Growth Factor Receptor
EORTC-QLQ-C30	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire 30-item Core Module
EORTC-QLQ-LC13	European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire 13-item Lung Cancer Module
FACT-G	Functional Assessment of Cancer Therapy – General (Item GP5)
GSK	GlaxoSmithKline
HR	Hazard Ratio
HRQoL	Health-related Quality of Life

Abbreviation	Description
IHC	Immunohistochemistry
irAE	Immune-Related Adverse Event
ITT	Intent-To-Treat
MedDRA	Medical Dictionary for Regulatory Affairs
NE	Not Evaluable
NQ	Non-quantifiable
NSCLC	Non-Small Cell Lung Cancer
OPS	Output and Programming Specification
ORR	Objective Response Rate
OS	Overall Survival
PCI	Potential Clinical Importance
PD	Progressive Disease
PDy	Pharmacodynamic
PFS	Progression Free Survival
PGIC	Patient Global Impression of Change
PGIS	Patient Global Impression of Severity
PK	Pharmacokinetic
PP	Per-Protocol
PR	Partial Response
PRO	Patient-Reported Outcomes
PRO-CTCAE	Patient-Reported Outcomes Version of the Common Terminology Criteria for Adverse Events
PT	Preferred Term
QTc	QT Interval Corrected for Heart Rate
Q3W	Every Three Weeks
RECIST	Response Evaluation Criteria in Solid Tumors
RNA	Ribonucleic Acid
SAP	Statistical Analysis Plan
SAE	Serious Adverse Event
SD	Stable Disease
SMQ	Standardised MedDRA Query

Abbreviation	Description
SOC	System Order Class
TEAE	Treatment-Emergent Adverse Event
TTD	Time to Deterioration
ULN	Upper Limit of Normal

6.1.2. Trademarks

Trademarks of the GlaxoSmithKline Group of Companies
None

Trademarks not owned by the GlaxoSmithKline Group of Companies
MedDRA

6.2. Appendix 2: Exclusions from Per Protocol Population

The per-protocol population include all participants in the ITT who do not have protocol deviations that may significantly impact the interpretation of efficacy results. All decisions to exclude participants from the per-protocol population will be made prior to database lock.

Specifically, a participant will be excluded from the per-protocol population if meeting any of the following criteria that could significantly impact the interpretation of efficacy results.

Exclusion criteria[1]	**Protocol deviation category**	**Protocol deviation subcategory**
Eligibility Criteria Not Met	Eligibility Criteria Not Met	Inclusion/Exclusion criteria not met
Incorrect treatment of Dostarlimab/Pembrolizumab was administered at any time in the study*	Wrong study treatment/administration/dose	Wrong study treatment or assignment administered
Use of other concomitant anti-cancer therapy or other concomitant investigational agents*	Excluded medication, vaccine or device	Medication, excluded by the protocol, was administered
Prohibited concomitant medication, non-drug therapy and vaccine administered (as noted in Section 6.9.1 of the protocol that could significantly impact the efficacy results)	Excluded medication, vaccine or device	All subcategories
Any unblinding event	Study procedures	Study blinding/unblinding procedures
Absence of measurable disease assessment at baseline per BICR*	N/A	N/A
Absence of at least one (1) evaluable post-baseline disease assessment per BICR* (defined as the following: either missing or NE)	N/A	N/A
Receipt of less than two (2) cycles of study treatment	N/A	N/A

[1] The criterion marked with the symbol * is considered as a firm criterion that will impact the interpretation of efficacy results significantly; the criterion marked without the symbol * will be evaluated case-by-case.

6.3. Appendix 3: Extended Time Without an Adequate Assessment

PFS
Given the scheduled disease assessment (i.e. starts at week 6, second at week 12 and then every 9 weeks; if study treatment continues after Week 48 then assessments will be every 12 weeks thereafter until discontinuation of study treatment due to disease progression with clinical instability, start of subsequent anticancer treatment, withdrawal of informed consent, or death, whichever comes first) the definition of 2 missed disease assessments will change. The following rules will be used for identifying the duration of extended time without an adequate assessment for PFS.

If the time difference between the event (PD/death) and last adequate disease assessment prior to the new anticancer therapy is more than the window, PFS will be censored at the last adequate disease assessment prior to the event (PD/death) and the new anticancer therapy.

- If the event is after Week 12 + 7 days and on or prior to week 48 - 7 days, then a subject will be identified as extended time without an adequate assessment if the subject did not have an adequate assessment during the time period of 140 days (18 weeks + 2-week windows) prior to the event;
- Else if the event is after Week 48 - 7 days then a subject will be identified as extended time without an adequate assessment if the subject did not have an adequate assessment during the time period of 182 days (24 weeks + 2-week windows).

7. REFERENCES

Brookmeyer R, Crowley J. A confidence interval for the median survival time. Biometrics. 1982 Mar 1:29-41.