A Phase 3, Global, Multi-Center, Double-Blind, Randomized, Efficacy Study of Zolbetuximab (IMAB362) Plus CAPOX Compared with Placebo Plus CAPOX as First-line Treatment of Subjects with Claudin (CLDN) 18.2-Positive, HER2-Negative, Locally Advanced Unresectable or Metastatic Gastric or Gastroesophageal Junction (GEJ) Adenocarcinoma

STATISTICAL ANALYSIS PLAN

ERRATUM

8951-CL-0302

20 Jan 2023

The purpose of this erratum is to correct the typo in the header of SAP final version 2.0

The original SAP final version 2.0 header states "SAP draft version 0.5". It should be "SAP final version 2.0".

Author: _____ Date: _____

PPD

PPD

PPD Biostatistics

STATISTICAL ANALYSIS PLAN

Final Version 2.0, 13-Sep-2022

A Phase 3, Global, Multi-Center, Double-Blind, Randomized, Efficacy Study of Zolbetuximab (IMAB362) Plus CAPOX Compared with Placebo Plus CAPOX as First-line Treatment of Subjects with Claudin (CLDN) 18.2-Positive, HER2-Negative, Locally Advanced Unresectable or Metastatic Gastric or Gastroesophageal Junction (GEJ) Adenocarcinoma

ISN/Protocol 8951-CL-0302

IND 129598

EudraCT 2018-000519-26

Sponsor:

Astellas Pharma Global Development, Inc. (APGD)

1 Astellas Way

Northbrook, IL 60062

Table of Contents

I. LIST OF ABBREVIATIONS AND KEY TERMS

List of Abbreviations

Abbreviations	Description of abbreviations
ADA	anti-drug antibody
AE	adverse event
AESI	adverse event of special interest
ALP	alkaline phosphatase
ALT	alanine aminotransferase (GPT)
AST	aspartate aminotransferase (GOT)
ATC	Anatomical Therapeutic Chemical Classification System
BMI	body mass index
BSA	body surface area
CAPOX	Capecitabine and Oxaliplatin
C1D1	Cycle 1 Day 1
CI	confidence interval
CLDN	Claudin
C_{trough}	trough concentration
CMH	Cochran-Mantel-Haenszel
CMQ	Customized Medical Dictionary for Regulatory Activities (MedDRA) Query
CR	complete response
CSR	Clinical study report
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
EORTC	European Organization for Research and Treatment of Cancer
EQ-5D-5L	EuroQOL Five Dimensions Questionnaire 5L
FAS	full analysis set
GEJ	gastroesophageal junction
GP	Global Pain
HER2	human epidermal growth factor receptor 2
HR	hazard ratio
HRQoL	health-related quality of life
HRU	Health Resource Utilization
IDAC	independent data analysis center
IDMC	independent data monitoring committee
INR	international normalized ratio

Abbreviations	Description of abbreviations
IRC	independent review committee
IRR	infusion-related reaction
IRT	interactive response technology
ISN	international study number
IV	Intravenous, intravenously
NCI	National Cancer Institute
NE	not evaluable
ORR	objective response rate
OS	overall survival
PD	progressive disease
PFS	progression free survival
PFS2	progression free survival following subsequent anti-cancer treatment
PGx	pharmacogenomics
PKAS	pharmacokinetics analysis set
PR	partial response
PT	preferred term
QLQ-C30	Quality of Life Questionnaire - Core Questionnaire
QLQ-OG25	Quality of Life Questionnaire - Oesophago-Gastric Module 25 (OG-25)
QoL	Quality of life
QTc	QT interval corrected
QTcF	Fridericia-corrected QT interval
RECIST	Response Evaluation Criteria in Solid Tumors
SAE	serious adverse event
SAF	safety analysis set
SAP	statistical analysis plan
SD	stable disease
SOC	system organ class
SOD	Sum of diameters
SSQ	Standardized Search Query
SMQ	Standardized MedDRA Query
TEAE	treatment-emergent adverse event
TSH	thyroid stimulating hormone
TTP	time to progression
ULN	upper limit of normal
VAS	visual analog scale
WHO	World health organization

List of Key Terms

Terms	Definition of terms
Baseline	Assessments of subjects as they enter a trial before they receive any treatment.
Endpoint	Variable that pertains to the efficacy or safety evaluations of a trial.
Enroll	To register or enter a subject into a clinical trial. NOTE: Once a subject has been randomized, the clinical trial protocol applies to the subject.
Intervention	The drug, device, therapy or process under investigation in a clinical study that is believed to have an effect on outcomes of interest in a study (e.g., health-related quality of life, efficacy, safety, pharmacoeconomics).
Investigational period	Period of time where major interests of protocol objectives are observed, and where the test drug or comparative drug (sometimes without randomization) is usually given to a subject, and continues until the last assessment after completing administration of the test drug or comparative drug.
Post investigational period	Period of time after the last assessment of the protocol. Follow-up observations for sustained adverse events and/or survival are done in this period.
Randomization	The process of assigning trial subjects to treatment or control groups using an element of chance to determine assignments in order to reduce bias.
Screening	A process of active consideration of potential subjects for enrollment in a trial.
Screen failure	Potential subject who did not meet 1 or more criteria required for participation in a trial.
Screening period	Period of time before entering the investigational period, usually from the time when a subject signs the consent until just before the test drug or comparative drug (sometimes without randomization) is given to a subject.
Study period	Period of time from the first site initiation date to the last site completing the study.
Study treatment	Includes Zolbetuximab/placebo and all components of CAPOX
Variable	Any entity that varies; any attribute, phenomenon or event that can have different qualitative or quantitative values.

1 INTRODUCTION

This Statistical Analysis Plan (SAP) contains technical and detailed elaboration of the principal features of the analysis described in the protocol, and includes procedures for executing the statistical analysis to fulfil the objectives of the study. The final SAP will be approved prior to database hard lock for the final PFS (and interim OS) analysis.

Changes that affected the statistical analyses from the planned analyses in the SAP will be documented in the Clinical Study Report (CSR).

2 STUDY OBJECTIVES AND DESIGN

2.1 Study Objectives

2.1.1 Primary Objective

The primary objective is to evaluate the efficacy of Zolbetuximab plus CAPOX compared with placebo plus CAPOX (as first-line treatment) as measured by progression free survival (PFS) in subjects with Claudin (CLDN)18.2 positive, human epidermal growth factor receptor 2 (HER2)–negative locally advanced unresectable or metastatic gastric and gastroesophageal junction (GEJ) adenocarcinoma.

2.1.2 Secondary Objectives

The secondary objectives are:

- To evaluate efficacy as measured by Overall Survival (OS) as a key secondary objective
- To evaluate the physical function (PF), OG25-Pain and GHS/QoL scores as measured by European Organization for Research and Treatment of Cancer (EORTC) as a key secondary objective
- To evaluate efficacy as measured by Objective Response Rate (ORR)
- To evaluate efficacy as measured by Duration of Response (DOR)
- To evaluate safety and tolerability of Zolbetuximab
- To further evaluate other health-related quality of life (HRQoL) using additional parameters as measured by EORTC QLQ-C30 and QLQ-OG25 plus STO22 Belching subscale, Global Pain (GP) and the EuroQOL Five Dimensions Questionnaire 5L (EQ-5D-5L) questionnaires
- To evaluate the pharmacokinetics of Zolbetuximab
- To evaluate the immunogenicity profile of Zolbetuximab

2.1.3 Exploratory Objectives

The exploratory objectives are:

- To evaluate efficacy as measured by Time to Progression (TTP)
- To evaluate PFS following subsequent anti-cancer treatment (PFS2)
- To evaluate Disease Control Rate (DCR)
- To evaluate potential genomic and/or other biomarkers that may correlate with treatment outcome to Zolbetuximab and CAPOX.

- To evaluate Health Resource Utilization (HRU)

2.2 Study Design

This global, multi-center, double-blind, 1:1 randomized, phase 3 study will evaluate efficacy of Zolbetuximab plus CAPOX versus placebo plus CAPOX as first-line treatment in subjects with CLDN18.2-positive, HER2-negative locally advanced unresectable or metastatic gastric and GEJ adenocarcinoma.

PFS as assessed by the Independent Review Committee (IRC) is the primary outcome. Secondary outcomes include OS, time to confirmed deterioration (TTCD) in PF, OG25-PA and GHS/QoL, ORR, DOR, safety and tolerability, other HRQoL, pharmacokinetic and the immunogenicity profile of Zolbetuximab. Exploratory outcomes include TTP, PFS2, DCR, biomarkers and HRU.

One interim analysis and one final analysis are planned for OS, while only one analysis is planned for PFS as the final analysis. The OS interim analysis will occur at the same time of the final PFS analysis (after pre-specified number of PFS events) and final OS analysis will be performed after the pre-specified number of OS events are observed. Refer to Section 6.9 for details on interim analysis.

Details of the study flow chart, dosing schedule, schedule of assessments are available in the protocol Section V.

2.3 Randomization

Subject randomization will be performed via IRT and treatment is assigned in a 1:1 ratio to:

- Arm A (Zolbetuximab in combination with CAPOX chemotherapy)
- Arm B (placebo in combination with CAPOX chemotherapy)

Prior to the initiation of the study treatment, the unblinded pharmacist/designee will be notified by the IRT system about the randomly assigned treatment. The unblinded pharmacist/designee will dispense the treatment according to the IRT system's assignment. Specific procedures for randomization through the IRT are contained in the IRT manual.

Randomization of subjects will use blocked randomization and be stratified by the following factors:

- Region (Asia vs. Non-Asia)
- Number of Organs with Metastatic Sites (0 to 2 vs. ≥ 3)
- Prior Gastrectomy (Yes vs. No)

3 SAMPLE SIZE

Approximately 500 subjects will be randomized in a 1:1 ratio to receive Zolbetuximab in combination with CAPOX chemotherapy (Arm A) or placebo in combination with CAPOX chemotherapy (Arm B). The planned 300 PFS events during the study will provide 93.4% power to detect a difference in PFS between Arm A (Zolbetuximab + CAPOX) with the assumption of 9 months median PFS time and Arm B (placebo + CAPOX) with the

assumption of 6 months median PFS time (hazard ratio = 0.67) at the overall 1-sided 0.025 significance level. Similarly, the planned 386 OS events during the study will provide 80% power to detect a difference in OS between Arm A (Zolbetuximab + CAPOX) with the assumption of 14.7 months median OS time and Arm B (placebo + CAPOX) with the assumption of 11 months median OS time (hazard ratio = 0.75) at the overall 1-sided 0.025 significance level.

4 ANALYSIS SETS

In accordance with International Conference on Harmonization (ICH) recommendations in guidelines E3 and E9, the following analysis sets will be used for the analyses.

The determination of whether subjects are included or excluded from the safety and efficacy analysis sets will be made prior to database hard-lock and unblinding. Inclusions and exclusions from the pharmacokinetics analysis set (PKAS) may occur after unblinding.

4.1 Full Analysis Set

The Full Analysis Set (FAS) will consist of all subjects who are randomized to 1 of the treatment arms. Subjects would be analyzed according to the treatment arm to which they were randomized to. The FAS will be used for summaries of demographic and baseline characteristics and all efficacy analyses. FAS in this study is identical to intent-to-treat (ITT) set (only the name “FAS” will be used).

4.2 Safety Analysis Set

The safety analysis set (SAF) will consist of all subjects who received at least 1 dose of any study drug (Zolbetuximab/placebo/CAPOX). Subjects would be analyzed according to the actual treatment they received. The SAF will be used for summaries of demographic and baseline characteristics and all safety and tolerability-related variables. In case that SAF and FAS are identical, summaries of demographic and baseline characteristics will not be repeated on SAF.

4.3 Pharmacokinetics Analysis Set (PKAS)

The PKAS consists of the subset of the SAF for which at least one concentration data is available. Additional subjects may be excluded from the PKAS at the discretion of the pharmacokineticist. The PKAS will be used for all summaries of the pharmacokinetic data.

5 ANALYSIS ENDPOINTS

5.1 Primary Efficacy Endpoint

The primary endpoint is PFS, which is defined as the time from the date of randomization until the date of radiological PD (per Response Evaluation Criteria In Solid Tumors [RECIST] 1.1 by independent review committee [IRC]) or death from any cause, whichever is earliest.

5.2 Secondary Efficacy Endpoints

- OS, defined as the time from the date of randomization until the date of death from any cause.
- Time to confirmed deterioration (TTCD) in PF, OG25-PA and GHS/QoL, as collected via EORTC QLQ-C30 and QLQ-OG25 plus STO22 Belching subscale, defined as the time from the date of randomization until the date of first clinically meaningful deterioration, which is also observed at the next consecutive scheduled visit, or followed by drop-out resulting in missing data or death.
- ORR, defined as the proportion of subjects who have a best overall response (BOR) of complete response (CR) or partial response (PR) as assessed by IRC per RECIST 1.1.
- DOR, defined as the time from the date of the first response (CR/PR) until the date of PD as assessed by IRC per RECIST 1.1 or date of death from any cause, whichever is earliest.
- Health-related Quality of Life questionnaires:
    - EORTC QLQ-C30 questionnaire - a 30-item cancer-specific instrument consisting of 5 functional scales (physical, role, emotional, social and cognitive), 9 symptom scales/items and a global health status scale;
    - EORTC QLQ-OG25 questionnaire - a 25-item instrument that evaluates gastric and GEJ cancer-specific symptoms. This module consists of 6 scales: dysphagia (3 items), eating restrictions (4 items), reflux (2 items), odynophagia (2 items), pain and discomfort (2 items) and anxiety (2 items), as well as 10 single items: eating in front of others, dry mouth, trouble with taste, body image, trouble swallowing saliva, choked when swallowing, trouble with coughing, trouble talking, weight loss and hair loss. In addition a belching and a bile or acid coming in your mouth question from the STO-22 follows the OG-25 questionnaire.
    - Global Pain (GP) questionnaire: a single assessment of overall pain.
    - EQ-5D-5L questionnaire: a standardized 6-item instrument that cover 5 main domains (mobility, self-care, usual activities, pain/discomfort and anxiety/depression) and a general visual analog scale (VAS) for health status.

5.3 Exploratory Efficacy Endpoints

The exploratory endpoints are:

- TTP, defined as the time from the date of randomization until the date of PD as assessed by IRC per RECIST 1.1
- PFS2, defined as the time from the date of randomization until the date of PD (per subject's local physician) following subsequent anti-cancer therapy, death from any cause, or start of any other anti-cancer therapy, whichever is earliest.
- DCR, defined as the proportion of subjects who have a BOR of SD, CR or PR as assessed by IRC per RECIST 1.1.
- Potential genomic and/or other biomarkers that may correlate with treatment outcome to Zolbetuximab and CAPOX.

- Health Resource Utilization (HRU) questionnaire: includes number of ER visits, number and duration of hospital stays, and number of doctor office visits that occur outside of the clinical trial.

5.4 Safety Endpoints

Safety and tolerability endpoints include AEs, laboratory test results, vital signs, electrocardiograms (ECGs) and Eastern Cooperative Oncology Group (ECOG) performance status.

5.4.1 AE

AE will be assessed by evaluation of the following variables:

- Treatment-emergent adverse events (TEAEs; frequency, severity, seriousness, and relationship to study drug)
  - TEAE is defined as an adverse event observed after starting administration of the study drug through 30 days after the last dose of study drug.
  - If the adverse event occurs on Day 1 and the onset check box is marked “Onset after first dose of study drug” or the onset check box is left blank, then the adverse event will be considered treatment emergent.
  - If the adverse event occurs on Day 1 and the onset check box is marked “Onset before first dose of study drug”, then the adverse event will not be considered treatment emergent.
  - If a subject experiences an event both during the pre-investigational period and during the investigational period, the event will be considered as TEAE only if it is reported with a new start date (i.e., as a new AE).
  - Any AEs with onset dates completely missing will be considered TEAEs in summaries. AEs with partially missing onset dates will be assumed TEAEs unless the available portion of the date indicates that the onset was strictly before start of study medication.
  - A drug-related TEAE is defined as any TEAE with possible relationship to study treatment as assessed by the investigator or with missing assessment of the causal relationship.
- Serious adverse events (SAEs) include adverse events that are flagged as serious by the investigator on eCRF or for which the SAE flag by the investigator on CRF is missing, or upgraded by the Sponsor based on review of the Sponsor’s list of Always Serious terms or Important Medical Events.

5.4.2 Clinical Laboratory Variables

Refer to protocol section 5.4.3 for a table of the laboratory tests that will be performed during the conduct of the study. Refer to the Protocol Schedule of Assessments for evaluation schedule.

5.4.3 Vital Signs

Vital signs will include systolic and diastolic blood pressure (mmHg), radial pulse (beats/min) and body temperature. Serial vital signs will be collected during Zolbetuximab dosing visits.

5.4.4 12-lead electrocardiogram (ECG)

A single 12-lead ECG will be performed at the time points outlined in the Protocol Schedule of Assessments. ECGs will be assessed locally.

5.4.5 ECOG performance score

ECOG performance scores will be collected.

5.4.6 Physical examination

Targeted (symptom driven) physical exams should be conducted every 3 weeks on day 1 of each cycle. If clinically significant worsening of findings from baseline is noted at any study visit, the changes will be documented as AEs on the AE eCRF.

5.5 Other Endpoints

5.5.1 Pharmacokinetic Endpoints

PK of Zolbetuximab as measured by C_{trough}.

5.5.2 Immunogenicity

Immunogenicity of Zolbetuximab as measured by the frequency of anti-drug antibody (ADA) positive subjects.

5.5.3 Biomarkers Endpoints

Potential genomic and/or other exploratory biomarkers that may be related to treatment outcome of
Zolbetuximab.

6 STATISTICAL METHODOLOGY

6.1 General Considerations

Continuous data will be summarized descriptively including the number of subjects, mean, standard deviation, median, minimum and maximum. Categorical data will be summarized by frequencies and percentages. Percentages by categories will be based on the number of subjects with no missing data, i.e. the percentages for the non-missing categories will add up to 100%. All non-coded free-text variables will be displayed in data listings only.

Summaries based on FAS (e.g. disposition, baseline characteristics and efficacy endpoints) will be presented by randomized treatment. Safety summaries based on SAF and summaries based on PKAS will be presented by actual treatment received.

All statistical comparisons will be made using one-sided test at the α=0.025 significance level unless specifically stated otherwise. All null hypotheses will be: Arm A is not better than

Arm B, all alternative hypotheses will be: Arm A is better than Arm B, unless specifically stated otherwise.

All data summarization and analyses will be performed using SAS® Version 9.3 or higher on LINUX. Specifications for table, figures, and data listing formats can be found in the TLF specifications document for this study.

Study day for safety assessments (e.g. laboratory assessment, onset of adverse events, vital signs, etc.) will be calculated in reference to the first dose date. For assessments conducted before the first dose, study day will be calculated as (assessment date – first dose date). For assessments conducted on or after the first dose, study day will be calculated as (assessment date – first dose date + 1). Study day for efficacy events (progression, death, tumor responses CR/PR) will be calculated in reference to the randomization date (event/assessment date – randomization date + 1).

For efficacy evaluation (except for PRO analysis), baseline is defined as the last available measurement before randomization. For safety evaluation and PRO analysis, baseline is defined as the last available measurement before the first dose. Unless otherwise specified, all summaries will be presented by treatment arm.

Study drug is defined as any one of the 3 components (Zolbetuximab, capecitabine and oxaliplatin) for Arm A and any one of the 3 components (placebo, capecitabine and oxaliplatin) for Arm B. Date of first dose of study drug is the date of start of infusion of the first component of study drug administered, or oral dosing of Capecitabine, whichever is earlier. Date of last dose of study drug is the date of stop of infusion of the last component of study drug administered or oral dosing of Capecitabine, whichever is later.

All by-visit summaries will use CRF visit (e.g., Cycle 3 Day 15) as described in Section 6.10.1.

6.2 Study Population

In general, data such as patient disposition, demographics and baseline characteristics will be summarized for FAS and SAF by treatment arm and overall, unless specifically stated otherwise. In the event when FAS is identical to SAF (i.e., no one received the wrong study drug), then these data summaries will not repeated for SAF.

6.2.1 Disposition of Subjects

The following summaries will be presented. A table may include one or more of the summaries.

- Number and percentage of subjects with informed consent, discontinued before randomization (screening failures), randomized (overall only);
- Number and percentage of randomized subjects in each analysis set, by treatment group and overall;
- Number and percentage of subjects completed/discontinued CAPOX regimen, by primary reason for treatment discontinuation;

- Number and percentage of subjects discontinued Zolbetuximab/placebo, by primary reason for treatment discontinuation;
- Number and percentage of subjects completed/not completed the 30-day post- CAPOX safety follow-up visit, by primary reason for not completing the visit;
- Number and percentage of subjects completed/not completed the 30-day post-Zolbetuximab/placebo safety follow-up visit, by primary reason for not completing the visit;
- Number and percentage of subjects completed/not completed the 90-day post- CAPOX follow-up visit, by primary reason for not completing the visit;
- Number and percentage of subjects completed/not completed the 90-day post-Zolbetuximab/placebo follow-up visit, by primary reason for not completing the visit;
- Number and percentage of subjects completed/not completed post-treatment follow-up, by primary reason for completing or not completing post-treatment follow-up;
- Number and percentage of subjects completed or discontinued survival follow-up period, by primary reason for survival follow-up discontinuation;

6.2.2 Protocol Deviations

The number and percentage of subjects with the following major protocol deviation criteria will be summarized for each criterion and overall, by treatment group and overall as well as by investigative site, for FAS and SAF. Subjects deviating from a criterion more than once will be counted once for the corresponding criterion. Any subjects who have more than one protocol deviation will be counted once in the overall summary.

The unique identifiers for major protocol deviation will be as follows:

PD1 - Entered into the study even though they did not satisfy entry criteria,

PD2 - Developed withdrawal criteria during the study and was not withdrawn,

PD3 - Received wrong treatment or incorrect dose,

PD4 - Received excluded concomitant treatment.

6.2.3 Demographic and Other Baseline Characteristics

Demographic variables (sex, age, age groups defined in the subgroup section, race, ethnicity, country), height, weight, BMI, BSA, tobacco history, baseline ECOG status, and the three stratification factors (listed in Section 2.3) will be summarized for FAS and SAF.

Primary diagnosis and tobacco history will be summarized for FAS.

Medical history is coded in MedDRA, and will be summarized by System Organ Class (SOC) and Preferred Term (PT) by treatment arm for FAS and SAF.

6.2.4 Transfusions

The blood product, duration, number of units will be summarized for FAS.

6.2.5 Previous and Concomitant Medications

Previous and concomitant medications will be summarized in separate tables by therapeutic subgroup (ATC 2nd level) and chemical subgroup (ATC 4th level) and preferred WHO name by treatment group for the FAS. Subjects taking the same medication multiple times will be counted once per medication and investigational period. A medication that can be classified into several chemical and/or therapeutic subgroups is presented in all chemical and therapeutic subgroups.

Previous medications are defined as medications that patients started prior to first administration of study medication. Concomitant medications are defined as any medications that patients took after the first dose of study medication (inclusive) and before the date of last dose of study medication (inclusive). Medications that started prior to first administration of study drug and continued while study drug was given will be counted as both previous and concomitant medications.

6.2.6 Non-medication Therapies

Reason for use will be summarized for FAS.

6.2.7 New Anti-Cancer Therapies

Subsequent anti-cancer therapies will be summarized by treatment arm by drug class for FAS.

6.2.8 Prior Radiation Therapy

The following variables will be summarized for FAS: whether subject received prior radiation therapy, area, duration, and reason of radiation therapy.

6.2.9 Prior Procedures for Primary Cancer

Frequency tabulations of subjects with surgery or procedures for the treatment of the primary cancer will be presented by treatment group and overall for FAS.

6.2.10 Prior Cancer Chemotherapy

The following variables will be summarized for FAS: whether subject received any chemotherapy medication, name and duration of chemotherapy, and reason for discontinuing medication.

6.3 Study Drug Exposure and Compliance

Duration and compliance of study drug will be summarized for SAF by treatment group and overall.

The following variables will be derived and summarized:

- Duration of Zolbetuximab or placebo and Oxaliplatin, defined as (date of last infusion – date of first infusion + 1). For those who did not receive any dose, duration will be 0.
- Duration of Capecitabine, defined as (date of last dose – date of first dose + 1). For those who did not receive any dose, duration will be 0.

- Duration of period for which all components are administered and duration of period for Any Component Administered
- Proportion of subjects who completed the 8 cycles of CAPOX treatment.
- Number of infusions administered, number of infusions entirely administered and number of infusion not entirely administered, number of infusions with dose adjustment, Number of infusions with dose adjustment due to AE, number of infusions with delay, Number of infusions with delay due to AE, number of infusions with interruption, Number of infusions with interruption due to AE, number of infusions prematurely discontinued for Zolbetuximab and Oxaliplatin and Number of infusions prematurely discontinued for Zolbetuximab and Oxaliplatin due to AE.
  For the interruptions that involve overnight infusion, only one infusion will be counted
- Number of cycle administered for oral dosing of capecitabine, number of cycle with dose adjustment, Number of cycle with dose adjustment due to AE, number of cycle with delay, number of cycle with delay due to AE, number of cycle with interruption, number of cycle with interruption due to AE, number of cycle prematurely discontinued for Capecitabine, number of cycle prematurely discontinued for Capecitabine due to AE.
- Average infusion time, calculated as (stop time – start time), for Zolbetuximab, placebo and Oxaliplatin. Interruption time is included in the infusion time if the infusion were finished within one day; for interruptions that goes overnight, interruption time will not be included.
- Cumulative actual dose of Zolbetuximab/placebo and each component of CAPOX.
- Average amount of dose per infusion for each component administered via infusion.
- Average amount of dose per planned dosing day for Capecitabine.
- Relative dose intensity (RDI), defined as (Cumulative actual dose/Planned cumulative dose). Where Planned Cummulative Dose is defined as Protocol specified planned cumulative dose.
- RDI category: <50, 50 to 80 inclusive, >80, unknown
- Number and percentages of subjects with the following cumulative categories of study drug (each component) duration will be summarized: ≥ 1d, > 6w, > 12w, > 24w, > 36w, > 48w, > 72w

6.4 Analysis of Efficacy

To address multiplicity, a gatekeeping testing strategy will be used for PFS (primary efficacy endpoint) and OS (key secondary endpoint). PFS will be tested once at 1-sided significance level of 0.025. Only if PFS is significant, hypothesis testing for OS interim and OS final analyses will be performed. An O'Brien-Fleming type alpha-spending function [Lan & DeMets, 1983] will be utilized to control the overall 1-sided 0.025 significance level for the OS interim and final analyses. Other secondary endpoints' testing will not be multiplicity adjusted.

For imaging assessments of tumor, date of assessment for each timepoint is defined as the date of last scan (if there are multiple scans over several days) for that timepoint, not the date when the overall timepoint response is recorded by radiologist in the system, with the

exception of PD date. PD date is defined in Table 1. Date of time point assessment of tumor as determined by IRC/investigators will be used for analysis.

6.4.1 Analysis of Primary Efficacy Endpoint (PFS)

6.4.1.1 Primary Analysis for Primary Efficacy Endpoint

The primary analysis of PFS will use radiological assessment of PD by the IRC and be performed for FAS. The hypothesis to be tested is:

- H_0: PFS of Arm A is not prolonged compared to that of Arm B
- H_a: PFS of Arm A is prolonged compared to that of Arm B

Comparison of Arm A and Arm B will be tested at 1-sided significance level of 0.025.

The distribution and median of PFS will be estimated for each treatment arm using Kaplan-Meier methodology. PFS rates at 6m, 12m, and 18m will be presented by Kaplan-Meier method too. In addition, numbers of subjects with PFS events and censored and 95% CI for median PFS and PFS rates will be presented.

Hypotheses testing between Arm A and Arm B will be performed using log-rank test stratified by:

- Region (Asia vs. Non-Asia)
- Number of Organs with Metastatic Sites (0 to 2 vs. ≥ 3)
- Prior Gastrectomy (Yes vs. No)

In addition, stratified Cox proportional hazards model will be used to estimate the hazard ratio and the corresponding 95% confidence interval. Table 1 defines PFS endpoint for the primary analysis. Evaluable radiological assessments include all assessments except those assessed by IRC as NE (not evaluable).

Table 1 PFS Primary Analysis Definition (based on IRC radiological assessments only)

Situation	Date of Event or Censoring	Outcome
No baseline imaging assessments	Date of randomization	Censored
No evaluable post-baseline imaging assessments, no death	Date of randomization	Censored
Subject did not receive new anti-cancer therapy (ACT):		
Radiological PD documented per RECIST v1.1	Date of first radiological PD (defined as earliest of date of scan showing new lesion if PD is based on new lesion or date of last scan of target lesions if PD is based on increase in sum of diameters (SOD) of target lesions)	Event
No radiological PD, but death recorded on eCRF	Date of death	Event
Neither radiological PD nor death	Date of last radiological assessment	Censored

Situation	Date of Event or Censoring	Outcome
Subject received new anti-cancer therapy (ACT)*:		
Radiological PD per RECIST v1.1 documented only after start of new ACT	Date of last radiological assessment before start of new ACT	Censored
Radiological PD documented per RECIST v1.1 before start of new ACT	Date of first radiological PD	Event
No radiological PD nor death	Date of last radiological assessment before start of new ACT	Censored
Missed >=2 scheduled radiological assessments:		
If radiological PD or death occurs after missing 2 or more scheduled radiological assessments**	Date of last radiological assessment	Censored

Note: PFS = date of event or censoring – date of randomization + 1. NE will be treated as missing in the derivation described in this table *New ACT includes new anti-cancer surgery, radiotherapy, chemo, immunotherapy (other than Zolbetuximab and CAPOX components) and on study tumor directed procedures after randomization. If a subject in Arm B switches from placebo to Zolbetuximab, it is also considered start of new ACT. **If the first radiological assessment after subject missed >=2 imaging assessments is SD or better and it's confirmed that subject did not take any other ACT during the missing period, the following imaging assessments will be used rather than censored.

6.4.1.2 Sensitivity Analysis 1 for Primary Efficacy Endpoint

The primary analysis will be repeated using radiologic PD assessments by local investigators only. In addition, a summary of discordance between IRC's and local investigator's PD assessments will be presented.

6.4.1.3 Sensitivity Analysis 2 for Primary Efficacy Endpoint

This analysis will treat likely informative censoring as PFS events. The primary analysis will be repeated with the following cases treated as PFS events rather than censored (whichever earliest):

- If subject dropped out of imaging follow-up without documented PD by IRC and there was investigator-reported radiological progression at the last imaging assessment, then the next scheduled date of imaging (i.e., date of last imaging assessment + 9w or 12w) will be treated as date of PFS event (even though the next scheduled imaging did not take place).
- If subject dropped out of imaging follow-up without documented PD by IRC and there was investigator-reported clinical progression around or after the time of last imaging assessment, then the next scheduled date of imaging (i.e., date of last imaging assessment + 9w or 12w) will be treated as date of PFS event (even though the next scheduled imaging did not take place).

- If subject dropped out of imaging follow-up without documented PD by IRC and there was ECOG performance status worsening from baseline (from 0-1 to >=2) around or after the time of last imaging assessment, then the next scheduled date of imaging will be treated as date of PFS event.
- Start date of new ACT will be treated as date of PFS event, when there is no prior documented PD by IRC.
- If subject missed >=2 scheduled imaging assessments without prior documented PD by IRC and there was investigator-reported clinical progression around or after the time of last imaging assessment prior to missing, then the next scheduled date of imaging after the last non-missing assessment will be treated as date of PFS event (even though the next scheduled imaging did not take place).
- If subject missed >=2 scheduled imaging assessments without prior documented PD by IRC and there was ECOG performance status worsening from baseline (from 0-1 to >=2) around or after the time of last imaging assessment prior to missing, then the next scheduled date of imaging after the last non-missing assessment will be treated as date of PFS event.

6.4.1.4 Sensitivity Analysis 3 for Primary Efficacy Endpoint

The primary analysis will be repeated where death after new ACT will be censored at date of last radiological assessment before start of new ACT.

6.4.1.5 ther Analysis of Primary Efficacy Endpoint

Imaging assessment interval (weeks) will be summarized for each treatment arm for the two periods: <=54 weeks and >54 weeks. If imbalance of imaging assessment interval is observed between the two arms, additional sensitivity analysis may be performed using protocol-planned dates of imaging assessments instead of actual dates of assessments. In addition, duration of imaging follow-up, defined as (date of the last on-study imaging assessment – date of randomization + 1), will be summarized for each treatment arm.

6.4.2 Analysis of Secondary Efficacy Endpoints

6.4.2.1 Overall Survival

A key secondary endpoint OS is defined as the time from date of randomization to the documented date of death from any cause. All deaths will be included, regardless of whether death occurred while the subject is still taking study drug or after the subject discontinue study drug. OS analysis will be performed for FAS.

Table 2 OS Definition

Situation	Date of Event or Censoring	Outcome
Death before analysis cutoff date	Date of death	Event
Last known alive date is before cutoff date	Last known alive date	Censored
Death after analysis cutoff date	Analysis cutoff date	Censored
Last known alive date is after cutoff date	Analysis cutoff date	Censored

OS = Date of Event or Censor – Date of Randomization +1

To maintain the overall Type I error rate at the 0.025 significance level, the hypothesis testing for OS interim and OS final analyses will be performed only if the null hypothesis in PFS primary analysis is rejected at the overall 1-sided 0.025 significance level.

The formal OS interim analysis is planned when the final PFS analysis occurs with the pre-specified number of PFS events. A group sequential design using the O'Brien-Fleming type alpha-spending function [Lan & DeMets, 1983] will be utilized to control the overall 1-sided 0.025 significance level (East®) for the OS analysis. The independent data monitoring committee (IDMC) may recommend terminating the trial for favorable results at the formal OS interim analysis. In the case of favorable results, the 1-sided significance level for superiority is 0.0074 for the interim OS analysis and 0.0228 for the final OS analysis. These alpha boundaries are based on an information factor of 70% and are subject to adjustment if observed information factor deviates from 70%. If the 1-sided P-value of the interim analysis is less than 0.0074, the IDMC may recommend terminating the trial for success without the need to conduct the final OS analysis. If the study is not stopped after the interim analysis, the final OS analysis will occur after 100% of the planned number of deaths have been observed.

The distribution and median of OS, as well as OS rates at 12m, 18m, 24m, 30m and 36m will be estimated for each treatment arm using Kaplan-Meier methodology. 95% CI for median OS and milestone OS rates will be presented. Arm A and Arm B will be compared using the log-rank test stratified by the same stratification factors used for PFS analysis. The hypothesis to be tested is:

- H_0: OS of Arm A is not prolonged compared to Arm B
- H_a: OS of Arm A is prolonged compared to Arm B

In addition, stratified Cox proportional hazard model will be used to estimate the hazard ratio and the corresponding 95% CI.

The median of time on study using reverse Kaplan-Meier approach and range (min and max) of time on study will be provided. Subjects alive up to date cut-off date will be considered as events and death of subjects on or prior to data cut-off date will be censored in the reverse Kaplan-Meier approach for estimation of median of time on study.

6.4.2.2 Health-Related Quality of Life

6.4.2.2.1 Scoring of HRQoL Questionnaires

For EORTC QLQ-C30, scores for 5 functional scales (physical, role, emotional, social and cognitive), 9 symptom scales/items, and global health status scale will be calculated according to the EORTC scoring manual in Appendix 9.2 (Fayers, et al, 2001). These scores will be standardized to a 0-100 scale. A high score for a functional scale and the global health status represents a healthy level of functioning (high QoL), while a high score for a symptom scale or item represents a severe level of symptoms (low QoL).

In addition, the QLQ-C30 Summary Score (Giesinger et al, 2016) will be calculated using 27 out of 30 items as follows (excluding 3 items on financial impact and global health status):

QLQ-C30 Summary Score = (Physical Functioning+ Role Functioning+ Social Functioning+ Emotional Functioning+ Cognitive Functioning+ 100-Fatigue+ 100-Pain+ 100-Nausea_Vomiting+ 100-Dyspnoea+ 100-Sleeping Disturbances+ 100-Appetite Loss+ 100-Constipation+ 100-Diarrhoea)/13.

The QLQ-C30 Summary Score should only be calculated if all of the required 13 scale/item scores are available (using scale scores based on the completed items, provided that at least 50% of the items in that scale have been completed). A high Summary Score represents high QoL.

EORTC QLQ-OG25 questionnaire consists of 6 scales: dysphagia (items 1-3), eating restrictions (items 4-7), reflux (items 8-9), odynophagia (items 10-11), pain and discomfort (items 12-13) and anxiety (items 14-15), as well as 10 single items: eating in front of others (item 16), dry mouth (item 17), trouble with taste (item 18), body image (item 19), trouble swallowing saliva (item 20), choked when swallowing (item 21), trouble with coughing (item 22), trouble talking (item 23), weight loss (item 24) and hair loss (item 25). Also, we added two items from EORTC QLQ-STO22 related to belching (item 26-27)

EORTC QLQ-OG25 7 scales plus 10 single items will be scored in the same way as QLQ-C30 symptoms scales/items. The scores will also be transformed to a 0-100 scale. Higher score means severer level of symptoms.

In addition, the EORTC QLQ-C30 and OG25 individual symptom scores and the Belching questionnaire will be categorized as the following:

a. For questions which have a 4 point scales with 1 being the best, and 4 being the worst, categorized as 1=”None”, 2=”Slight”, 3=”Moderate”, 4= “Severe”.

The EQ-5D-5L scores (on 1-5 scale) will be categorized as: 1=”None”, 2=”Slight”, 3=” Moderate”, 4=”Severe”, 5=”Extreme”.

6.4.2.2.2 Analysis of HRQoL Questionnaires

Analyses will be by treatment arm and based on FAS. All the PRO assessments while on study drug, as well as end of treatment will be considered.

6.4.2.2.2.1 Patient disposition

The subject disposition by treatment group for all PRO assessment time-points (e.g. analysis visits) will be provided:

- The number of subjects with PRO assessment expected
- The number and % of subjects with PRO assessment not expected due to progression
- The number and % of subjects with PRO assessment not expected due to death
- The number and % of subjects with PRO assessment not expected due to other reasons

The subject disposition by treatment group per analysis visit will also be provided graphically by means of a stacked bar chart.

6.4.2.2.2.2 Completion rate

Instrument completion rate at each analysis visit will be reported for each instrument:

- Completion rate (i.e. unadjusted) at each analysis visit will be calculated as the number of subjects meeting the minimum requirements for scoring at least one domain of the instrument divided by the number of subjects in the FAS population.
- Compliance rate (i.e. adjusted) at each analysis visit will be calculated among subjects who are expected to have PRO assessments. The following will be provided:
    - The number and % of subjects with all questions completed
    - The number and % of subjects meeting at least the minimum requirements for scoring of the instrument; these requirements are as follows:
    - EQ-5D-5L: 1) the utility index 2) the EQ-VAS is calculated
    - EORTC QLQ-C30: at least one subscale can be calculated
    - EORTC QLQ-OG25 plus STO22 Belching: at least one subscale can be calculated
    - The number and % of subjects with at least one question completed

The completion (unadjusted) and compliance (adjusted) rates by treatment group at each analysis visit will also be provided graphically by means of a line graph.

6.4.2.2.2.3 Time to deterioration for GHS/QoL, PF and OG25-Pain

Definition

Time to clinically meaningful symptom worsening or HRQoL deterioration (PRO deterioration) will be analysed for each scale of the PRO instruments collected in this study separately. For convenience, a generic term “time to clinically meaningful deterioration” will be used both for symptom worsening and HRQoL deterioration, with an understanding of a specific meaning depending on the scale or subscale analysed. The following definition will be considered for the secondary endpoints of time to deterioration in GHS/QoL, PF and OG25-Pain.

Time to confirmed deterioration (TTCD) will be defined as the duration of time from the date of randomization to the date of the first deterioration in PRO scores of at least one threshold unit as compared to the baseline score if the deterioration of at least one threshold unit as compared to the baseline score is also observed at the next consecutive scheduled visit (e.g., after the first deterioration is observed) or if the patient dropped out after deterioration or died, resulting in missing data.

For those patients who experienced first confirmed clinically meaningful deterioration, TTCD will be computed as follows and then converted to months:

TTCD = Date of assessment when first confirmed clinically meaningful deterioration was observed – Date of randomization + 1

Patients with a non-missing baseline assessment will be censored at the last available PRO assessment. Patients with no baseline PRO assessment, or without post-baseline PRO questionnaire, or whose baseline scores do not allow for further deterioration will be

censored at the randomization. Death or progression will not be considered deterioration events.

Given the absence of reliable and well-accepted thresholds for within-patient clinically meaningful change, the clinically meaningful threshold denoting a deterioration will be defined based on anchor-based analyses performed on the same trial data. These analyses are described in a separate pre-specified analysis plan for an exit survey dated 09-Mar-2022. Sensitivity analyses with the next higher threshold value that will provide a different classification for deterioration for each scale will also be performed. The reason of this approach is that EORTC values are discrete in nature due to the transformation of raw scores to 0-100, therefore not all values are possible within this range and certain threshold values will result in the same categorization for patients, e.g., for the two-item OG25-Pain domain, the possible values are 0, 16.7, 33.3, 50, 66.7, 83.3 and 100, therefore applying a threshold of 10 and 14 will result in the same classification of patients into deterioration and no deterioration. If the resulting threshold value from the anchor-based analysis is for example 10, then the sensitivity threshold will be 17.

The above analyses will be also repeated by employing the sensitivity/secondary thresholds.

Alternative definitions for the time to deterioration, such as time to first deterioration, or time to definitive deterioration, will also be described in a separate SAP for PRO-related analyses. Briefly, time to first deterioration (TTFD) will be defined as the duration of time from the date of randomization to the date of the first deterioration in PRO scores of at least one threshold point as compared to the baseline score. Time to first definitive deterioration (TTDD) will be defined as the duration of time from the date of randomization to the date of the first deterioration in PRO scores of at least one threshold unit as compared to the baseline score, which is:

- also observed at all time points thereafter (e.g., after the first deterioration is observed) or
- followed by drop-out, resulting in monotone missing data.

In addition, the TTCD, TTFD and TTDD definition will be repeated where death (due to any cause) will be counted as an event if the patient does not experience PRO deterioration prior to death and where death occurs within 2 scheduled assessments (e.g., 42 days; this corresponds to the maximal time interval between 2 consecutive scheduled visits for PRO assessment) after the last available PRO assessment; progression will not be considered deterioration event.

Details of the sensitivity/secondary analysis will be included in the PRO related SAP.

Analysis

If the null hypothesis on the OS is rejected at the overall 1-sided 0.025 significance level, then the TTCD will be tested using the gatekeeping procedure with the following order:

1. Non-inferiority testing for TTCD in PF at 0.025 significance level
2. Non-inferiority testing for TTCD in OG25-Pain at 0.025 significance level

3. Non-inferiority testing for TTCD in GHS/QoL at 0.025 significance level
4. Superiority testing for TTCD in PF at 0.025 significance level
5. Superiority testing for TTCD in OG25-PA at 0.025 significance level
6. Superiority testing for TTCD in GHS/QoL at 0.025 significance level

A stratified Cox proportional hazard model will be used to estimate the hazard ratio and the corresponding 95% CI for the three time-to-event PRO endpoints.

To determine whether Zolbetuximab (IMAB362) in combination with CAPOX chemotherapy is inferior to Placebo in combination with CAPOX chemotherapy, a non-inferiority margin of 1.33 in terms of hazard ratio with respect to experiencing a definitive deterioration as defined in the previous section will be used.

The primary hypothesis (H1) is defined as:

H0: HR $\geq$ 1.33 vs. H1: HR < 1.33

The hypothesis of non-inferiority will be tested at a one-sided significance level of 0.025 in the FAS population using a stratified Cox Proportional Hazards model with adjusting for randomization stratification factors: Region (Asia vs Non-Asia), Number of organs with metastatic sites (0 to 2 vs $\geq$ 3), Prior gastrectomy (Yes or No). The 95% confidence interval for the hazard ratio of Zolbetuximab (IMAB362) Plus CAPOX to Placebo Plus CAPOX will be estimated. If the upper limit of the 95% confidence interval for the estimated HR for the stratified Cox model is below the non-inferiority margin of 1.33, then non-inferiority of the TTCD endpoints in patients treated with Zolbetuximab (IMAB362) Plus CAPOX compared to that of patients treated with Placebo Plus CAPOX will be declared.

If the non-inferiority hypothesis are met for all 3 HEOR endpoints, the hypothesis of superiority for these endpoints will be tested.

The hypothesis (H1) is defined as: H0: HR $\geq$ 1 vs. H1: HR < 1

Kaplan-Meier curves will be used to estimate the distribution of TTCD. The 50th percentile of Kaplan-Meier estimates will be used to estimate the median duration of TTCD, respectively. A two-sided 95% confidence interval will be provided for these estimates. Median TTCD will be compared using stratified log rank test adjusting for randomization stratification factors: Region (Asia vs Non-Asia), Number of organs with metastatic sites (0 to 2 vs $\geq$ 3), Prior gastrectomy (Yes or No). A Kaplan-Meier plot by treatment group will be presented.

Additionally, the benefit of Zolbetuximab (IMAB362) Plus CAPOX compared to Placebo Plus CAPOX will be evaluated by a single hazard ratio (HR) (Zolbetuximab vs Placebo) with its 95% confidence interval based on a stratified Cox regression model with the same strata as above. The proportional hazards assumption will be tested by examining plots of complementary log(-log(survival)) versus log(time). In case departures from the assumption are observed, only the KM and the quartiles of the survival distribution will be presented.

Kaplan-Meier analysis will be performed using PROC LIFETEST (SAS procedure). Cox proportional hazard regression model will be performed using PROC PHREG (SAS procedure).

6.4.2.2.2.4 Descriptive analysis

For QLQ-C30's 5 functional scales, global health status, Summary Score, and EQ-5D-5L VAS

Scores (on 0-100 range) as well as change from baseline will be summarized by visit for each visit where at least 10 subjects are evaluable for any treatment arm. Summary statistics will not be provided for the treatment arm for the visit if that arm has less than 10 subjects at that visit. Plot of mean change from baseline over time will be presented by treatment arm for each score. The highest and lowest change from baseline of each subject (highest value post baseline –baseline value and lowest value post baseline – baseline value) will be summarized by treatment arm.

For QLQ-C30 and QLQ-OG25 plus STO-22 belching subscale symptoms:

Scores (on 0-100 range) will be summarized separately in the same way as above. Frequency distribution of categorized QLQ-C30 and QLQOG25 symptom scores will be presented by treatment arm and visit. Shift from baseline category to subject's worst post-baseline category will be tabulated (except for 7 points scale questionnaire).

EQ-5D-5L 5 questions

Scores (on 1-5 range) will be analyzed using descriptive statistics including change from baseline. Categorized scores will be tabulated using shift table (shift from baseline to worst category post baseline).

Global pain

Score and change from baseline will be summarized by visit.

6.4.2.3 Objective Response Rate

The ORR is defined as the proportion of subjects with a BOR of complete response (CR) or partial response (PR) based on IRC per RECIST V1.1.

Best overall response (BOR) will be determined once all tumor response data for the subject is available. Subjects' BOR will be determined as outlined in RECIST V1.1 criteria based on IRC assessments. BOR is defined as the best response among all timepoints' overall responses excluding NE responses (CR is better than PR and PR is better than SD and SD is better than PD). If all timepoint overall responses are NE, BOR is NE. For BOR of SD, SD must be documented as present at least once and at least 8 weeks after randomization. If the first assessment of SD does not meet the minimum "8 weeks from randomization" time window, that assessment of SD will be treated as NE in analysis. Timepoint responses after start of new ACT will not be used in determining BOR.

The comparison of ORR between Arm A and Arm B will be performed using stratified Cochran-Mantel-Haenszel (CMH) test with the same stratification factor used for the PFS

analysis. In addition, ORR for each arm will be estimated and corresponding 95% CI will be constructed using Clopper-Pearson method.

In addition, percent of subjects with BOR of CR, PR or SD as well as ORR and DCR will be summarized.

Sensitivity analyses include:

- Analysis of ORR with confirmation, defined as the proportion of subjects with best overall response as confirmed CR or confirmed PR based on the RECIST v1.1 as assessed by IRC. Confirmation of CR or PR should occur at the next scheduled assessment (>= 4 weeks following the initial assessment at which CR or PR is observed). See Table 3 for rules of confirmation of CR and PR.

Table 3 Confirmation of overall response:

Overall response at current timepoint	Overall response at next timepoint	Confirmed overall response at current timepoint
CR	CR	CR
PR	CR/PR	PR
CR/PR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
PR	SD	SD
SD	CR/PR	SD
SD	SD/PD/NE	SD provided minimum criteria for SD duration met, otherwise, PD or NE.
CR	NE, followed by CR at the following assessment	CR
PR	NE, followed by CR/PR at the following assessment	PR
CR	NE	SD provided minimum criteria for SD duration met, otherwise, NE
PR	NE	SD provided minimum criteria for SD duration met, otherwise, NE
NE	NE	NE

- Analysis of ORR per investigator assessment for both confirmed and unconfirmed response.

6.4.2.4 Duration of Response

DOR is defined as the time from the date of the first CR or PR (whichever is first recorded) as assessed by IRC to the date of radiological PD as documented by IRC or death, whichever is earlier. The DOR analysis will be performed on the subset of FAS who have at least one CR or PR documented by IRC. If a subject has not progressed, the subject will be censored at the date of last evaluable radiological assessment or at the date of first CR/PR if no later

evaluable radiological assessment is available. Other censoring used for the PFS analysis (see Table 1) will apply to DOR too.

DOR = Date of Event or Censor – Date of the first CR/PR + 1.

The distribution of DOR will be estimated for each treatment arm using Kaplan-Meier methodology and compared between Arm A and Arm B using the log-rank test stratified by the same stratification factors used for the PFS analysis. In addition, stratified Cox proportional hazard model will be used to estimate the hazard ratio and the corresponding 95% CI.

As sensitivity analyses, DOR analysis will be repeated for confirmed responses. For confirmed CR/PR, date of first confirmed CR/PR is defined as the date of first assessment of CR/PR (not the later assessment confirming the previous CR/PR).

DOR analysis will also be repeated using investigator assessment, and it will include both analysis for confirmed and unconfirmed response.

6.4.3 Analysis of Exploratory Efficacy Endpoints

6.4.3.1 Time to Progression

TTP is defined as the time from the date of randomization until the date of PD per RECIST 1.1 by IRC. TTP does not include deaths as event. For deaths before the first documented PD by IRC, subjects will be censored at the time of last evaluable radiological assessment. The TTP analysis assumes that deaths without documented PD are not related to tumor progression and estimates TTP for these dead subjects (who did not have documented PD prior to death) as if the subjects had not die. See Table 4 for derivation of the TTP endpoint. Kaplan-Meier and log-rank methods will be applied to the TTP endpoint.

Table 4 TTP Definition (based on IRC radiological assessments only)

Situation	Date of Event or Censoring	Outcome
No baseline imaging assessments	Date of randomization	Censored
No evaluable post-baseline imaging assessments	Date of randomization	Censored
Subject did not receive new anti-cancer therapy (ACT):		
Radiological progression documented per RECIST v1.1	Date of radiological PD (defined as earliest of date of scan showing new lesion if PD is based on new lesion or date of last scan of target lesions if PD is based on increase in sum of diameters (SOD) of target lesions)	Event
No radiological progression, but death recorded on eCRF	Date of last radiological assessment	Censored
Neither radiological progression nor death	Date of last radiological assessment	Censored

Situation	Date of Event or Censoring	Outcome
Subject received new anti-cancer therapy (ACT) *:		
Radiological progression documented per RECIST v1.1 after new ACT	Date of last radiological assessment before start of new anti-cancer therapy	Censored
Radiological progression documented per RECIST v1.1 before new ACT	Date of radiological PD	Event
No radiological progression before new ACT but death recorded	Date of last radiological assessment before start of new anti-cancer therapy	Censored
No radiological progression nor death	Date of last radiological assessment before start of new anti-cancer therapy	Censored
Missed >=2 scheduled radiological assessments:		
If radiological progression occurs after missing 2 or more scheduled radiological assessments**	Date of last radiological assessment	Censored

NE will be treated as missing in the derivation described in this table *New ACT includes new anti-cancer surgery, radiotherapy, chemo, immunotherapy (other than Zolbetuximab and CAPOX components) and on study tumor directed procedures after randomization. If a subject in Arm B switches from placebo to Zolbetuximab, it is also considered start of new ACT. **If the first radiological assessment after subject missed >=2 imaging assessments is SD or better and it's confirmed that subject did not take any other ACT during the missing period, the following imaging assessments will be used rather than censored.

As sensitivity analyses, TTP analysis will be repeated using investigator assessment.

6.4.3.2 Progression Free Survival After Subsequent Therapy (PFS2)

PFS2 is defined as the time from the date of randomization until the date of radiological/objective PD (per subject's local physician) following subsequent (2nd line) anti-cancer therapy or death from any cause, whichever is earliest. In cases where PFS2 cannot be reliably determined, end date of subsequent (2nd line) ACT or start date of 3rd line ACT may be used as the event date. Otherwise, subjects will be censored. Subjects who are alive and for whom a PFS2 event has not been observed should be censored at the last time of time known to be alive. Here last know alive date is used as a surrogate to the last radiological assessment date as the later may not be available in the database after 2nd line ACT.

The distribution of PFS2 will be estimated for each treatment arm using Kaplan-Meier methodology and compared between Arm A and Arm B using stratified log-rank test with the same stratification factors used for the PFS analysis. In addition, stratified Cox proportional hazard model will be used to estimate the hazard ratio and the corresponding 95% CI.

Table 5 PFS2 Definition (as compared to PFS)

Situation	PFS2 event/censor	PFS2 date
Subject did not take new ACT:		
Subject died	Event	Death date
Subject did not die	Censor	Last known alive date
Subject started new ACT:		
PD (based on investigator) after new ACT or Death	Event	Date of PD after new ACT or Date of death, whichever earlier
No PD (based on investigator) after new ACT, No Death. Ended 2nd or started 3rd line ACT.	Event	End date of 2nd line ACT, or start date of 3rd line ACT, whichever earlier.
No PD (based on investigator) after new ACT, No Death. Not record of End date of 2nd line ACT or start date of 3rd line ACT.	Censor	Last known alive date

6.4.3.3 Disease Control Rate

The DCR is defined as the proportion of subjects with a BOR of SD, CR or PR based on RECIST 1.1 by IRC.

The comparison of DCR between Arm A and Arm B will be performed using CMH test with the same stratification factor used for PFS analysis. In addition, DCR for each arm will be estimated and corresponding 95% CI will be constructed.

As sensitivity analyses, DCR analysis will be repeated using investigator assessment.

6.4.3.4 Health Resource Utilization

HRU variables by visit and their sums over all visits, will be summarized by treatment arm:

- Number of ER visits
- Number of ER visits with hospitalization
- Duration of ER hospitalizations
- Number of hospital admissions
- Duration of stay in hospital
- Number of general practitioner visits
- Number of specialist visits

The above summaries will be repeated for the time period before subjects' radiological PD.

6.5 Analysis of Safety

Safety analyses will be based on SAF and by treatment arm and overall. All SAF subjects will be analyzed according to the actual treatment they received. Astellas Standard TLF templates should be followed wherever applicable.

Safety analyses will only include data collected during and after the first dose of the first study drug component given, up to 30 days after the last dose of the last study drug component.

6.5.1 Adverse Events

AEs will be coded using the Medical Dictionary for Regulatory Activities (MedDRA) and graded using NCI-CTCAE. MedDRA Version used will be presented in the title of the related TLFs.

TEAE is defined as an AE observed after starting administration of the study treatment and within 30 days after the last dose of the last administered component of study treatment. Late AE/SAE is defined as AE/SAE that is collected after 30 days post last dose of study drug. Serious TEAE summaries include both investigator-assessed and Astellas upgraded SAEs.

Separate summaries for AE/SAE related to any component of the study drug and AE/SAE related to each component of study drug will be provided for Drug-related AE/SAE summaries.

An overview summary table will include the following details:

- Number and percentage of subjects with TEAEs,
- Number and percentage of subjects with drug-related TEAEs
- Number and percentage of subjects with serious TEAEs,
- Number and percentage of subjects with serious drug-related TEAEs,
- Number and percentage of subjects with TEAEs leading to permanent discontinuation of any component of study drug and by component
- Number and percentage of subjects with drug-related TEAEs leading to permanent discontinuation of any component of study drug and by component
- Late adverse event occurred beyond 30 days from the last study treatment (all components)
- Late serious adverse event occurred beyond 30 days from the last study treatment (all components)
- Number and percentage of subjects with NCI CTCAE grade 3 or higher TEAEs
- Number and percentage of subjects with drug-related NCI CTCAE grade 3 or higher TEAEs
- Number of deaths from first dose of study drug up to 30 days after the last dose of the last administered component of study drug
- Number and percentage of subjects with TEAE leading to death
- Number and percentage of subjects with drug-related TEAE leading to death
- Number of all deaths up to analysis cutoff date

The number and percentage of subjects with TEAEs, as classified by SOC and PT will be summarized. Summaries will be provided for:

- TEAEs
- Drug-related TEAEs,
- AEs collected after 30 days post last dose of study drug,
- Serious TEAEs,
- Serious AEs collected after 30 days post last dose of study drug,
- Drug-related serious TEAEs,

- TEAEs leading to permanent discontinuation of any component of study drug and by component,
- Drug-related TEAEs leading to permanent discontinuation of any component of study drug and by component,
- TEAE Leading to Dose Interruption of any component of study drug and by component
- Drug-Related TEAE Leading to Dose Interruption of any component of study drug and by component
- TEAE Leading to Dose Reduction of any component of study drug and by component
- Drug-Related TEAE Leading to Dose Reduction of any component of study drug and by component
- TEAE Leading to Dose Rate reduction of any component of study drug and by component
- Drug-Related TEAE Leading to Dose Rate reduction of any component of study drug and by component
- TEAEs excluding serious adverse events that have a frequency of >=10% in any treatment arm, and
- TEAE with NCI CTCAE Grade 3 or higher
- Drug-related TEAE with NCI CTCAE Grade 3 or higher
- Drug-related TEAEs with treatment group difference in incidence (Arm A - Arm B) >10%
- Drug-related serious TEAEs with treatment group difference in incidence (Arm A - Arm B) >5%

The number and percentage of subjects with TEAEs and TEAEs leading to death, as classified by PT only, will be summarized by treatment group and overall.

AE summary tables will include subject counts as opposed to AE counts, except for serious TEAE and TEAE leading to death where AE counts will also be presented. If a subject experiences more than one episode of a particular AE, that subject will be counted only once for that event. If a subject has more than one AE that code to the same preferred term, the subject will be counted only once for that preferred term. Similarly, if a subject has more than one AE within a body system, the subject will be counted only once in that body system.

The number and percentage of subjects with TEAEs, as classified by SOC and PT, will also be summarized by NCI-CTCAE severity grade and by relationship to study drug. Drug-related TEAEs will be presented in a similar way by severity grade only. If an adverse event changes in severity grade or relationship, then the subject will be counted only once with the worst severity grade and highest degree of relationship. If a subject has an event more than once with missing severity grade and non-missing severity grade, then the subject will be counted as the highest non-missing grade. If a subject has an event more than once with missing relationship and non-missing relationship, then the subject will be counted as "related".

The number and percentage of subjects with TEAEs of interest [listed below], as classified by SOC and PT, will also be summarized. The list of adverse events of interest to be

summarized may change during the course of the study due to ongoing pharmacovigilance. It will be finalized before the database hard lock.

6.5.2 AE of Special interest

Adverse events of special interest (AESI) include the following:

- Nausea
- Vomiting
- Abdominal pain
- Hypersensitivity reactions
- Infusion-related reactions (IRRs)
- Anemia
- Neutropenia

The list of adverse events of special interest to be summarized may change during the course of the study due to ongoing pharmacovigilance. AESIs except for infusion related reactions (IRRs), are classified by SSQ/CMQ or SMQ and PT. IRRs are classified as follows:

- Infusion-related reactions (IRRs)
  - Infusion-related reactions (IRRs) flagged by investigators
  - Potential IRR defined as all AE that have a start date the same as a study treatment day

The number and percentage of subjects with AESI (AE of Special Interest), as classified by SOC and PT will be summarized. Summaries will be provided for:

- AESI
- Serious AESI
- AESI by NCI-CTCAE Grade
- AESI leading to permanent discontinuation
- AESI leading to dose interruption
- AESI leading to dose reduction
- AESI leading to dose rate reduction
- AESI leading to death
-

6.5.3 Clinical Laboratory Evaluation

Quantitative values evaluated by the central laboratory including hematology, biochemistry, urinalysis and coagulation will be summarized using mean, standard deviation, minimum, maximum and median by treatment group at each visit. Additionally, a within-subject change will be calculated as the post-baseline measurement minus the baseline measurement and summarized in the same way. Frequency tabulations of selected qualitative clinical laboratory variables (i.e. urinalysis) will be presented by treatment arm at each visit.

Central laboratory results will also be graded using NCI-CTCAE, where possible. Laboratory parameters that have criteria available for both low and high values, i.e., hypo- and hyper-, will be summarized for both criteria. The same subject can be counted for both criteria if the

subject has different laboratory values meeting each criterion. NCI-CTCAE grade of laboratory results will be summarized by number and percentage of subjects for each visit. Shift table of NCI-CTCAE grade change from baseline to each post-baseline visit will be presented by treatment arm and visit. Shift table of NCI-CTCAE grade change from baseline to worst post-baseline grade will also be presented by treatment arm. Number and percent of subjects with treatment-emergent NCI-CTCAE grade of 3 or 4 laboratory results will be presented by laboratory parameter. "Treatment-emergent NCI-CTCAE grade of 3 or 4 laboratory results" are defined as: subject having a maximum NCI-CTCAE grade of 3 or 4 post baseline for a parameter and that grade is higher than the subject's baseline grade for that parameter (or his/her baseline grade is missing).

The list of laboratory parameters to be summarized may change during the course of the study due to ongoing pharmacovigilance. It will be finalized before the database hard lock.

6.5.3.1 Liver Safety Assessment

The liver safety assessments will be summarized by the following categories below based on the measurements from Alkaline Phosphatase (ALP), Alanine Transaminase (ALT), total bilirubin, Aspartate Transaminase (AST) and their combination as defined.

The subject's highest value post-baseline of each parameter will be used.

- ALT: > 3xULN, > 5xULN, > 10xULN, > 20xULN
- AST: > 3xULN, > 5xULN, > 10xULN, > 20xULN
- ALT or AST: > 3xULN, > 5xULN, > 10xULN, > 20xULN
- ALP: > 1.5xULN
- Total Bilirubin: > 2xULN
- (ALT or AST > 3xULN) and Total Bilirubin > 2xULN
- (ALT or AST > 3xULN) and Total Bilirubin > 2xULN and ALP < 2xULN

The last 2 criteria where 2 or more parameters are evaluated will use the measurements on the same day or up to 1 day apart. The number and percentage of subjects meeting the criteria post-baseline will be summarized by treatment arm.

6.5.4 Vital Signs

Vital signs will be summarized using mean, standard deviation, minimum, maximum and median by treatment arm and visit. Additionally, a within-subject change will be calculated per visit as the post-baseline measurement minus the baseline measurement and summarized by treatment group and visit. For serial vital signs measured values and within-subject change from pre-dose value on that visit will be summarized by visit and timepoint in separate tables.

Tables for potentially clinically significant vital signs will be generated using baseline value and highest value obtained during treatment for each subject.

The following potentially clinically significant criteria are defined for each parameter:

Vital Sign Variable	Criteria
SBP	≥180 mmHg AND ≥20 mmHg change from baseline
SBP	≤ 80 mmHg
DBP	≥105 mmHg AND ≥15 mmHg change from baseline
Pulse Rate	≥120 bpm AND ≥15 bpm change from baseline

6.5.5 Electrocardiograms

ECG variables will be summarized using mean, standard deviation, minimum, maximum and median for each treatment group at each visit, including changes from baseline.

Number and percentage of subjects with normal and abnormal results for the overall interpretation will be tabulated by treatment group and time point. A shift analysis table showing shift in overall ECG interpretation from baseline to each time point will be provided. Percent of subjects on different kind of abnormality will also be reported.

The QT interval corrected by Fridericia's Correction formula (QTcF interval) will be summarized using frequency tables for values of clinical importance using the range criteria below.

QTcF Interval Criteria	QTcF Interval Value (msec)
Normal	≤ 450
Borderline	> 450 to <=480
Prolonged	> 480 to <=500
Clinically significant	> 500

QTcF interval: Fridericia-corrected QT interval

Cumulative tabulation for >450, >480 and >500 msec will also be presented.

The QTcF interval will also be summarized by the frequencies of subjects with a change from baseline of clinical importance using the criteria identified below.

Variable	Change from Baseline
QTcF Interval (msec)	<=0
	>0 to <=30
	>30 to <=60
	> 60

QTcF interval: Fridericia-corrected QT interval

Baseline value is from Cycle 1 Day 1 pre-dose assessment. Cumulative tabulation for >0, >30 and >60 msec will also be presented.

6.5.6 ECOG Performance Status

Number and percent of subjects of for each ECOG performance status grade will be presented at each visit. The change from baseline to each post baseline visit will be summarized. Negative change scores indicate an improvement. Positive change scores

indicate a decline in performance. Shift tables of ECOG performance status change from baseline to worst post-baseline grade will also be presented.

6.5.7 Physical Examination

Weight and change from baseline will be summarized by treatment arm and visit.

6.5.8 Pregnancy

A listing of all pregnancy tests will be provided.

6.6 Analysis of Pharmacokinetics

Descriptive statistics (e.g., N, mean, standard deviation, minimum, median, maximum, coefficient of variation [CV], geometric mean, and geometric CV) will be provided for serum concentrations of Zolbetuximab by scheduled sampling visit and time.

Trough concentrations versus visit profile will be described using box and whisker plots.

Additional model-based analyses may be performed and reported separately.

6.7 Subgroup Analyses

PFS, ORR and DOR based on IRC assessments and OS will be summarized for the following subgroups:

- Age group 1: <=65, >65 years
- Age group 2: <=75, >75 years
- Sex: male, female
- Race: White, Asian
- Tobacco history: Never, current, former
- Region: Asia, Non-Asia
- Number of organs with metastatic sites: 0-2, ≥ 3
- Prior gastrectomy (total or partial): No, Yes
- Histology (tumor type): diffuse vs. intestinal vs. mixed/other
- Tumor location: Gastric vs GEJ; Gastric proximal vs. Gastric distal; GEJ proximal vs GEJ distal;
- Country: Japan vs non-Japan; China vs non China

Forest plots for PFS and OS will be produced to summarize the treatment effect (HR) across subgroups.

6.8 Other Analyses

6.8.1 Immunogenicity

Immunogenicity of Zolbetuximab will be summarized using the frequency of ADA positive subjects. The potential relationship between Zolbetuximab immunogenicity and Zolbetuximab pharmacokinetics, efficacy, safety profile in subjects may be assessed.

Listing of efficacy variable will also be generated for the following subgroup:

- Subjects with negative ADA at baseline and positive at after receiving Zolbetuximab, or subjects with positive ADA and increased ADA titer after receiving Zolbetuximab

6.8.2 Exploratory Biomarkers

Biomarkers may be summarized graphically or descriptively, and summary statistics may be tabulated.

CLDN18.2 status will be summarized by different demographic characteristics as well as by some primary diagnosis variables. Associations between biomarkers and clinical (e.g., efficacy, safety, pharmacodynamics, or pharmacokinetics) measures may be performed on subjects who have sufficient baseline and on-study measurements to provide interpretable results for specific parameters. Analysis will be further described in a biomarker SAP.

6.9 Interim Analysis (and Early Discontinuation of the Clinical Study)

To evaluate whether Zolbetuximab + CAPOX (Arm A) is beneficial compared to the concurrent placebo + CAPOX (Arm B) while the study is ongoing, a formal OS interim analysis is planned when the final PFS analysis occurs with the pre-specified number of PFS events. A group sequential design using the O'Brien-Fleming type alpha-spending function [Lan & DeMets, 1983] will be utilized to control the overall 1-sided 0.025 significance level (East®) for the OS endpoint. The OS interim and final analyses will be performed only if primary PFS analysis is significant.

The IDMC may recommend terminating the study for favorable results at the formal efficacy interim analysis using OS. In the case of favorable results, the 1-sided significance level for superiority is 0.0112, assuming about 78% of the target number of OS events is obtained, for the interim OS analysis and 0.0217 for the final OS analysis (Note: The OS significance level will be adjusted depending on the number of OS event at the time of interim analysis). If the 1-sided P value of the interim analysis is less than the significance level (and PFS is also significant at 1-sided 0.025 alpha),) the IDMC may recommend terminating the study for success. If the study is not stopped after the interim analysis, a final OS analysis will occur after 100% of the planned death events have been observed.

The interim analysis will be run by an Independent Data Analysis Center (IDAC) for IDMC. In addition, safety data reviews during the trial will be conducted by the IDMC on a periodic basis. For example, the IDMC will have its first safety data review 6 weeks after the 40th subject has been randomized and on study drug for 2 cycle (6 weeks) and meetings will be conducted regularly thereafter, as specified in the IDMC Charter.

The full procedures for IDMC safety review will be described in a separate IDMC Charter. An interim analysis plan (IAP) will describe specific analyses to be presented for safety and the efficacy interim reviews.

6.10 Additional Conventions

6.10.1 Analysis Visits

Nominal visits as recorded on eCRF will be used in the by-visit summaries. Values from unscheduled visits will be included in the summary of extreme cases (e.g., summary of worst value post-baseline, summary of minimum value post-baseline, summary of maximum value post-baseline). For efficacy endpoints, all values (scheduled and unscheduled) will be included in the analysis.

For time course plots, actual study day (calculated using actual visit date – date of first dose +1) will be used.

6.10.2 Imputation Rules for Incomplete Dates

Every effort will be made to resolve missing or incomplete dates for adverse events and concomitant medications. If a partial date cannot be resolved, the most conservative imputation methods will be used to complete the missing information. As a general rule, if the month or year is missing, imputation should be avoided if possible. More details on date imputation, if needed, would be placed in the TLF specifications.

7 REVISION AND RATIONALE

7.1 List of Changes in SAP Version 2.0 from Version 1.0 (if applicable)

The changes from the approved SAP Version 1.0 (Dated dd-MMM-yyyy) to Version 2.0 that impact analyses are listed with the rationale in the table below.

SAP Section(s)	Description of Change(s)	Rationale
Section 4.3 in V1.0	Removal of section 4.3 Per Protocol set	The PPS is defined as the subjects who do not meet predetermined study entry and treatment criteria as well as those with lack of imaging assessment. The data from subjects meeting these criteria are unlikely to allow adequate assessment of potential impact on treatment benefit, possibly resulting in risk of bias. Therefore, the robustness of treatment benefit in Primary Endpoint will instead be assessed through sensitivity analyses applying different censoring rules.
Section 3	The number of PFS events required for the interim analysis of overall survival is reduced from 344 to 300.	The number of required PFS events has been adjusted based on the enrollment and event accrual rates to maintain the timing of Primary Analysis with adequate power which is > 93%.
Section 2.1.2, 2.2, 5.2, 6.4.2.2	Addition of health economics and outcomes research (HEOR) related key secondary endpoints, including physical function, Pain, and Global Health Score.	A key secondary endpoint for QOL measures has been added after FDA interaction in order to more specifically address the effect of Zolbetuximab in gastric/GEJ cancer, which impacts the risk/benefit assessment.
Section 6.4.1	Censoring rule has been updated from 'NACT before radiologic progression or death' to 'NACT before radiologic progression'	To be consistent with protocol v5.0, Section 7.4.1.1

8 REFERENCES

ICH Harmonized Tripartite Guideline E 3. Structure and Content of Clinical Study Reports, November 1995. (www.ich.org; Guidelines; "Efficacy" Topics)

ICH Harmonized Tripartite Guideline E 9. Statistical Principles for Clinical Trials, February 1998. (www.ich.org; Guidelines; "Efficacy" Topics)

Fayers PM, Aaronson NK, Bjordal K, Groenvold M, Curran D, Bottomley A, on behalf of the EORTC Quality of Life Group. The EORTC QLQ-C30 Scoring Manual (3rd Edition). Published by: European Organisation for Research and Treatment of Cancer. Brussels 2001.

Giesinger JM et al. Replication and validation of higher order models demonstrated that a summary score for the EORTC QLQ-C30 is robust. J. Clin. Epidemiol. 69:79-88, 2016.

Lan KKG, DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika.1983;70:659-63.

9 APPENDICES

9.1 EORTC QLQ-C30 questionnaire (version 3)

		Not at All	A Little	Quite a Bit	Very Much
1.	Do you have any trouble doing strenuous activities, like carrying a heavy shopping bag or a suitcase?	1	2	3	4
2.	Do you have any trouble taking a long walk?	1	2	3	4
3.	Do you have any trouble taking a short walk outside of the house?	1	2	3	4
4.	Do you need to stay in bed or a chair during the day?	1	2	3	4
5.	Do you need help with eating, dressing, washing yourself or using the toilet?	1	2	3	4

During the past week:		Not at All	A Little	Quite a Bit	Very Much
6.	Were you limited in doing either your work or other daily activities?	1	2	3	4
7.	Were you limited in pursuing your hobbies or other leisure time activities?	1	2	3	4
8.	Were you short of breath?	1	2	3	4
9.	Have you had pain?	1	2	3	4
10.	Did you need to rest?	1	2	3	4
11.	Have you had trouble sleeping?	1	2	3	4
12.	Have you felt weak?	1	2	3	4
13.	Have you lacked appetite?	1	2	3	4
14.	Have you felt nauseated?	1	2	3	4
15.	Have you vomited?	1	2	3	4
16.	Have you been constipated?	1	2	3	4

Please go on to the next page

During the past week:	**Not at All**	**A Little**	**Quite a Bit**	**Very Much**
17. Have you had diarrhea?	1	2	3	4
18. Were you tired?	1	2	3	4
19. Did pain interfere with your daily activities?	1	2	3	4
20. Have you had difficulty in concentrating on things, like reading a newspaper or watching television?	1	2	3	4
21. Did you feel tense?	1	2	3	4
22. Did you worry?	1	2	3	4
23. Did you feel irritable?	1	2	3	4
24. Did you feel depressed?	1	2	3	4
25. Have you had difficulty remembering things?	1	2	3	4
26. Has your physical condition or medical treatment interfered with your family life?	1	2	3	4
27. Has your physical condition or medical treatment interfered with your social activities?	1	2	3	4
28. Has your physical condition or medical treatment caused you financial difficulties?	1	2	3	4

For the following questions please circle the number between 1 and 7 that best applies to you

29. How would you rate your overall health during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

30. How would you rate your overall quality of life during the past week?

1	2	3	4	5	6	7
Very poor						Excellent

9.2 EORTC QLQ-30 Scoring

Scoring the EORTC QLQ-C30 version 3.0

Table 1: Scoring the QLQ-C30 version 3.0

	Scale	Number of items	Item range*	**Version 3.0** Item numbers	Function scales
Global health status / QoL					
Global health status/QoL (revised)†	QL2	2	6	29, 30	
Functional scales					
Physical functioning (revised)†	PF2	5	3	1 to 5	F
Role functioning (revised)†	RF2	2	3	6, 7	F
Emotional functioning	EF	4	3	21 to 24	F
Cognitive functioning	CF	2	3	20, 25	F
Social functioning	SF	2	3	26, 27	F
Symptom scales / items					
Fatigue	FA	3	3	10, 12, 18	
Nausea and vomiting	NV	2	3	14, 15	
Pain	PA	2	3	9, 19	
Dyspnoea	DY	1	3	8	
Insomnia	SL	1	3	11	
Appetite loss	AP	1	3	13	
Constipation	CO	1	3	16	
Diarrhoea	DI	1	3	17	
Financial difficulties	FI	1	3	28	

* *Item range* is the difference between the possible maximum and the minimum response to individual items; most items take values from 1 to 4, giving *range* = 3.
† (revised) scales are those that have been changed since version 1.0, and their short names are indicated in this manual by a suffix "2" – for example, PF2.

For all scales, the *RawScore*, *RS*, is the mean of the component items:

$$RawScore = RS = (I_1 + I_2 + ... + I_n)/n$$

Then for **Functional scales**:

$$Score = \left\{1 - \frac{(RS - 1)}{range}\right\} \times 100$$

and for **Symptom scales / items** and **Global health status / QoL:**

$$Score = \{(RS - 1)/range\} \times 100$$

Examples:

Emotional functioning

$$RawScore = (Q_{21} + Q_{22} + Q_{23} + Q_{24})/4$$
$$\text{EF } Score = \{1 - (RawScore - 1)/3\} \times 100$$

Fatigue

$$RawScore = (Q_{10} + Q_{12} + Q_{18})/3$$
$$\text{FA } Score = \{(RawScore - 1)/3\} \times 100$$

9.3 EORTC QLQ-OG25 questionnaire

EORTC QLQ – OG25 plus STO22 Belching subscale

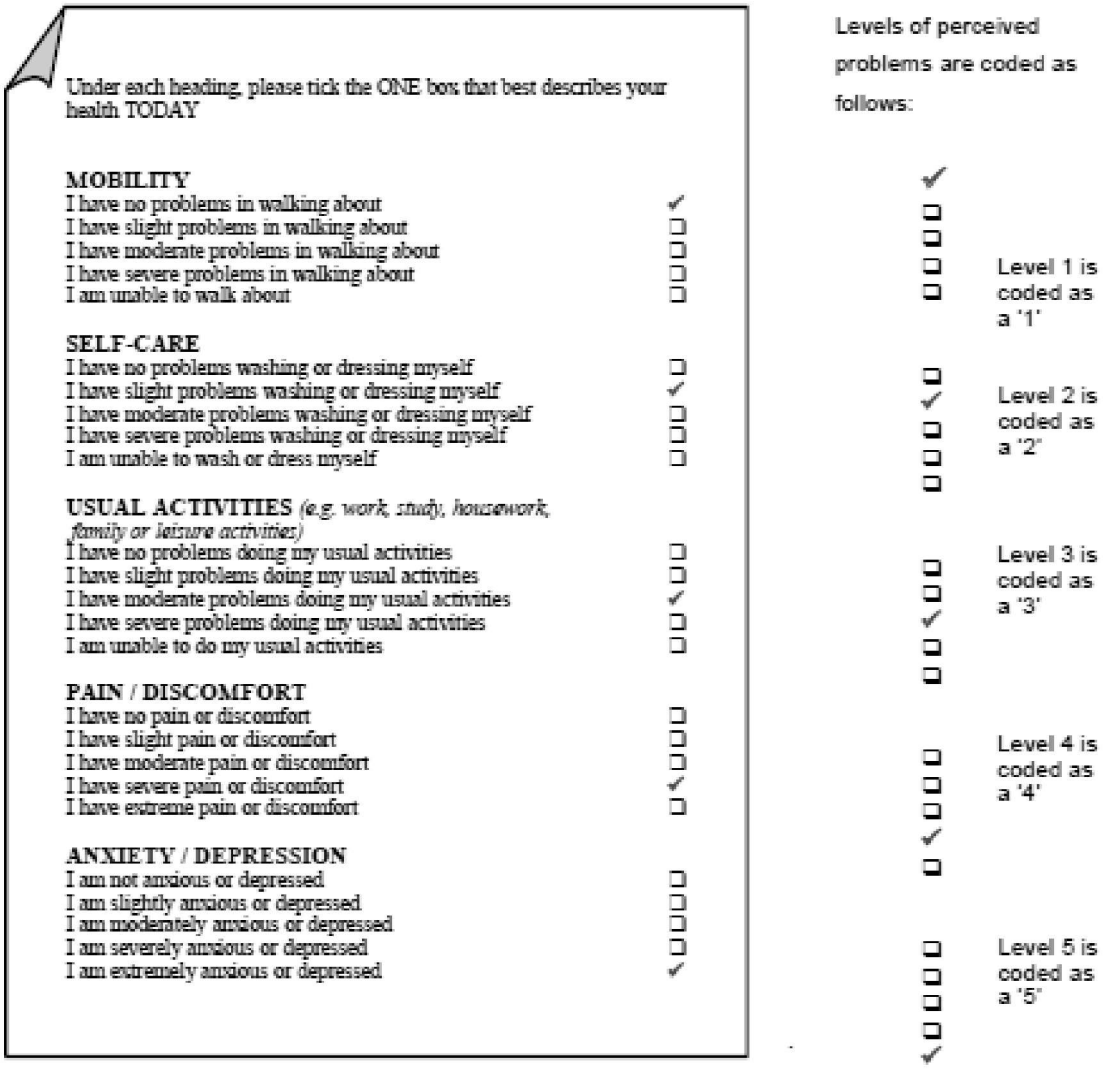
Patients sometimes report that they have the following symptoms or problems. Please indicate the extent to which you have experienced these symptoms or problems during the past week. Please answer by circling the number that best applies to you.

During the past week	Not at all	A little	Quite a bit	Very much
1. Have you had problems eating solid foods?	1	2	3	4
2. Have you had problems eating liquidised or soft foods?	1	2	3	4
3. Have you had problems drinking liquids?	1	2	3	4
4. Have you had trouble enjoying your meals?	1	2	3	4
5. Have you felt full up too quickly after beginning to eat?	1	2	3	4
6. Has it taken you a long time to complete your meals?	1	2	3	4
7. Have you had difficulty eating?	1	2	3	4
8. Have you had acid indigestion or heartburn?	1	2	3	4
9. Has acid or bile coming into your mouth been a problem?	1	2	3	4
10. Have you had discomfort when eating?	1	2	3	4
11. Have you had pain when you eat?	1	2	3	4
12. Have you had pain in your stomach area?	1	2	3	4
13. Have you had discomfort in your stomach area?	1	2	3	4
14. Have you been thinking about your illness?	1	2	3	4
15. Have you worried about your health in the future?	1	2	3	4
16. Have you had trouble with eating in front of other people?	1	2	3	4
17. Have you had a dry mouth?	1	2	3	4
18. Have you had problems with your sense of taste?	1	2	3	4
19. Have you felt physically less attractive as a result of your disease or treatment?	1	2	3	4
20. Have you had difficulty swallowing your saliva?	1	2	3	4
21. Have you choked when swallowing?	1	2	3	4
22. Have you coughed?	1	2	3	4
23. Have you had difficulty talking?	1	2	3	4
24. Have you worried about your weight being too low?	1	2	3	4
25. Answer this question only if you lost any hair: If so, were you upset by the loss of your hair?	1	2	3	4
26. Have you had trouble with bile or acid coming into your mouth?	1	2	3	4
27 have you had trouble with belching?	1	2	3	4

9.4 EQ-5D-5L Scoring

2. Scoring the EQ-5D-5L descriptive system

The EQ-5D-5L descriptive system should be scored, for example, as follows:



Under each heading, please tick the ONE box that best describes your health TODAY

MOBILITY
I have no problems in walking about ✓
I have slight problems in walking about ☐
I have moderate problems in walking about ☐
I have severe problems in walking about ☐
I am unable to walk about ☐

SELF-CARE
I have no problems washing or dressing myself ☐
I have slight problems washing or dressing myself ✓
I have moderate problems washing or dressing myself ☐
I have severe problems washing or dressing myself ☐
I am unable to wash or dress myself ☐

USUAL ACTIVITIES *(e.g. work, study, housework, family or leisure activities)*
I have no problems doing my usual activities ☐
I have slight problems doing my usual activities ☐
I have moderate problems doing my usual activities ✓
I have severe problems doing my usual activities ☐
I am unable to do my usual activities ☐

PAIN / DISCOMFORT
I have no pain or discomfort ☐
I have slight pain or discomfort ☐
I have moderate pain or discomfort ☐
I have severe pain or discomfort ✓
I have extreme pain or discomfort ☐

ANXIETY / DEPRESSION
I am not anxious or depressed ☐
I am slightly anxious or depressed ☐
I am moderately anxious or depressed ☐
I am severely anxious or depressed ☐
I am extremely anxious or depressed ✓

Levels of perceived problems are coded as follows:

Level 1 is coded as a '1'

Level 2 is coded as a '2'

Level 3 is coded as a '3'

Level 4 is coded as a '4'

Level 5 is coded as a '5'

This example identifies the health state '12345'.

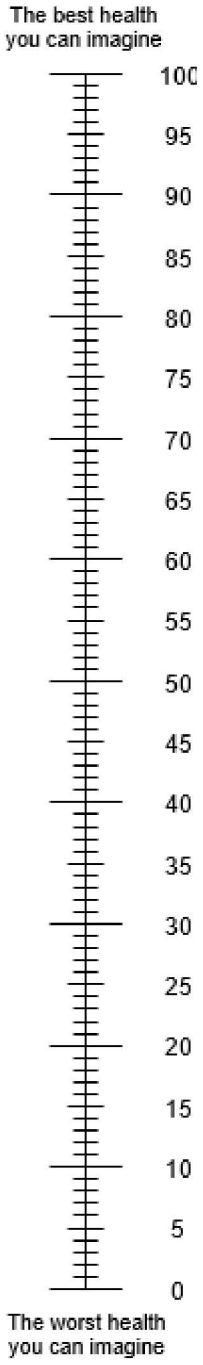
***NB:** There should be only ONE response for each dimension*

***NB: Missing values** can be coded as '9'.*

***NB: Ambiguous values** (e.g. 2 boxes are ticked for a single dimension) should be treated as missing values.*

- We would like to know how good or bad your health is TODAY.
- This scale is numbered from 0 to 100.
- 100 means the best health you can imagine.
  0 means the worst health you can imagine.
- Mark an X on the scale to indicate how your health is TODAY.
- Now, please write the number you marked on the scale in the box below.

YOUR HEALTH TODAY = ☐



9.5 Global Pain

The Global Pain item is below (the eCOA team can format with radial dials or check boxes)

Please rate your pain by selecting the one number that best describes your pain at its worst in the last 24 hours

0	1	2	3	4	5	6	7	8	9	10
No Pain										Pain as bad as you can imagine

9.6 Health Resource Utilization

1. Since your last study visit, have you had any visits to the emergency room (ER)?

   ☐ No (if no, go to 4)
   ☐ Yes (if yes, go to 2)

2. Since your last visit, how many emergency room visits have you had?

   _____

3. For each emergency room visit please complete the following:

	Result in a Hospital Admission (more than a 24 hour stay)?	Length of stay in hospital (number of days)
ER Visit 1	Yes/No	
ER Visit 2	Yes/No	
ER Visit 3	Yes/No	

4. Since your last study visit, have you had any hospital admissions (more than a 24 hour stay) that occurred without first going to the emergency room (ER)?

☐ No (if no, go to 7)
☐ Yes (if yes, go to 5)

5. How many hospital admissions (more than 24 hour stay; without previous ER transferal)? _____

6. For each hospital admission (more than 24 hour stay; without previous ER transferal) visit please complete the following:

	Length of stay in hospital (number of days)
Hospital Visit 1	
Hospital Visit 2	
Hospital Visit 3	

7. Since your last study visit, have you had any visits to a general practitioner (primary care physician)?

☐ No (if no, go to 9)
☐ Yes (if yes, go to 8)

8. How many visits have you had to a general practitioner (primary care physician)? _____

9. Since your last study visit, did you have any visits to a specialist physician (e.g., oncologist, rheumatologist, endocrinologist, orthopedic surgeon, etc.)?

☐ No
☐ Yes (if yes, go to 10)

How many visits have you had to a specialist physician (e.g., oncologist, rheumatologist, endocrinologist, orthopedic surgeon, etc.)?

9.7 Author and Approver Signatories

E-signatures are attached at end of document (next page). Wet signatures, if any, are provided on this page.

Author:	_____	Date:	_____
	Ran Li, PhD Study Statistician/ Associate Director, Biostatistics		
Approved by:	_____	Date:	_____
	Jung Wook Park, PhD GSTATL/Senior Director Biostatistics		
Approved by:	_____	Date:	_____
	Maria Matsangou, MD GML/Senior Medical Director		

Electronic Signature Page

Document Name: 8951-cl-0302-clsap-en-text

Signed By: Ran Li Capacity: Biostatistician	Signature Date: 03-Oct-2022 13:48:32 GMT+0000

Signed By: Jung Wook Park Capacity: Biostatistician	Signature Date: 03-Oct-2022 13:50:04 GMT+0000

Signed By: Maria matsangou Capacity: Therapeutic Area - Medical	Signature Date: 05-Oct-2022 15:14:24 GMT+0000

Document Type - Subtype: Clinical - Statistics and Data Management

Vault Document Number: VV-CLIN-024867 v3.0

Electronic Signature Page

Document Name: 8951-cl-0302-clsap-en-text-erratum

Signed By: PPD Capacity: Biostatistician	Signature Date: 20-Jan-2023 15:19:12 GMT+0000

Document Type - Subtype: Clinical - Statistics and Data Management

Vault Document Number: VV-CLIN-055975 v1.0