CLINICAL STUDY PROTOCOL

Study Title:	A Phase 3 Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of GS-5745 Combined with mFOLFOX6 as First Line Treatment in Patients with Advanced Gastric or Gastroesophageal Junction Adenocarcinoma	
Sponsor:	Gilead Sciences, Inc. 333 Lakeside Drive Foster City, CA 94404	
IND Number:	116561	
EudraCT Number:	2015-001526-42	
Clinical Trials.gov Identifier:	NCT02545504	
Indication:	Gastric Adenocarcinoma	
Protocol ID:	GS-US-296-1080	
Gilead Clinical Program Manager:	Name: Telephone: Fax: Email:	PPD PPD PPD PPD
Gilead Medical Monitor:	Name: Telephone: Email:	PPD PPD PPD
Protocol Version/Date:	Original: Amendment 1 Amendment 2 Amendment 3 Amendment 4 Amendment 5 Amendment 6	02 December 2014 15 June 2015 17 August 2015 14 December 2015 22 June 2016 14 December 2016 06 March 2017

TABLE OF CONTENTS

LIST OF IN-TEXT TABLES

PROTOCOL SYNOPSIS
Gilead Sciences, Inc.
333 Lakeside Drive
Foster City, CA 94404

Study Title:	A Phase 3 Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of GS-5745 Combined with mFOLFOX6 as First Line Treatment in Patients with Advanced Gastric or Gastroesophageal Junction Adenocarcinoma
IND Number: **EudraCT Number:** **Clinical Trials.gov Identifier:**	116561 2015-001526-42 NCT02545504
Study Centers Planned:	Approximately 188 centers globally
Study Phase:	Phase 3
Objectives:	The primary objective of this study is: • To compare the efficacy of GS-5745 versus placebo in combination with mFOLFOX6 as measured by overall survival (OS) The secondary objectives of this study are: • To compare the efficacy of GS-5745 versus placebo in combination with mFOLFOX6 as measured by progression-free survival (PFS) • To compare the efficacy of GS-5745 versus placebo in combination with mFOLFOX6 as measured by objective response rate (ORR) per Response Evaluation Criteria In Solid Tumors Version 1.1 (RECIST v1.1) • To compare the safety of GS-5745 versus placebo in combination with mFOLFOX6

The exploratory objectives of this study are:

Study Design:

This is a Phase 3, randomized, double-blind, multicenter study of GS-5745 combined with mFOLFOX6 in subjects with untreated gastric and gastroesophageal junction (GEJ) adenocarcinoma. A total of 430 eligible subjects with advanced gastric and GEJ adenocarcinoma will be randomized in a 1:1 manner to mFOLFOX6 plus GS-5745 or mFOLFOX6 plus placebo. Treatment assignment will be stratified by Eastern Cooperative Oncology Group (ECOG) 0 or 1, geographic region (Latin America or other participating countries), and primary tumor site (gastric or GEJ).

Computed Tomography (CT) or Magnetic Resonance Imaging (MRI) scans will be performed every 8 weeks to evaluate response to treatment by RECIST v1.1.

Dosage and frequency will be as follows:

mFOLFOX6 on Days 1 and 15 of each 28-day treatment cycle (1 cycle of mFOLFOX consists of 2 infusions) for a total of 6 cycles followed by leucovorin (LV) and 5-fluorouracil (5-FU) dosing on Days 1 and 15 of each 28-day treatment cycle until disease progression. The mFOLFOX6 dosing regimen will consist of *l*-LV 200 mg/m^2 or *dl*-LV 400 mg/m^2 and oxaliplatin 85 mg/m^2 followed by bolus 5-FU 400 mg/m^2 and a 46-hour infusion of 5-FU 2400 mg/m^2. Minor modifications to the duration of the infusion time are permitted as per institutional standard. Adjustments to the dose of mFOLFOX6 are permitted in response to treatment emergent adverse events. The preferred formula for calculating body surface area (BSA) will be the Mosteller formula below and rounded to the nearest tenth:

$\sqrt{}$([Height (cm) × Weight (kg)]/3600)

	However, institutional guidelines/practice for calculating BSA will also be allowed. • GS-5745/placebo 800 mg every 2 weeks, on Days 1 and 15 of each 28 day cycle, until disease progression An independent data monitoring committee (DMC) will review the progress of the study and perform interim reviews of safety data. The first safety review by the DMC will be performed after the equivalent of 4 treatment cycles from the time the 60th subject is randomized. Thereafter, review of safety data will be performed at regular intervals as described in the DMC charter. In addition, the DMC will meet after approximately 33.3% and 66.7% of the expected number of events have occurred to review the results from the futility and efficacy interim analysis, respectively.
Number of Subjects Planned:	Approximately 430 subjects
Target Population:	Subjects who are ≥18 years old with histologically confirmed, inoperable, locally advanced or metastatic adenocarcinoma of the stomach or GEJ and who have received no prior treatment for advanced or metastatic gastric cancer.
Duration of Treatment:	Eligible subjects will be randomized to receive study drug (GS-5745/placebo) in combination with mFOLFOX6 until disease progression, unacceptable toxicity, or withdrawal of consent.
Duration of Study:	Each cycle will consist of 28 days and will continue in the absence of disease progression or unacceptable toxicity.
Eligibility Criteria:	<u>Inclusion Criteria:</u> Subjects must meet all of the following inclusion criteria to be eligible for participation in this study: 1) Male or female ≥18 years of age 2) Histologically confirmed adenocarcinoma of the stomach or GEJ with inoperable, locally advanced or metastatic disease, not amenable to curative therapy Adenocarcinoma of the GEJ is defined as tumors that have their center within 5 cm proximal and distal of the anatomical esophagogastric junction as described in Siewert's classification system 3) Eastern Cooperative Oncology Group (ECOG) ≤ 1

4) Measurable disease or non-measurable but evaluable disease, according to RECIST v1.1. Subjects with peritoneal disease would generally be regarded as having evaluable disease and allowed to enter the trial
5) Subjects not receiving anticoagulant medication must have an international normalized ratio (INR) ≤ 1.5 and activated partial thromboplastin (aPTT) ≤ 1.5 x upper limit of normal (ULN)

   The use of full-dose oral or parenteral anticoagulants is permitted as long as the INR or aPTT is within therapeutic limits (according to the medical standard in the institution) and the subject has been on stable dose of anticoagulants for at least 1 week at the time of randomization
6) Adequate hematologic function:
   a) neutrophils ≥ 2.0 x 10^9/L
   b) platelets ≥ 100 x 10^9/L
   c) hemoglobin ≥ 9 g/dL
7) Adequate hepatic function:
   a) Direct or total bilirubin ≤ 1.5 x ULN
   b) ALT and AST ≤ 2.5 x ULN, in case of liver metastases ≤ 5 x ULN
8) Creatinine clearance (CL_{cr}) should be ≥ 30 mL/min based on the Cockroft-Gault formula. Subjects with a CL_{cr} just below30 mL/min may be eligible if a measured creatinine clearance (based on 24 hour urine collection or other reliable method) is ≥ 30 mL/min
9) For female subjects of childbearing potential, willingness to use a protocol-recommended method of contraception from the screening visit throughout the study treatment period, for 90 days following the last dose of study drug (GS-5745/placebo), and for 4 months after the last dose of oxaliplatin or 6 months after the last dose of 5-FU whichever occurs later unless the subject chooses continuous heterosexual abstinence as a lifestyle-choice (see Appendix 3 for more information)
10) For male subjects of reproductive potential having intercourse with females of childbearing potential, willingness to use a protocol recommended method of contraception and to refrain from sperm donation from the start of study drug, throughout the study treatment period, for 90 days after administration of the last dose of any study drug, and for 6 months after the last dose of oxaliplatin or 6 months following the last dose of 5-FU whichever occurs later (see Appendix 3)

11) Breastfeeding females must agree to discontinue nursing before study drug administration

12) In the judgment of the investigator, participation in the protocol offers an acceptable benefit-to-risk ratio when considering current disease status, medical condition, and the potential benefits and risks of alternative treatments for the subject's cancer

13) Willingness to comply with scheduled visits, drug administration plan, imaging studies, laboratory tests, other study procedures, and study restrictions

14) Evidence of a signed informed consent prior to implementation of any protocol specific procedure

Exclusion Criteria

Subjects who meet any of the following exclusion criteria are not to be randomized in this study:

1) Previous chemotherapy for locally advanced or metastatic gastric or GEJ cancer. Subjects may have received prior neoadjuvant or adjuvant chemotherapy as long as it was completed at least 6 months prior to randomization

2) Human Epidermal Growth Factor Receptor 2 (HER2) cancer (primary tumor or metastatic lesion). HER2-positivity is defined as either IHC3+ or IHC2+/ISH+ (ISH positivity is defined as a HER2:CEP17 ratio of ≥2.0.)

3) Patients who have received palliative radiation and have not recovered from all acute, reversible effects

4) Uncontrolled intercurrent illness including, but not limited to, active uncontrolled infection, active gastrointestinal bleeding, uncontrolled cardiac arrhythmia, or psychiatric illness/social situation that would limit compliance with study requirements as judged by treating physician

5) History of a concurrent or second malignancy except for adequately treated local basal cell or squamous cell carcinoma of the skin, cervical carcinoma in situ, superficial bladder cancer, asymptomatic prostate cancer without known metastatic disease and with no requirement for therapy or requiring only hormonal therapy and with normal prostate-specific antigen for ≥ 1 year prior to randomization, adequately treated Stage 1 or 2 cancer currently in complete remission, or any other cancer that has been in complete remission for ≥ 5 years

6) Major surgery, defined as any surgical procedure that involves general anesthesia and a significant incision (ie, larger than what is required for placement of central venous access, percutaneous feeding tube, or biopsy), within 28 days of first dose of study drug
7) Known positive status for human immunodeficiency virus (HIV)
8) Known acute or chronic-active infection with hepatitis B virus (HBV) or hepatitis C virus (HCV)
9) Peripheral neuropathy ≥ Grade 2 according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) v.4.03
10) Chronic daily treatment with oral corticosteroids (dose of > 10 mg/day methylprednisolone equivalent). Inhaled steroids and short courses of oral steroids for anti-emesis or as an appetite stimulant are allowed
11) Pregnant or breastfeeding women (pregnancy needs to be excluded by testing of beta-human chorionic gonadotropin [β-hCG])
12) Known or suspected central nervous system metastases
13) Known dihydropyrimidine dehydrogenase-deficiency (special screening not required)
14) Known alcohol or drug abuse or any other medical or psychiatric condition which contraindicates participation in the study
15) Documented myocardial infarction or unstable/uncontrolled cardiac disease (ie, unstable angina, congestive heart failure [New York Heart Association > Class II]) within 6 months of randomization
16) Active tuberculosis or history of latent tuberculosis that has not been treated
17) Any chronic medical condition that, in the opinion of the Investigator, would make the subject unsuitable for the study or would prevent compliance with the study protocol.
18) Serious systemic fungal, bacterial, viral, or other infection that is not controlled or requires intravenous antibiotics
19) Experimental medical treatment within 28 days prior to randomization
20) Known hypersensitivity to any of the study drugs or excipients or to Chinese hamster ovary cell products or to recombinant human or humanized antibodies, or known allergic reactions to products that contain platinum

21) History of long QT syndrome or whose corrected QT interval (QTc) measured using Fridericia's formula ($QTcF = QT/RR^{0.333}$) at screening is prolonged (> 450 ms for males and > 470 ms for females)

22) Subjects with potassium, magnesium or calcium less than the lower limit of normal (LLN); electrolyte replacement is permitted during screening

Study Visits: Screening will commence with obtaining the subject's signed informed consent and will occur up to 28 days prior to the first dosing of study drug on Day 1. Screening procedures will include the following: medical history review, physical exam, vital signs, 12-lead ECG, ECOG performance status, prior/concomitant medication review, blood collection for pregnancy test (females), chemistry, hematology and coagulation, adverse event (AE) assessment, and CT or MRI. Baseline tumor lesions will be measured and characterized prior to randomization to assess the subject's disease status prior to beginning treatment.

Treatment:

Treatment will occur over cycles comprised of 28 days. Subjects who meet eligibility will undergo CT or MRI scan performed every 8 weeks. Beginning with Day 1 of Cycle 1, subjects will receive GS-5745 or placebo by intravenous infusion over 30 minutes every 2 weeks for a total of 2 infusions per cycle (Day 1 and Day 15 of each cycle). All subjects will also receive mFOLFOX6 for the first 6 cycles and 5-fluorouracil (5-FU) and leucovorin (LV) thereafter administered via intravenous infusion after GS-5745 or placebo on Days 1 and 15 of each cycle. The mFOLFOX6 dosing regimen will consist of *l*-LV 200 mg/m^2 or *dl*-LV 400 mg/m^2 and oxaliplatin 85 mg/m^2 followed by bolus 5-FU 400 mg/m^2 and a 46-hour infusion of 5-FU 2400 mg/m^2.

Treatment will continue with subsequent 28-day treatment cycles in combination with intravenous mFOLFOX6 for a total of 12 infusions over 6 cycles via the same regimen defined above. After the subject has received 6 cycles of mFOLFOX6, the oxaliplatin component will be discontinued, and 5-FU, leucovorin, and GS-5745/placebo will be continued as maintenance therapy in the absence of disease progression or toxicity warranting discontinuation of therapy. CT or MRI for assessment of tumor status will be conducted every 8 weeks. However, tumor response may be assessed prior to the specified every 8-week time point, if clinically indicated.

Investigational Drug:	Subjects will be administered 800 mg GS-5745/placebo intravenous over approximately 30 minutes and in advance of mFOLFOX6 on Days 1 and 15 of each cycle.
Reference Therapy, Dose, and Mode of Administration:	The dosing regimen for each subject will adhere to the protocol specifications. mFOLFOX6 will consist of *l*-LV 200 mg/m^2 or *dl*-LV 400 mg/m^2 and oxaliplatin 85 mg/m^2 followed by bolus 5-FU 400 mg/m^2 and a 46-hour infusion of 5-FU 2400 mg/m^2.
Criteria for Evaluation:	All subjects meeting the eligibility criteria that have signed a consent form and have begun treatment will be evaluated for response. Efficacy Overall survival (OS) is the primary endpoint of the study and is measured as time from date of randomization to death from any cause. Progression-free survival (PFS) is measured as the interval of time from randomization to the earlier of first documentation of definitive disease progression or death from any cause. Subjects who discontinue the study drug before disease progression will continue to be followed-up until they have documented disease progression. Objective response is assessed by the RECIST v1.1 as Complete Response (CR), Partial Response (PR), Stable Disease (SD) Non-CR/Non-PD (NN) or Progressive Disease (PD). The response of Not Evaluable (NE) will be recorded for subjects who drop out early before the scheduled imaging is performed, or for images with poor quality. Objective response rate (ORR) is the proportion of subjects who achieve a CR or PR. Safety The safety evaluation will be based on incidence of adverse events, clinically relevant changes in laboratory values and vital signs. Pharmacokinetics Blood samples to measure GS-5745 and oxaliplatin/5-FU concentrations will be collected at the timepoints specified in the protocol. Blood samples to measure anti-GS-5745 antibodies will also be collected at the time points specified in the protocol. Pharmacodynamic and Exploratory Biomarkers CCI [redacted].

Statistical Methods:

CCI [redacted]

Analysis Methods

The primary efficacy analysis set will be intent-to-treat (ITT). The ITT analysis set includes all randomized subjects and will be analyzed according to treatment assigned.

The Kaplan-Meier method and stratified log-rank test will be used to compare the 2 treatment groups for OS and PFS. A Cox proportional hazard model will be used to estimate the hazard ratio and corresponding 95% confidence interval (CI).

For the analysis of ORR, a Cochran-Mantel-Haenszel (CMH) Chi-square test will be performed to compare the 2 treatment groups.

In general, count and percent of subjects will summarize categorical and ordinal data. Mean, standard deviation, minimum, quartiles, median and maximum will summarize continuous data.

Interim Analysis

Two interim analyses are planned: the futility interim will be performed at 33.3% information and the efficacy interim at 66.7% information. The final analysis will occur when 286 OS events have been observed.

The Lan-DeMets approach with O'Brien-Fleming type alpha spending function will be used for efficacy data monitoring. The stopping boundaries at each analysis time are provided in the table below.

Stopping Boundaries for Efficacy Analyses

Efficacy Analysis	Events (%)	Stopping Boundary	
		Z scale	One-sided p-value scale
Interim	191 (66.7%)	2.509	0.006
Final	286 (100%)	1.993	0.023

Sample Size

Assuming a median OS time for the mFOLFOX6 + placebo group of 11.5 months, 286 events are needed to detect a hazard ratio (HR) of 0.70 with 85% power at a 1-sided significance level of 0.025. With an accrual period of 18 months, minimum follow-up of 18 months, and a 10% annual dropout rate, a total sample size of 430 subjects (215 subjects per treatment group) is needed to observe the required 286 events within the 36-month time frame.

As the targeted number of deaths is large (~286), if the null hypothesis of the primary end point of OS is rejected, it will convincingly demonstrate clinical treatment effect and would provide a narrow confidence interval. Based on the number of deaths and the assumed treatment effect on OS in the protocol, the expected 95% confidence interval on the HR of OS between the 2 treatment groups is ~ (0.627, 0.996). At the time of final analysis and assuming OS is significant, PFS will be tested at one-sided type I error of 0.016. Assuming that the hazard ratio in PFS is 0.7, which is expected to be on par or better than the treatment effect in OS, and that median PFS in control is 9 months, a sample size of 430 subjects (322 PFS events) will provide 85% power under the planned study enrolment (18 months), follow-up duration (18 month) and drop-out rate (annually 10%).

This study will be conducted in accordance with the guidelines of Good Clinical Practices (GCPs), including archiving of essential documents.

GLOSSARY OF ABBREVIATIONS AND DEFINITION OF TERMS

β-hCG	beta-human chorionic gonadotropin
λz	terminal elimination rate constant; estimated by linear regression of the terminal elimination phase of the log serum/plasma/PBMC concentration versus time curve of the drug
5-FU	5-fluorouracil
AE	adverse event
ALT	alanine aminotransferase
ANC	absolute neutrophil count
AST	aspartate aminotransferase
AUC	area under the concentration versus time curve
AUC_{last}	area under the plasma/serum/PBMC concentration versus time curve from time zero to the last quantifiable concentration
AUC_{tau}	area under the plasma/serum/PBMC concentration versus time curve over the dosing interval
BSA	body surface area
C_{max}	maximum observed concentration of drug
CFR	Code of Federal Regulations
CI	confidence interval
C_{last}	last observed quantifiable serum/plasma/PBMC concentration of the drug
CL_{cr}	creatinine clearance
C_{max}	maximum observed serum/plasma/PBMC concentration of drug
CR	complete response
CRO	contract research organization
CT	computed tomography
Ctau	observed drug concentration at the end of the dosing interval
CTCAE	Common Terminology Criteria for Adverse Events
DCR	disease control rate
DMC	data monitoring committee
DNA	deoxyribonucleic acid
DOR	duration of response
DSPH	Drug Safety and Public Health
ECG	electrocardiogram
ECOG	Eastern Cooperative Oncology Group
eCRF	electronic Case Report Form
EORTC	European Organization for Research and Treatment of Cancer
EOS	end of study
EOT	end of treatment
eSAE	electronic serious adverse event
EU	European Union
FDA	Food and Drug Administration
GCP	good clinical practice

GEJ	gastroesophageal junction
HBV	hepatitis B virus
HCV	hepatitis C virus
HER2	human epidermal growth factor receptor 2
HIV	human immunodeficiency virus
HLT	high-level term
HR	hazard ratio
IB	investigator's brochure
IC_{50}	concentration necessary to achieve 50% inhibition of target
ICH	International Conference on Harmonisation
IHC	immunohistochemistry
IMP	investigational medicinal product
INR	international normalized ratio
IRB/IEC	institutional review board or independent ethics committee
ISH	in situ hybridization
ITT	intent-to-treat
IUD	intrauterine device
IWRS	interactive web response system
KM	Kaplan-Meier
LV	leucovorin
LTFU	long term follow up
MMP	matrix metalloproteinases
MRI	magnetic resonance imaging
MSS	musculoskeletal syndrome
ORR	objective response rate
OS	overall survival
PD	progressive disease
PE	physical examination
PFS	progression-free survival
PK	pharmacokinetic
PR	partial response
PRO	patient-reported outcomes
QoL	quality of life
RECIST	Response Evaluation Criteria In Solid Tumors
PT	preferred term
RNA	ribonucleic acid
SADR	serious adverse drug reactions
SAE	serious adverse event
SD	stable disease
SOC	system organ class

SOP	standard operating procedure
SUSAR	suspected unexpected serious adverse reactions
t½	estimate of the terminal elimination half-life of the drug in serum/plasma/PBMC, calculated by dividing the natural log of 2 by the terminal elimination rate constant (λz)
TIMPs	tissue inhibitors of metalloproteinases
T_{last}	time (observed time point) of C_{last}
TTR	time to response
ULN	upper limit of normal

1. INTRODUCTION

1.1. Background

Matrix metalloproteinases (MMP) comprise a family of at least 23 Zn^{2+}-dependent proteases which are primarily involved in the degradation and remodeling of the extracellular matrix and basement membranes in many normal as well as pathologic biological processes. They are typically grouped based on their structure or their primary substrates and include the gelatinases, collagenases, stromelysins, matrilysins, an elastase, and membrane-type MMP, a group of cell surface tethered proteases {Hu 2007, Mott 2004}. The gelatinases comprise MMP2 and MMP9, sometimes referred to as type IV collagenases, which are named for their ability to degrade type IV collagen and gelatin, a denatured form of collagen {Chen 2002, Kridel 2001}. The contrasting roles of MMP9 and MMP2 have been revealed in a variety of studies which support a more ubiquitous expression pattern and associated role for MMP2 in normal tissue homeostasis, as compared to disease-induced and pathology-associated expression and activity of MMP9 {Agrawal 2006, Garg 2009, Hu 2007, Itoh 2002, Castaneda 2005, Dubois 1999, Li 2009, Miyazaki 2011, Naito 2005, Santana 2006}. Additional substrates have been identified for MMP9, and the active enzyme can release cytokines, growth factors, and bioactive fragments which in turn modulate inflammation, neovascularization, and matrix remodeling {Hijova 2005}. MMP9 is an inducible MMP that is secreted as a zymogen and activated in a "cysteine switch" mechanism by the cleavage of the peptidoglycan binding domain {Van Wart 1990}. While activation of MMP9 appears to be carried out by other MMPs, the protease's activity is also regulated by the binding of tissue inhibitors of metalloproteinases (TIMPs), primarily by TIMP1 {Imai 1995, Olson 1997, Vempati 2007}. Elevated MMP9 expression in diseased tissue and plasma is associated with several human diseases. The health and largely normal development of the MMP9 knockout mouse has enabled evaluation in a variety of disease models, and these data support a significant role for MMP9 in a variety of inflammatory, fibrotic, and oncologic processes {Dubois 1999, Hu 2007, Itoh 2002, Itoh 1999, Opdenakker 2003}.

More recent studies in the MMP field have revealed diversity in the functional roles of MMPs in disease and normal homeostasis, suggesting a therapeutic opportunity for selective inhibitors. Despite their structural similarities, expression analysis in human disease and data from knockout mice reveal contrasting roles for MMP9 and MMP2 regulation and activity in normal homeostasis and in disease. MMP9 expression is restricted to limited cell types in healthy tissues whereas MMP2 is found to be more constitutively expressed {Hu 2007}. The disease-associated induction and functions of MMP9 render it an attractive therapeutic target.

1.1.1. Gastric Adenocarcinoma

Adenocarcinoma of the stomach is the most common gastrointestinal cancer in the world and the third leading cause of cancer death worldwide{Ferlay 2013}. Approximately 22,220 patients are diagnosed annually in the United States, of whom 10,990 are expected to die. While the incidence of distal gastric adenocarcinoma has recently declined in the United States, gastric cancer remains quite frequent in certain minority populations and it is still the second most

common cause of cancer death worldwide. In addition, adenocarcinoma of the gastroesophageal junction (GEJ) is one of the most rapidly increasing solid tumors in the United States and Western Europe.

Most patients with gastric cancer in the United States are symptomatic and already have advanced incurable disease at the time of presentation. At diagnosis, approximately 50 percent have disease that extends beyond locoregional confines, and only one-half of those who appear to have locoregional tumor involvement can undergo a potentially curative resection. Surgically curable early gastric cancers are usually asymptomatic and only infrequently detected outside the realm of a screening program. Screening is not widely performed, except in countries which have a very high incidence, such as Japan, Venezuela, and Chile. The common presenting symptoms and diagnostic approaches to gastric cancer include weight loss (usually results from insufficient caloric intake rather than increased catabolism) and may be attributable to anorexia, nausea, abdominal pain, early satiety, and/or dysphagia. Abdominal pain is often present which tends to be epigastric, vague and mild early in the disease but more severe and constant as the disease progresses. Dysphagia is a common presenting symptom in patients with cancers arising in the proximal stomach or at the esophagogastric junction. Patients may also present with nausea or early satiety from the tumor mass or in cases of an aggressive form of diffuse-type gastric cancer called linitis plastica, from poor distensibility of the stomach. They may also present with a gastric outlet obstruction from an advanced distal tumor.

In metastatic disease, multiple single agent studies using drugs such as cisplatin, docetaxel, 5-fluorouracil (5-FU), and irinotecan among others have demonstrated modest activity. Treatment of metastatic HER2 negative cancer in the first line setting consists of combination chemotherapy with a triplet or doublet regimen. The triplet regimen of docetaxel/cisplatin/5-FU is the only approved combination in the United States. Because of significant toxicity, this combination is reserved for patients who are medically fit and have good performance status. The National Comprehensive Cancer Network guidelines for the treatment of metastatic disease recommend a 2-drug regimen. A common regimen includes a fluoropyrimidine (5-FU or capecitabine) and platinum agent (cisplatin or oxaliplatin). The current standard of care for patients who progress on front line regimens remains controversial. While many patients are treated in the second line setting with taxanes, some are treated with irinotecan either by itself or in combination with a taxane. Ramucirumab, an anti-vascular endothelial growth factor receptor 2 antibody, was recently approved in the United States as monotherapy or in combination with paclitaxel to treat gastric and GEJ cancer that has progressed after treatment with a fluoropyrimidine- or platinum-containing regimen. While there is no clear standard regimen for this disease, there is also a lack of clinical trials addressing this patient population. The majority of patients with adenocarcinomas of the esophagus and stomach are diagnosed with either stage III or IV disease and the prognosis is very poor with 5-year survival rates between 5-15%. Treatment options after first-line therapy are still limited in the metastatic setting. Traditional cytotoxic chemotherapy regimens for second and third-line treatment in patients with metastatic disease have shown a median overall survival ranging from 3-5 months compared to 2.4-3.6 months for best supportive care {Ford 2014, Kang 2012, Thuss-Patience 2011}.

1.1.2. MMP9 Expression in Oncology

Matrix metalloproteinase 9 is expressed by tumor epithelia as well as infiltrating macrophages, other inflammatory cells, fibroblastic stroma, and tumor-associated endothelial cells. Expression of MMP9 by tumor epithelia in particular has been implicated in many pro-tumorigenic processes and is associated either with loss of tumor suppressor or gain of oncogenic activity, as a temporal response to oncogenic changes in local tumor environment, or during processes such as invasion and proliferation. MMP9 expression by tumor-associated macrophages, neutrophils, mesenchymal-derived suppressor cells and other cell types in in the tumor microenvironment is also associated with local pro-tumorigenic immunomodulation and angiogenesis {Farina 2014}. In gastric tumors, MMP9 expression is consistently observed, often in both tumor epithelia and stromal compartments.

1.2. General Information

GS-5745 is a humanized monoclonal antibody that binds with high affinity to human MMP9, but not to other human MMPs. GS-5745 was derived from the murine anti-human MMP9 monoclonal antibody, AB0041, and shares the same binding characteristics. GS-5745 and AB0041 cross-react with and inhibit rat and cynomolgus monkey MMP9 but not murine MMP9. AB0046, which cross-reacts with and inhibits murine MMP9, was generated via immunization in MMP9 knockout mice. Epitope mapping analysis revealed that AB0046 binds a similar region in murine MMP9 to that bound by GS-5745 and AB0041 on human MMP9.

1.2.1. Preclinical Pharmacology and Toxicology

1.2.1.1. Pharmacology

The therapeutic potential of inhibitory antibodies targeting human MMP9 (AB0041) and mouse MMP9 (AB0046) was evaluated in a surgical orthotopic xenograft mouse model of colorectal carcinoma in which tumors were derived from the human tumor cell line HCT116. In this treatment model, selective inhibition of MMP9 using a cocktail of anti-human MMP9 and anti-mouse MMP9 antibodies significantly reduced growth of the primary tumor and reduced the incidence of metastases in multiple independent studies. In xenografts, treatment with either an anti-tumor-MMP9 (human) or anti-stromal MMP9 (murine) antibody yielded significant tumor growth reduction highlighting important roles for tumor epithelial-derived and stromal-derived MMP9 in primary tumor outgrowth. However, targeting of stromal MMP9 was necessary for maximum efficacy with respect to incidence of metastases, highlighting the disease-associated role of other cellular sources of MMP9 in tumorigenesis.

The major dose limiting toxicity observed in clinical studies with pan-MMP inhibitors, such as marimastat, was musculoskeletal syndrome (MSS) consisting of tendonitis manifested by joint stiffness, edema, reduced mobility, and skin discoloration. A study to evaluate the potential of an anti-MMP9 antibody to induce MSS was conducted in Lewis rats. Unlike the pan-MMP inhibitor, Marimastat, AB0041 did not induce any evidence of MSS or other toxicities in this Lewis rat MSS model.

Further details on the non-clinical pharmacology are available in the GS-5745 Investigator's Brochure (IB).

1.2.1.2. Toxicology

The toxicity profile of GS-5745 has been assessed in rats and monkeys administered GS-5745 intravenous once weekly for up to 4 weeks (5 doses) at up to 100 mg/kg/dose. Target organs identified in the 4-week studies were the physeal bone (rat) and adrenal gland (monkey). The physeal hypertrophy in rats is likely directly attributable to inhibition of MMP9 as similar findings were observed in MMP9 null mice {Vu 1998} and in children with mutations in MMP9 and MMP13 {Lausch 2009}. In both mice and children, this is a transient finding that spontaneously regresses as the bone matures. The physeal hypertrophy noted in rats is not considered relevant to adult humans because the growth plates are closed and longitudinal bone growth is no longer ongoing in adults. The relationship of the adrenal gland weight increase to MMP9 inhibition is unknown. Both the physeal bone and adrenal gland findings were reversible after discontinuing GS-5745 treatment. Because these findings were minimal and/or did not impact the overall health of the animals, these findings were not considered adverse. The no observed adverse effect levels in both the rat and monkey studies were 100 mg/kg/dose intravenous once weekly for 4 weeks (5 doses).

1.2.2. Clinical Trials of GS-5745

In addition to the clinical development program in solid tumors, GS-5745 is being developed for the treatment of rheumatoid arthritis and cystic fibrosis. Details of the clinical studies in these diseases can be found in the IB.

Study GS US-296-0101 is a Phase 1, open-label, dose-escalation study to evaluate safety, pharmacokinetics (PK), and pharmacodynamics following multiple intravenous administrations of GS-5745 every 2 weeks in subjects with advanced solid tumors that are refractory to or intolerant of standard treatment, or for which no standard treatment is available. Multiple doses between 200 and 1800 mg are being evaluated.

1.3. Rationale for This Study

The preliminary clinical efficacy data in patients with gastric and GEJ adenocarcinoma treated with GS-5745 (800 mg every 2 weeks) and mFOLFOX6 in the Phase 1 study (GS-US-296-0101) suggests that the combination is well tolerated and has the potential to provide benefit over treatment with mFOLFOX6 alone. Serum biomarkers including the collagen neo-epitopes that are likely to reflect MMP9 enzymatic activity decrease during the first cycle of treatment with these agents. The 800 mg every 2 weeks regimen of GS-5745 is expected to reach linear range of pharmacokinetics (ie, saturate target-mediated drug disposition at trough concentrations) and to achieve steady trough concentrations greater than 26 μg/mL in majority of subjects, which is 260-fold over the IC_{50} (0.691 nM or 0.1 μg/mL) evaluated in enzymatic inhibitory activities of GS-5745 in vitro. Together, these data support the hypothesis that GS-5745 treatment inhibits MMP9 activity and that the inhibition may lead to improved clinical outcomes. The study aims

to assess in a randomized, blinded fashion whether GS-5745 in combination with mFOLFOX6 improves therapeutic outcomes in subjects with previously untreated locally advanced or metastatic gastric and GEJ adenocarcinoma.

1.3.1. Information about mFOLFOX-6

Treatment of metastatic HER2 negative esophagogastric adenocarcinoma consists of combination chemotherapy with a triplet or doublet cytotoxic regimen. The triplet regimen of docetaxel/cisplatin/5-FU is the only approved combination in the US. Because of significant toxicity, this combination is reserved for patients who are medically fit and have good performance status. The National Comprehensive Cancer Network and European Society of Medical Oncology guidelines for the treatment of metastatic disease in the first line setting recommend a 2-drug regimen. A common regimen includes a fluoropyrimidine (5-FU or capecitabine) and platinum agent (cisplatin or oxaliplatin). There are no definitive studies demonstrating the superiority of either fluoropyrimine or platinum. An mFOLFOX6 regimen is commonly used and is acceptable in the first line setting to treat advanced esophagogastric adenocarcinoma and will be evaluated in combination with GS-5745.

1.4. Risk/Benefit Assessment for the Study

GS-5745 when administered at 800 mg IV every 2 weeks in combination with standard of care chemotherapy appears to be safe and well-tolerated. The 800 mg every 2 weeks regimen of GS-5745 IV is expected to achieve steady trough concentrations greater than 26 µg/mL in the majority of subjects, which is 260-fold over the IC_{50} (0.691 nM or 0.1 µg/mL) evaluated in enzymatic inhibitory activities of GS-5745 in vitro. There were no dose-limiting toxicities observed at monotherapy doses up to 1800 mg IV every 2 weeks. The AE profile observed in study GS-US-296-0101 is consistent with the profile typically seen in subjects with advanced cancers including pancreatic, gastric, or non-small cell lung adenocarcinoma receiving standard of care chemotherapy. Preliminary efficacy data indicates that the objective response rate is 58% (33 evaluable subjects, data on file as of 11 September 2015) compared to the historical response rate of approximately 40% in gastric cancer patients with measurable disease treated with mFOLFOX6.

For the treatment of advanced gastric cancer, a common regimen includes a fluoropyrimidine (5-FU or capecitabine) and platinum agent (cisplatin or oxaliplatin). There are no definitive studies demonstrating the superiority of either fluoropyrimine or platinum. In the EU, in line with ESMO recommendations clinical practice guidelines {Waddell 2013}, combination regimens based upon a platinum fluoropyrimidne doublet are generally used for the treatment of metastatic advanced gastric cancer.

There is no clear or demonstrably superior standard regimen for the treatment of advanced gastric cancer and few clinical trials addressing this target patient population. Furthermore, there are limited first-line therapy treatment options in the metastatic setting. In addition, prognosis is poor in patients with adenocarcinoma of the stomach diagnosed with Stage III or IV disease.

Thus, the favorable nonclinical and clinical data outweigh the risks associated with administration of GS-5745 and hence support the evaluation of GS-5745 in combination with mFOLFOX6 as a first-line treatment in patients with advanced gastric or gastroesophageal junction adenocarcinoma.

1.5. Compliance

The sponsor will conduct this study in compliance with this protocol, Good Clinical Practice (GCP), and all applicable regulatory requirements.

2. OBJECTIVES AND ENDPOINTS

2.1. Primary Objective

- To compare the efficacy of GS-5745 versus placebo in combination with mFOLFOX6 as measured by overall survival (OS)

2.2. Secondary Objectives

- To compare the efficacy of GS-5745 versus placebo in combination with mFOLFOX6 as measured by progression-free survival (PFS)
- To compare the efficacy of GS-5745 versus placebo in combination with mFOLFOX6 as measured by objective response rate (ORR) per Response Evaluation Criteria In Solid Tumors version 1.1 (RECIST v1.1)
- To compare the safety of GS-5745 versus placebo in combination with mFOLFOX6

2.3. Exploratory Objectives

2.4. Primary Endpoint

- Overall survival (OS) - defined as the time from date of randomization to death from any cause

2.5. Secondary and Exploratory Endpoints

The following secondary endpoints will be defined and analyzed in this study:

- Progression free survival (PFS) – defined as the time from randomization to the earlier of first documentation of definitive disease progression or death from any cause

- Objective response rate (ORR) – defined as the proportion of subjects who achieve a CR or PR as assessed by RECIST v1.1

- Safety measurements including incidence of adverse events, clinical relevant changes in laboratory values and vital signs.

The following exploratory endpoints will be defined and analyzed in this study:

3. STUDY DESIGN

3.1. Study Design

This is a Phase 3, randomized, double-blind, multicenter study of GS-5745 combined with mFOLFOX6 in subjects with untreated gastric and GEJ adenocarcinoma. A total of 430 eligible subjects with advanced gastric and GEJ cancer will be randomized in a 1:1 manner to mFOLFOX6 plus GS-5745 or mFOLFOX6 plus placebo. Treatment assignment will be stratified by ECOG status (0 or 1), geographic region (Latin America or other participating countries), and primary tumor site (gastric or GEJ).

Computed tomography (CT) or magnetic resonance imaging (MRI) scans will be performed every 8 weeks to evaluate response to treatment by RECIST v1.1.

Dosage and frequency will be as follows:

- mFOLFOX6 on Days 1 and 15 of each 28-day treatment cycle for a total of 6 cycles followed thereafter by leucovorin (LV) and 5-fluorouracil (5-FU) dosing on Days 1 and 15 of each 28-day treatment cycle until disease progression . The mFOLFOX6 dosing regimen will consist of *l*-LV 200 mg/m^2 or *dl*-LV 400 mg/m^2 and oxaliplatin 85 mg/m^2 followed by bolus 5-FU 400 mg/m^2 and a 46-hour infusion of 5-FU 2400 mg/m^2.

- GS-5745/placebo 800 mg every 2 weeks until disease progression

An independent data monitoring committee (DMC) will review the progress of the study and perform interim reviews of safety data. Safety review by the DMC will be performed after the equivalent of 4 treatment cycles from the time the 60th subject is randomized. Thereafter, review of safety data will be performed at regular intervals as described in the DMC charter. In addition, the DMC will meet after approximately 33.3% and approximately 66.7% of the expected number of events have occurred to review the results from the futility and efficacy interim analysis, respectively.

3.2. Study Treatments

Subjects will receive study drug (GS-5745/placebo) on Day 1 and Day 15 of each 28 day cycle of mFOLFOX6 via intravenous infusion over approximately 30 minutes. One cycle of mFOLFOX6 consists of 2 infusions. Administration of mFOLFOX6 will be immediately after administration of study drug as described above in Section 3.1.

3.3. Duration of Treatment

Each cycle will consist of 28 days and will continue in the absence of disease progression or unacceptable toxicity, consent withdrawal or subject's refusal of treatment. There will be a screening period of up to 28 days. Following completion of treatment, subjects will be followed for safety for 55 days and survival approximately every 3 months for up to 5 years.

3.4. Discontinuation Criteria from Study Treatment

GS-5745/Placebo and all of the components from mFOLFOX6 will be discontinued for any of the following reasons:

- Adverse events
- Pregnancy
- Investigator decision to remove the subject from the study treatment, in consultation with Gilead Medical Monitor
- Disease progression
- Intercurrent illness that would, in the judgment of the investigator, affect assessments of clinical status to a significant degree
- Initiation of non-study specific anti-neoplastic therapy in the absence of progression
- Subject request to discontinue treatment
- Withdrawal of consent.
  - A subject may withdraw consent solely from active participation in the study but still participate in follow up for disease progression and survival
- Death
- Discontinuation of the study at the request of Gilead, a regulatory agency, or an institutional review board or independent ethics committee (IRB/IEC)
- Lost to follow up

Should this occur, the corresponding study drug completion CRF should be entered to document the reason for discontinuation. In addition, the subject should continue with the rest of the treatment regimen and the study related procedures per protocol.

3.5. Premature Discontinuation from Study Treatment

If a subject has discontinued all study treatments prior to definitive disease progression, the subject shall remain on study until at least 1 of the criteria for discontinuation from study is met (Section 3.6). Every attempt should be made to keep the subject in the study and continue to perform tumor evaluation by CT or MRI every 8 weeks until disease progression. If this is not possible or acceptable to the subject or investigator, the subject may be withdrawn from the study. It is recommended that the investigator consults with the medical monitor prior to removing the subject from study for any reason except subject withdrawal of consent.

3.6. Discontinuation Criteria from Study

Subject study participation will be ended due to any of the following reasons, and subjects should then enter long-term follow-up if applicable:

- Initiation of non-study specific anti-neoplastic therapy in the absence of progression
- Disease progression
- Intercurrent illness that would, in the judgment of the investigator, affect assessments of clinical status to a significant degree
- Withdrawal of consent.
  - A subject may withdraw consent solely from active participation in the study but participate in follow up for survival
- Investigator decision to remove the subject from the study, in consultation with Gilead Medical Monitor
- Death
- Discontinuation of the study at the request of Gilead, a regulatory agency, or an IRB/IEC
- Lost to follow up

3.7. Long-Term Follow-Up for Overall Survival and Study Completion

After the End of Study (EOS), long-term follow-up (LTFU) will be initiated for subjects who discontinue from study due to reasons other than death. The subject shall remain on LTFU for OS until:

- Death
- Withdrawal of consent to participate in LTFU
- Lost to follow up
- End of LTFU period

Every attempt should be made to keep the subject in the LTFU for OS.

The end of the trial will be defined as when all subjects have completed LTFU or discontinued their participation in the study due to death, withdrawal of consent or lost to follow up.

3.8. Source Data

The subject identification number and randomization number captured by the interactive web response system (IWRS), as well as the patient reported outcomes data captured are considered source data.

3.9. Biomarker Testing

3.9.1. Biomarker Samples to Address the Study Objectives

The biological specimens described below will be collected in this study and will be used to evaluate the association of exploratory systemic and/or tissue specific biomarkers with study drug response including efficacy, as well as to increase knowledge and understanding of the mechanism of action of drug activity in human tumors, biology of resistance and highlight possible new combination opportunities. The specific analyses may include, but will not be limited to, the biomarkers and assays listed below. Because biomarker science is a rapidly evolving area of investigation, it is not possible to specify prospectively all tests that will be done on the specimens provided. Testing may be modified during or after the end of the study to remove tests no longer indicated and/or to add new tests based upon the growing state of the art knowledge.

These samples will be destroyed no later than 10 years, or as per local regulations, after the end of study unless the subject gives specific consent for the remainder of the samples to be stored for optional future research. If the patient provides consent for optional future research, the samples will be destroyed no later than 15 years after the end of the study, or as per local regulations.

3.9.1.1. Pharmacodynamic

CCI [redacted]

3.9.1.2. CCI [redacted]

[redacted]

Table 3-1. Biomarker Objectives and Testing

SAMPLE TYPE	OBJECTIVE	TEST
Blood	To evaluate pharmacodynamics of GS-5745	MMP9 cleavage products (eg, C1M, C4M) Circulating MMP9 protein
	To evaluate the effect of GS-5745 and chemotherapy on inflammation	Circulating cytokines and inflammatory markers (eg, interleukin-8, interleukin-2 receptor, interleukin 6, interferon gamma, etc)
	To evaluate the effect of GS-5745 and chemotherapy on circulating immune cells	Immune monitoring assay
	To evaluate other biomarkers of GS-5745 activity	Circulating growth factors (eg, vascular endothelial growth factor , epidermal growth factor etc)
	To evaluate disease burden and identify mutations correlated with resistance to therapy	Circulating tumor DNA isolation and sequencing for disease-specific markers (eg, mKRAS, mEGFR, mPIK3CA etc) and other genes
Tissue Biopsy: Archival Tissue	To evaluate baseline markers that correlate with response	MMP9, other MMPs and Immune Cells by immunohistochemistry (IHC) Gene expression patterns (RNA) DNA mutations may be evaluated
Tissue Biopsy: At progression (if medically feasible)	To evaluate markers of response or resistance	MMP9, other MMPs and Immune Cells by IHC Gene expression patterns (RNA) and DNA mutations may be evaluated and compared with data from the archival biopsy

3.9.2. Optional Blood Sample for Future Genomic Analysis

CCI [redacted]

3.9.3. Biologic Samples for Optional Future Research

CCI

[redacted]

4. SUBJECT POPULATION

4.1. Number of Subjects and Subject Selection

Approximately 430 subjects will be randomized to receive double-blind study drug. The target population is subjects with previously untreated advanced gastric or GEJ adenocarcinoma.

4.2. Inclusion Criteria

Subjects must meet all of the following inclusion criteria to be eligible for participation in this study.

1) Male or female ≥18 years of age

2) Histologically confirmed adenocarcinoma of the stomach or GEJ with inoperable, locally advanced or metastatic disease, not amenable to curative therapy

   Adenocarcinoma of the GEJ is defined as tumors that have their center within 5 cm proximal and distal of the anatomical esophagogastric junction as described in Siewert's classification system

3) Eastern Cooperative Oncology Group (ECOG) ≤ 1

4) Measurable disease or non-measurable but evaluable disease, according to RECIST v1.1. Subjects with peritoneal disease would generally be regarded as having evaluable disease and allowed to enter the trial

5) Subjects not receiving anticoagulant medication must have an international normalized ratio (INR) ≤ 1.5 and activated partial thromboplastin time (aPTT) ≤ 1.5 X upper limit of normal (ULN)

   The use of full-dose oral or parenteral anticoagulants is permitted as long as the INR or aPTT is within therapeutic limits (according to the medical standard in the institution) and the subject has been on stable dose of anticoagulants for at least 1 week at the time of randomization.

6) Adequate hematologic function:

   a) neutrophils ≥ 2.0 x 10^9/L

   b) platelets ≥ 100 x 10^9/L

   c) hemoglobin ≥ 9 g/dL

7) Adequate hepatic function:

   a) Direct or total bilirubin ≤ 1.5 x ULN.

   b) ALT and AST ≤ 2.5 x ULN, in case of liver metastases ≤ 5 x ULN

8) Creatinine clearance (CL_{cr}) should be ≥ 30 mL/min based on the Cockroft -Gault formula. Subjects with a CL_{cr} just below 30 mL/min may be eligible if a measured CL_{cr} (based on 24 hour urine collection or other reliable method) is ≥ 30 mL/min

9) For female subjects of childbearing potential, willingness to use a protocol-recommended method of contraception from the screening visit throughout the study treatment period, for 90 days following the last dose of study drug (GS-5745/placebo), and for 4 months after the last dose of oxaliplatin or 6 months after the last dose of 5-FU whichever occurs later unless the subject chooses continuous heterosexual abstinence as a lifestyle-choice (see Appendix 3 for more information)

10) For male subjects of reproductive potential, willingness to use a protocol-recommended method of contraception and to refrain from sperm donation from the start of study drug, throughout the study treatment period, for 90 days after administration of the last dose of any study drug, and for 6 months after the last dose of oxaliplatin or 6 months following the last dose of 5-FU whichever occurs later (see Appendix 3)

11) Breastfeeding females must agree to discontinue nursing before study drug administration

12) In the judgment of the investigator, participation in the protocol offers an acceptable benefit-to-risk ratio when considering current disease status, medical condition, and the potential benefits and risks of alternative treatments for the subject's cancer

13) Willingness to comply with scheduled visits, drug administration plan, imaging studies, laboratory tests, other study procedures, and study restrictions

14) Evidence of a signed informed consent prior to implementation of any protocol specific procedure

4.3. Exclusion Criteria

Subjects who meet *any* of the following exclusion criteria are not to be randomized in this study.

1) Previous chemotherapy for locally advanced or metastatic gastric or GEJ cancer. Subjects may have received prior neoadjuvant or adjuvant chemotherapy as long as it was completed at least 6 months prior to randomization

2) Human Epidermal Growth Factor Receptor 2 (HER2)-positive gastric cancer (primary tumor or metastatic lesion). HER2-positivity is defined as either IHC3+ or IHC2+/ISH+ (ISH positivity is defined as a HER2:CEP17 ratio of ≥2.0.)

3) Patients who have received palliative radiation and have not recovered from all acute, reversible effects.

4) Uncontrolled intercurrent illness including, but not limited to, active uncontrolled infection, active gastrointestinal bleeding, uncontrolled cardiac arrhythmia, or psychiatric illness/social situation that would limit compliance with study requirements as judged by treating physician

5) History of a concurrent or second malignancy except for adequately treated local basal cell or squamous cell carcinoma of the skin, cervical carcinoma in situ, superficial bladder cancer, asymptomatic prostate cancer without known metastatic disease and with no requirement for therapy or requiring only hormonal therapy and with normal prostate-specific antigen for ≥ 1 year prior to randomization, adequately treated Stage 1 or 2 cancer currently in complete remission, or any other cancer that has been in complete remission for ≥ 5 years

6) Major surgery, defined as any surgical procedure that involves general anesthesia and a significant incision (ie, larger than what is required for placement of central venous access, percutaneous feeding tube, or biopsy), within 28 days of first dose of study drug

7) Known positive status for human immunodeficiency virus (HIV)

8) Known acute or chronic-active infection with hepatitis B virus (HBV) or hepatitis C virus (HCV)

9) Peripheral neuropathy ≥ Grade 2 according to National Cancer Institute Common Terminology Criteria for Adverse Events version 4.03 (NCI CTCAE v.4.03)

10) Chronic daily treatment with oral corticosteroids (dose of > 10 mg/day methylprednisolone equivalent). Inhaled steroids and short courses of oral steroids for anti-emesis or as an appetite stimulant are allowed

11) Pregnant or breastfeeding women (pregnancy needs to be excluded by testing of beta-human chorionic gonadotropin [β-hCG])

12) Known or suspected central nervous system metastases

13) Known dihydropyrimidine dehydrogenase-deficiency (special screening not required)

14) Known alcohol or drug abuse or any other medical or psychiatric condition which contraindicates participation in the study

15) Documented myocardial infarction or unstable/uncontrolled cardiac disease (ie, unstable angina, congestive heart failure [New York Heart Association > Class II]) within 6 months of randomization

16) Active tuberculosis or history of latent tuberculosis that has not been treated

17) Any chronic medical condition that, in the opinion of the Investigator, would make the subject unsuitable for the study or would prevent compliance with the study protocol.

18) Serious systemic fungal, bacterial, viral, or other infection that is not controlled or requires intravenous antibiotics

19) Experimental medical treatment within 28 days prior to randomization

20) Known hypersensitivity to any of the study drugs or excipients or to Chinese hamster ovary cell products or to recombinant human or humanized antibodies**,** or known allergic reactions to products that contain platinum

21) History of long QT syndrome or whose corrected QT interval (QTc) measured using Fridericia's formula ($QTcF = QT/RR^{0.333}$) at screening is prolonged (> 450 ms for males and > 470 ms for females)

22) Subjects with potassium, magnesium or calcium less than the lower limit of normal (LLN); electrolyte replacement is permitted during screening

5. INVESTIGATIONAL MEDICINAL PRODUCTS

5.1. Randomization, Blinding and Treatment Codes

Randomization to GS-5745 or placebo will be based on a randomization schedule prepared by Gilead and/or a designee before the start of the study. Eligible subjects will be randomized via an interactive web response system (IWRS).As described in Section 8.5.1, randomization will be stratified by ECOG status (0 or 1) at screening, geographic region (Latin America or other participating countries), and primary tumor site (gastric or GEJ).

The IWRS will be used to maintain a central log documenting screening, to implement randomization, to assess current inventories of study drug, to initiate any necessary resupply of study drug, and to document discontinuation of the study drug.

The IWRS will assign kit numbers and provide instructions for dispensing of blinded study drug (GS-5745/placebo). It is anticipated that subjects will usually begin study drug immediately after randomization.

5.1.1. Procedures for Breaking Treatment Codes

In the event of a medical emergency where breaking the blind is required to provide medical care to the subject, the investigator may obtain treatment assignment directly from the IWRS system for that subject. Gilead recommends but does not require that the investigator contact the Gilead medical monitor before breaking the blind. Treatment assignment should remain blinded unless that knowledge is necessary to determine subject emergency medical care. The rationale for unblinding must be clearly explained in source documentation and on the electronic case report form (eCRF), along with the date on which the treatment assignment was obtained. The investigator is requested to contact the Gilead medical monitor promptly in case of any treatment unblinding.

Blinding of study treatment is critical to the integrity of this clinical trial and therefore, if a subject's treatment assignment is disclosed to the investigator, the subject will have study treatment discontinued. All subjects will be followed until study completion unless consent to do so is specifically withdrawn by the subject.

Gilead Drug Safety and Public Health (DSPH) may independently unblind cases for expedited reporting of suspected unexpected serious adverse reactions (SUSARs).

5.2. Description and Handling of Study Treatments

5.2.1. Formulation

5.2.1.1. GS-5745

GS-5745 is formulated as a sterile, aqueous buffered solution containing acetate at pH 5.0, with sucrose and polysorbate 20 added for stabilization. Each 10 mL vial contains 400 mg GS-5745 at a concentration of 40 mg/mL.

5.2.1.2. Placebo to Match GS-5745

Placebo to match GS-5745 is formulated as a sterile, aqueous buffered solution containing acetate at pH 5.0, with sucrose and polysorbate 20 added for stabilization in a 10 mL vial.

5.2.1.3. mFOLFOX6

The mFOLFOX6 dosing regimen will consist of *dl*-LVor *l*-LV, oxaliplatin and 5-FU.

5.2.1.3.1. *dl*-Leucovorin and *l*-Leucovorin

dl-Leucovorin and *l*-LV are commercially sourced. Information regarding the formulation can be found in the current prescribing information.

5.2.1.3.2. Oxaliplatin

Oxaliplatin is commercially sourced. Information regarding the formulation can be found in the current prescribing information

5.2.1.3.3. 5-Fluorouracil

5-Fluorouracil is commercially sourced. Information regarding the formulation can be found in the current prescribing information.

5.2.2. Packaging and Labeling

Study drug (GS-5745/placebo) solution will be supplied in 10 mL glass vials with coated elastomeric stoppers and aluminum crimp overseals with a flip-off cap.

Study drug(s) to be distributed to centers in the US and other participating countries shall be labeled to meet applicable requirements of the United States Food and Drug Administration (FDA), European Union (EU) guideline to Good Manufacturing Practice - Annex 13 (investigational medicinal Product(s)), and/or other local requirements. Gilead or designated distribution depots will distribute study drug vials to sites as per current Good Manufacturing Practices (cGMP) requirements.

mFOLFOX6 will be supplied by Gilead per local country regulations.

5.2.3. Storage and Handling

Study drug (GS-5745/placebo) should be stored at 2 - 8 ºC. Storage conditions are specified on the study drug label. Until dispensed to the subject, all study drug should be stored in a securely locked area, accessible only to authorized site personnel. To ensure stability and proper identification, the study drug should be stored in the containers in which they were supplied until dosing to the subject.

Components of mFOLFOX6 regimen are commercially sourced. Information regarding the storage and handling of *dl*-LV/*l*-LV, oxaliplatin and 5-FU can be found in the current prescribing information

5.3. Dosage and Administration of Study Drug and mFOLFOX6

Study drug (GS-5745/placebo) will be administered 800 mg via intravenous infusion over approximately 30 (± 5) minutes at the research clinic by a qualified staff member on Days 1 and 15 of each 28-day treatment cycle of mFOLFOX6. One cycle of mFOLFOX6 consists of 2 infusions. mFOLFOX6 will be administered immediately following study drug on Days 1 and 15 of each cycle as well. After 6 cycles of mFOLFOX6, the oxaliplatin component will be discontinued. Leucovorin and 5-FU will be continued as maintenance therapy, in combination with GS-5745/placebo, in the absence of disease progression or toxicity warranting discontinuation of therapy. The mFOLFOX6 dosing regimen will consist of *l*-LV 200 mg/m^2 or *dl*-LV 400 mg/m^2 and oxaliplatin 85 mg/m^2 followed by bolus 5-FU 400 mg/m^2 and a 46-hour infusion of 5-FU 2400 mg/m^2. Minor modifications to the duration of the infusion time are permitted as per institutional standard. Adjustments to the dose of mFOLFOX6 are permitted in response to treatment emergent adverse events. Refer to the local prescribing information for care of subjects including contraindications, subject monitoring, and the medicinal products prohibited or to be used with care for all the components of mFOLFOX6.

The Investigator or a qualified designee must be present during the administration of GS-5745/placebo. Subjects should be observed following end of infusion and discharged at the discretion of the Investigator or qualified designee.

Documentation of the study drug and mFOLFOX6 administration will be noted on the eCRF and in the source documentation.

5.3.1. Dose Interruption and Reduction

If an adverse event is attributed to only 1 drug (ie, GS-5745/placebo or 1 or more component of mFOLFOX6), the investigator's discretion will be used to determine if the drug(s) not attributed to the adverse event will be withheld based on the investigator's assessment of risk-benefit of withholding 1 or more drugs (please see sections 5.3.1.1 and 5.3.1.2 for additional guidance).

Careful monitoring must be ensured for subjects with a history of allergic reactions to other platinum based chemotherapy agents. In the event of anaphylactic shock symptoms, interrupt infusion immediately and initiate appropriate treatment. Resumption of oxaliplatin treatment following anaphylactic reaction is contraindicated.

Clinically significant, abnormal 12-lead safety ECGs should be repeated. Subjects who have 2 consecutive ECGs showing a new absolute QTc duration > 500 msec, or a QT/QTc > 60 msec over the corresponding baseline value must discontinue any medications that could prolong the QT interval (including oxaliplatin). Subject's concomitant medications should be reviewed to determine a potential etiology for the ECG changes. Appropriate intervention (ie cardiology evaluation, telemetry monitoring, management of electrolyte abnormalities) in response to treatment emergent QT interval prolongation should be initiated.

5.3.1.1. GS-5745/Placebo (Study Drug)

If a subject experiences a Grade 3 or greater toxicity considered secondary to study drug (GS-5745/placebo), treatment will be postponed until the toxicity returns to Grade 0-1 (as defined by NCI CTCAE v 4.03) or returns to the subject's baseline value. If the toxicity returns to Grade 0-1 or the subject's baseline value, the subject may resume GS-5745 at the originally assigned dose level.

If the event does not resolve to Grade 0-1 or the subject's baseline within 28 days, treatment with study drug (GS-5745/placebo) must be permanently discontinued. If the subject experiences a recurrence of the Grade 3 or greater toxicity after restarting study drug, treatment with GS-5745/placebo will be discontinued. Subjects who permanently discontinue treatment with study drug (GS-5745/placebo) may continue treatment with mFOLFOX6 after consultation with the medical monitor.

Subjects who are not receiving any chemotherapy (mFOLFOX6) should not receive study drug (GS-5745/placebo) as monotherapy. Investigators should contact the Gilead Medical Monitor with any questions regarding study drug dose modification, interruption or discontinuation.

5.3.1.2. mFOLFOX6

Recommended dose reduction for the components of mFOLFOX6 is described in Table 5-1 and is based on the adverse event (AE) table described in Appendix 5. Sites may also follow their institutional practice for dose reductions. Leucovorin doses may be adjusted per institutional guidelines in the event of a supply shortage.

Subjects who interrupt all components of mFOLFOX6 for greater than 28 days secondary to drug toxicity(s) that do not resolve should permanently discontinue all study treatment.

Subjects who permanently discontinue all components of mFOLFOX6 should also discontinue GS-5745/placebo.

Investigators should contact the Gilead Medical Monitor with any questions regarding study drug dose modification, interruption or discontinuation.

Table 5-1. Dose Reduction Levels[a] for mFOLFOX6

	Dose Level		
Drug	**Starting Dose**	**-1**	**-2 [b]**
Oxaliplatin	85 mg/m^2	65 mg/m^2	50 mg/m^2
5-FU bolus	400 mg/m^2	OMIT	OMIT
5-FU continuous infusion over 46-48 hours	2400 mg/m^2	1900 mg/m^2	1500 mg/m^2
dl-Leucovorin/l-Leucovorin[c]	400/200 mg/m^2	100%	100%

a If an AE is believed likely to be due to 1 drug, it is permissible to decrease dose of that drug only.
b Further dose levels (-3, -4, etc.) will be 20% dose reductions from the previous level for oxaliplatin and 5-FU continuous infusion. In addition, the bolus dose of 5-FU will continue to be omitted, and the leucovorin dose will remain unadjusted (100%).
c Dosing of leucovorin will remain fixed at 100% of recommended dose.

5.4. Prior and Concomitant Medications

During the course of the clinical trial, study subjects are anticipated to continue the use of prescribed medications identified during the screening procedures, consistent with study inclusion and exclusion criteria.

Non-study anticancer chemotherapy or immunotherapy (approved or investigational) are not permitted during the trial. If administered, the subject may be removed from the trial.

Concomitant medications that are known to lengthen QTc interval must be used carefully while the subject is receiving oxaliplatin. Please refer to https://www.crediblemeds.org/ for a listing of such medications.

5.5. Accountability for Study Drug

The investigator or designee (eg, pharmacist) is responsible for ensuring adequate accountability of all used and unused investigational medicinal product during the study. This includes acknowledgement of receipt of each shipment of study drug (quantity and condition) and tracking of vials assigned/utilized for subject dosing.

Study drug (GS-5745/placebo) accountability records will be provided to each study site to:

- Record the date received and quantity of study drug vials
- Record the date, subject number, subject initials, the vial number dispensed
- Record the date, quantity of used and unused vials returned, along with the initials of the person recording the information.

Dispensing records will include the initials of the person dispensing the study drug or supplies.

5.5.1. Investigational Medicinal Product & mFOLFOX6 Return or Disposal

The study drug and mFOLFOX6 should be disposed of at the site as per local standard operating procedures. Please see Section 9.1.7 for additional instructions.

6. STUDY PROCEDURES

The study procedures to be conducted for each subject enrolled in the study are presented in tabular form in Appendix 2 and described in the text that follows.

The investigator must document any deviation from protocol procedures and notify the sponsor or contract research organization (CRO).

Safety and tolerability assessments will include regular monitoring of AEs, changes from baseline in laboratory variables, physical examinations, vital signs, and special safety assessment like ECGs.

From the time of obtaining informed consent through the first administration of study drug, record all serious adverse events (SAEs), as well as any non-serious AEs related to protocol-mandated procedures on the AEs eCRF. All other untoward medical occurrences observed during the Screening period, including exacerbation or changes in medical history are to be captured on the medical history eCRF.

6.1. Subject Enrollment and Treatment Assignment

Subject eligibility will be established at the conclusion of the screening evaluations. The screening number and/or subject ID will be assigned for that individual subject by the designated IWRS. Subject eligibility must be determined by results received from the central lab with exceptions made for re-screening and human epidermal growth factor receptor 2 (HER2) testing as described in Sections 6.2.2 and 6.2.8.2.

It is the responsibility of the investigator to ensure that each subject is eligible for the study before randomization. A subject will be considered enrolled once he or she has completed randomization.

Subjects in both treatment arms will undergo all the same procedures. Details regarding randomization and treatment assignment are in Section 5.

6.2. Study Procedure Descriptions

The sections below describe the individual study procedures outlined in subsequent sections and the schedule of assessments. During the treatment period, all visits may be performed within the windows identified in Appendix 2.

6.2.1. Informed Consent

All subjects must sign and date the most recent IRB/IEC-approved informed consent form before any study specific procedures are performed except where noted in the protocol in relation to standard of care procedures.CCI [redacted]

6.2.2. Re-Screening Criteria

Subjects who do not randomize within 28 days of screening will be screen failed. The screening period may be extended beyond 28 days with sponsor approval if the only outstanding eligibility criterion is the human epidermal growth factor receptor 2 (HER2) status.

Re-screening may be allowed. Subjects who are re-screened after 30 days, must be re-consented with a new screening number, and repeat the screening assessments. For subjects re-screened within 30 days, assessments with results that would exclude the subject will need to be repeated. Subjects with abnormal electrolyte values during screening will be allowed to repeat testing for the purposes of study eligibility - replacement therapy is permitted. If this is the only inclusion/exclusion criteria not met, these analytes may be retested (locally or centrally) so that values above the lower limit of normal could be achieved without the need to screen fail the subject.

Subject eligibility must be determined by results received from the central lab. However, if there have been 2 failed attempts to obtain test results from the central lab, eligibility may be determined using local lab results, with documentation of failed attempts, local lab results, and sponsor approval.

6.2.3. Medical & Medication History

A complete medical and surgical history will be obtained by the investigator or designee at screening, including disease history, and recorded on the eCRF.

All medications taken within 30 days prior to screening and during the screening period will be obtained prior to randomization and recorded on the eCRF. At each study visit, the site will capture any and all medications taken by the subject since the last visit or during the visit (as applicable). Concomitant medications include prescription and non-prescription medications, vitamins and minerals.

In addition, supportive therapies given during the course of the study (eg, blood transfusion, growth factor) should be collected and recorded on the eCRF.

6.2.4. Physical Examination

A physical examination (PE) will be performed at screening, end of treatment (EOT) and EOS. This will include assessment of clinical signs and symptoms. The exam will be performed by a physician, a physician's assistant, or nurse practitioner qualified to perform assessments. Breast, genital, and rectal examinations are not required, unless warranted in opinion of the healthcare provider.

A modified physical exam capturing changes from prior exams will be performed at Day 1 of each cycle and at the 30-day Safety Follow-Up. Height will be collected at Screening only and entered in the CRFs in centimeters. Body surface area (BSA) to determine the dose for mFOLFOX6 will be calculated using height and weight; the Mosteller formula: $\sqrt{}$([Height (cm) ×Weight (kg)]/3600) is preferred. However, institutional guidelines/practice for calculating BSA will also be allowed.

6.2.5. Vital Signs & Weight

Vital signs ie, blood pressure, heart rate, respiratory rate, and oral temperature, will be measured by the investigator or qualified designee as per standard institutional guidelines at each study visit as indicated in Appendix 2. Weight will be collected at the same visits vital signs are taken and entered in the CRFs in kilograms. BSA to determine the dose for mFOLFOX6 will be calculated using weight in kilograms. The Mostellar formula is preferred for calculating BSA. However, institutional guidelines/practice for calculating BSA will also be allowed.

6.2.6. Electrocardiogram Assessment

A single 12-lead electrocardiogram (ECG) will be collected at Screening, Day 1 of each cycle, at the EOT and EOS visits, per standard practice. Additional ECGs following administration of oxaliplatin will be collected at visits when oxaliplatin is administered after the infusion is completed (i.e., Days 1 and 15 of each cycle, up to a maximum of six 28-day cycles) as required per country label guidance regarding cardiac monitoring. The investigator will review all ECGs and retain the tracing with the source documents. ECGs obtained after oxaliplatin infusions must be reviewed by the investigator/sub investigator before subjects leave the clinic/infusion center.

6.2.7. Performance Status

Performance status will be scored using the ECOG performance status scale index (refer to Appendix 4), at Screening, Day 1 of each cycle, EOT, EOS and 30-day Safety Follow-up visits. ECOG used to determine eligibility must be the performance status during the screening period. ECOG performance status on Cycle 1 Day 1 may be waived if it was conducted during screening within 4 days of Cycle 1 Day 1.

6.2.8. Laboratory Assessments

The central laboratory will be responsible for chemistry, hematology, coagulation, urinalysis, and serum pregnancy testing (per Table 6-1) as well as processing and/or storage of other study samples. Specific instructions for processing, labeling, and shipping samples will be provided in a central laboratory manual. The date and time of sample collection will be reported to the central laboratory.

If central laboratory results are not available, local laboratories may be used for dosing decisions. Local laboratory assessments resulting in a dose change or as part of an adverse event assessment, which is not supported by central lab results, will be reported on the eCRF.

Urine pregnancy test will be performed at the site.

Table 6-1. Analytes

Chemistry	Urinalysis	Hematology	Other
Albumin Alkaline phosphatase ALT AST Bicarbonate BUN Calcium Chloride Creatinine[a] Glucose Lipase Magnesium Phosphorus Potassium Sodium Total bilirubin Direct bilirubin Total protein CA19-9 CEA	Color and appearance Specific gravity pH Occult blood Protein Glucose Bilirubin Leukocyte esterase Nitrite Urobilinogen Ketones Microscopic[b]	WBC Hemoglobin Hematocrit Platelet ANC Differential Eosinophils Lymphocytes Monocytes Neutrophils **Coagulation** PT/INR aPTT	Serum β-hCG or urine pregnancy test[c]

ANC = absolute neutrophil count; ALT = alanine aminotransferase; aPTT = Activated Partial Thromboplastin Time; AST = aspartate aminotransferase; BUN = blood urea nitrogen; CEA=carcinoembryonic antigen; β-hCG = beta-human chorionic gonadotropin; HER2 = human epidermal growth factor receptor 2; INR = International Normalized Ratio; PT = Prothrombin Time; SGOT = Serum Glutamic-Oxaloacetic Transaminase; SGPT = Serum Glutamic-Pyruvic Transaminase; WBC = white blood cell

a Estimated creatinine clearance (CL_{cr})/glomerular filtration rate will be calculated based on the Cockroft-Gault formula using actual body weight: CL_{cr} (mL/min) = (140 – age [years])* weight (kg) / (serum creatinine [mg/dL]*72). If the subject is female, multiply the quantity by 0.85.

b Reflex testing based on other abnormalities

c Females of child-bearing potential only. Serum pregnancy will be conducted at Screening. Urine pregnancy will be conducted pre-dose on Day 1 of each cycle

Screening laboratory samples should be obtained within 28 days prior to randomization. Blood samples will be obtained for hematology, chemistry, coagulation, and pregnancy testing for female subjects. A urine sample will also be obtained at screening for urinalysis.

Blood samples for hematology and chemistry will be obtained at pre-dose of Days 1 and 15 of each cycle, EOT, EOS, and 30-day Safety Follow-up. A urine sample for urinalysis and pregnancy testing for female subjects will also be obtained at Day 1 of each cycle, EOT, EOS, and 30-day Safety Follow-up (urinalysis only). Blood samples for coagulation will be obtained at EOT and EOS.

At any time during the study, abnormal laboratory parameters that are clinically relevant (eg, lead to clinical symptoms or signs, require therapeutic intervention), and constitute an AE must be recorded in the eCRF.

6.2.8.1. Pregnancy Test

All females of childbearing potential (see Appendix 3) will have a serum pregnancy test at screening. Urine pregnancy tests will be performed on Day 1 of every cycle, EOT, EOS, and 30-day Safety Follow-up.

6.2.8.2. Human Epidermal Growth Factor Receptor 2 (HER2) Testing

Prior to randomization, the subject's tumor should have been tested for HER2 status with approved immunohistochemistry (IHC) and in situ hybridization (ISH) kits. HER2 positivity is defined as IHC3+ or IHC2+/ISH+ (ISH positivity is defined as HER2:CEP17 ratio of ≥ 2.0). Results for HER2 status obtained prior to signing informed consent are acceptable if obtained with approved IHC and ISH kits. HER2 status may be determined during screening by testing the tumor at the central lab or a local lab with approved IHC and ISH kits.

6.2.8.3. Archival Tumor Tissue

Archival tumor tissue formalin-fixed paraffin embedded (FFPE) blocks will be collected from all subjects for biomarker analysis and shipped to central laboratory for sectioning after Day 1 of Cycle 1. If FFPE blocks are not available, unstained slides are also acceptable (see Covance manual for details). If an archival block was submitted for HER2 testing, and the subject is randomized, the block will be stored for biomarker analysis. However, if slides were submitted for HER2 testing, additional slides will be required after Day 1 of Cycle 1 for biomarker analysis.

6.2.8.4. Anti-GS-5745 Antibody

Blood samples for anti-GS-5745 antibody will be collected prior to dosing on Day 1 of cycles 1, 2, 3, 5, 7, and every 3 cycles thereafter, at the EOT, EOS, and 30-Day Safety Follow Up visits.

6.2.8.5. GS-5745 Pharmacokinetics

Blood plasma samples will be collected for GS-5745 PK at 30 (± 15) minutes after the end of infusion on Cycle 1 Day 1, prior to dosing, and 30 (± 15) minutes after the end of infusion on Day 1 of Cycles 2, 3, 5, 7, and every 3 cycles thereafter, at the EOT and EOS visits.

6.2.8.6. Optional PK Sub-study

CCI [redacted]

CCI

6.2.8.7. Biomarkers

Samples for biomarker analysis as listed on Table 3-1 will be collected as specified below and in Appendix 2:

- Blood biomarker samples will be collected prior to dosing on Cycle 1 Day 1, Cycle 1 Day 15, and Day 1 of Cycles 2 and 4, and every 2 cycles thereafter.
- A pharmacodynamic biomarker sample and immune-monitoring sample will be collected prior to dosing on Cycle 1 Day 1 and Cycle 3 Day 1.
- A biomarker sample for circulating tumor DNA will be collected prior to dosing on Cycle 1 Day 1, Cycle 3 Day 1, and at disease progression.
- CCI

6.2.8.8. Core Biopsy

At disease progression, a core biopsy will be collected (fine needle aspirate is not acceptable) if medically feasible. The sample should be collected by the last clinic visit on study, ie, EOS or 30-Day Safety Follow-up visit.

6.2.9. Disease and Response Assessment

6.2.9.1. Tumor Imaging

Either contrast-enhanced CT or gadolinium-enhanced MRI of the chest, abdomen, and pelvis will be performed at screening, every 8 weeks during the study (starting from Cycle 1 Day 1) and at the EOS visit if one has not been performed within the last 8 weeks. Tumor burden will be evaluated solely based on radiographic imaging per RECIST v 1.1. Chest x-ray, ultrasound, endoscopy, laparoscopy, positron-emission tomography, radionuclide scans, or tumor markers will not be considered for response assessment.

For radiographic evaluations, the same method of assessment and the same technique (eg, scan type, scanner, subject position, dose of contrast, injection/scan interval) should be used to characterize each identified and reported lesion at baseline and during study treatment and follow-up.

Scans taken as part of standard medical practice up to 42 days prior to randomization can be used for Screening as long as they meet all study requirements. During the treatment phase, scans may be performed at time points other than every 8 weeks as clinically indicated to assess tumor progression.

For subjects who stop study treatment in the absence of disease progression (eg, experienced unexpected toxicity), scans should continue to be collected approximately every 8 weeks until disease progression or initiation of systemic anti-tumor therapy other than the study treatment, whichever is earlier.

All relevant clinical and radiographic information required to make each assessment must be made available for source verification and submission to a central reader. Scans will be transferred to a central reader for collection and future analysis. Disease progression will be determined by the investigator or qualified designee.

6.2.10. Patient-reported Outcomes Assessments

Patient-reported outcomes (PRO) will be collected at Day 1 of each cycle, EOT and EOS visits, and at the 30 day safety follow up visit prior to any assessments in the clinic.

6.2.10.1. EuroQol-5D

The EuroQol-5D (EQ-5D) is a general health quality of life self-report instrument that assesses 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each dimension has 3 levels (EQ-5D-3L), ranging from no health problem, moderate health problem(s), and extreme health problem(s). It also includes a single visual analog scale for assessment of current general health.

6.2.10.2. Quality of Life Questionnaire

The European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire (QLQ-C30) is a self-report questionnaire consisting of 30 questions assessing 15 sections of the subject's wellbeing. The accompanying module for gastric cancer, QLQ-STO22, has an additional 22 questions.

6.2.11. Study Drug and Chemotherapy Administration

Study drug (GS-5745/placebo) will be dosed on Days 1 and 15 of every 28-day cycle via intravenous infusion over approximately 30 (± 5) minutes followed by chemotherapy (mFOLFOX6) as described in Section 5. One cycle of mFOLFOX6 consists of 2 infusions.

6.2.12. Adverse Events

From the time of obtaining informed consent through the first administration of study drug, record all SAEs as well as any AEs related to protocol-mandated procedures on the AE eCRF. All other untoward medical occurrences observed during the screening period, including exacerbation or changes in medical history are to be captured on the medical history eCRF.

From the time of the first administration of study drug through 55 days post the last administration of treatment (GS-5745/placebo and mFOLFOX6), record any SAEs, and AEs including exacerbation or changes in medical history, on the AE eCRF.

See Section 7 Adverse Events and Toxicity Management for additional details.

6.3. Assessments for Premature Discontinuation from Study

If a subject has discontinued all study treatments prior to definitive disease progression, the subject shall remain on study for follow-up for progression-free survival (see Sections 3.5 and 3.6). Every attempt should be made to keep the subject in the study and continue to perform tumor evaluation by CT or MRI every 8 weeks. Subjects will remain on study until disease progression or initiation of non-study specific anti-neoplastic therapy in the absence of progression, whichever occurs earlier.

If it is not possible to keep the subject on study, or acceptable to the subject or investigator, the subject may be withdrawn. It is recommended that the investigator consults with the medical monitor prior to removing the subject from study for any reason except subject withdrawal of consent.

6.4. Criteria for Discontinuation of Study Treatment

See Sections 3.4 and 3.6 for discontinuation criteria.

6.5. End of Treatment

End of treatment (EOT) assessments will be completed only by subjects who discontinue all treatment prior to disease progression. These assessments should be completed as soon as possible after the decision is made. Every attempt should be made to keep the subject in the study and continue to perform tumor evaluation by CT or MRI every 8 weeks until disease progression.

6.6. End of Study

End of study (EOS) assessments will be completed when the subject meets at least 1 of the criteria for study discontinuation (Section 3.6).

6.7. Safety Follow-Up

A follow-up visit will be performed 30 days (± 7 days) following the last dose of GS-5745/placebo or all components of mFOLFOX6, whichever is discontinued later. A follow-up phone call will be performed 55 days (± 7 days) following the last dose of GS-5745/placebo or all components of mFOLFOX6, whichever is discontinued later to assess any AEs and concomitant mediations.

6.8. Long-Term Follow-up

Long-term follow-up (LTFU) for overall survival begins after the EOS visit, or the last visit on study if EOS does not occur. Subjects will be contacted via phone call every 3 months for determination of long term survival status and record of any other anti-cancer therapy, cancer related surgery for up to 5 years after the EOS visit.

Subjects who are not deceased by the time Gilead has made the determination the study will be ended will receive a final follow-up phone call to assess survival status and communicate the Sponsor's decision.

The investigator will make every effort to contact the subject or a close relative or caretaker by phone to collect survival information. The investigator should show due diligence by documenting in the source documents steps taken to contact the subject (ie, dates of phone calls, registered letters, etc).

See Section 3.7 for reasons for discontinuing long-term follow-up and study completion.

6.9. Unscheduled visits

Unscheduled visits may occur at any time while the subject is enrolled on study. Vital signs, laboratory assessments, ECG, and physical examination may be conducted at these visits. Data generated during an unscheduled visit will be collected on the eCRF.

6.10. Protocol Deviations

Gilead's policy prohibits exemptions from protocol inclusion/exclusion criteria. In the event of a significant deviation related to gross non-compliance from the protocol or incidences that impose significant risk to subject safety, the investigator or designee must notify Gilead and/or its designee immediately. The site will be required to document deviations in accordance with Gilead's procedures and in accordance with the site's procedures and processes.

7. ADVERSE EVENTS AND TOXICITY MANAGEMENT

7.1. Definitions of Adverse Events, Adverse Reactions, and Serious Adverse Events

7.1.1. Adverse Events

An AE is any untoward medical occurrence in a clinical study subject administered a medicinal product, which does not necessarily have a causal relationship with the treatment. An AE can therefore be any unfavorable and/or unintended sign, symptom, or disease temporally associated with the use of a medicinal product, whether or not considered related to the medicinal product. AEs may also include pre- or post-treatment complications that occur as a result of protocol specified procedures, lack of efficacy, overdose, drug abuse/misuse reports, or occupational exposure. Preexisting events that increase in severity or change in nature during or as a consequence of participation in the clinical study will also be considered AEs.

An AE does not include the following:

- Medical or surgical procedures such as surgery, endoscopy, tooth extraction, and transfusion. The condition that led to the procedure may be an adverse event and must be reported.
- Pre-existing diseases, conditions, or laboratory abnormalities present or detected before the screening visit that do not worsen
- Situations where an untoward medical occurrence has not occurred (eg, hospitalization for chemotherapy infusion per institutional guidelines, elective surgery, social and/or convenience admissions)
- Overdose without clinical sequelae
- Any medical condition or clinically significant laboratory abnormality with an onset date before the consent form is signed and not related to a protocol-associated procedure is not an AE. It is considered to be pre-existing and should be documented on the medical history eCRF.

7.1.2. Serious Adverse Events

An SAE is defined as an event that, at any dose, results in the following:

- Death
- Life-threatening (Note: The term “life-threatening” in the definition of “serious” refers to an event in which the subject was at risk of death at the time of the event; it does not refer to an event that hypothetically might have caused death if it were more severe.)

- In-patient hospitalization or prolongation of existing hospitalization (Note: Hospitalization for chemotherapy infusion per institutional guidelines will not be considered an SAE.)

- Persistent or significant disability/incapacity

- A congenital anomaly/birth defect

- A medically important event or reaction: such events may not be immediately life-threatening or result in death or hospitalization but may jeopardize the subject or may require intervention to prevent one of the other outcomes constituting SAEs. Medical and scientific judgment must be exercised to determine whether such an event is a reportable under expedited reporting rules. Examples of medically important events include intensive treatment in an emergency room or at home for allergic bronchospasm; blood dyscrasias or convulsions that do not result in hospitalization; and development of drug dependency or drug abuse. For the avoidance of doubt, infections resulting from contaminated medicinal product will be considered a medically important event and subject to expedited reporting requirements.

7.1.2.1. Protocol-Specific Serious Adverse Event Instructions

To maintain the integrity of the study, disease progression and death from disease progression should be reported as SAEs by the investigator only if it is assessed that the study drugs caused or contributed to the disease progression (ie, by a means other than lack of effect). Unrelated disease progression should be captured on the eCRF.

In addition, events that are indicative of the following disease-related SAEs that are assessed as unrelated to study drugs will not be reported as expedited reports by Gilead during the study:

- Progression of gastric cancer

- Death related to disease progression

These events will be exempt from global expedited reporting requirements for the duration of the study as they are the primary endpoints of this study. They will be reported as appropriate in the final clinical study report as well as any relevant aggregate safety report.

7.2. Assessment of Adverse Events and Serious Adverse Events

The investigator or qualified subinvestigator is responsible for assessing AEs and SAEs for causality and severity, and for final review and confirmation of accuracy of event information and assessments.

7.2.1. Assessment of Causality for Study Drugs and Procedures

The investigator or qualified subinvestigator is responsible for assessing the relationship to IMP therapy using clinical judgment and the following considerations:

- **No:** Evidence exists that the adverse event has an etiology other than the study drug. For SAEs, an alternative causality must be provided (eg, pre-existing condition, underlying disease, intercurrent illness, or concomitant medication).

- **Yes:** There is reasonable possibility that the event may have been caused by the investigational medicinal product.

It should be emphasized that ineffective treatment should not be considered as causally related in the context of adverse event reporting.

The relationship to study procedures (eg, invasive procedures such as venipuncture or biopsy) should be assessed using the following considerations:

- **No:** Evidence exists that the adverse event has an etiology other than the study procedure.

- **Yes:** The adverse event occurred as a result of protocol procedures, (eg., venipuncture)

7.2.2. Assessment of Severity

The severity of AEs will be graded using the National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE), Version 4.03 in the study manual.

If an NCI CTCAE term is not available for the AE/SAE, the severity will be graded per the General Guidance rules found in the introduction section of the NCI CTCAE , Version 4.03 document.

The distinction between the seriousness and the severity of an AE should be noted. Severe is a measure of intensity; thus, a severe (Grade 3) reaction is not necessarily a serious adverse event (SAE).

7.3. Investigator Requirements and Instructions for Reporting Adverse Events and Serious Adverse Events to Gilead

Requirements for collection prior to study drug initiation:

After informed consent, but prior to initiation of study medication, the following types of events should be reported on the eCRF:

- All SAEs and adverse events related to protocol-mandated procedures

7.3.1. Adverse Events

Following initiation of study medication, collect all AEs, regardless of cause or relationship, until 55 days after last administration of GS-5745/placebo or all components of mFOLFOX6, whichever is discontinued later. These must be reported to the eCRF database as instructed.

All AEs should be followed up until resolution or until the adverse event is stable, if possible. Gilead Sciences may request that certain AEs be followed beyond the protocol defined follow-up period.

7.3.2. Serious Adverse Events

All SAEs, regardless of cause or relationship, that occurs after the subject first consents to participate in the study (ie, signing the informed consent) and throughout the duration of the study, including the protocol-required post treatment follow-up period, must be reported to the eCRF database and Gilead DSPH as instructed. This also includes any SAEs resulting from protocol-associated procedures performed after informed consent is signed.

Any SAEs and deaths that occur after the post treatment follow-up visit but within 55 days of the last dose of GS-5745/placebo or all components of mFOLFOX6, whichever is discontinued later, regardless of causality, should also be reported.

Investigators are not obligated to actively seek SAEs after the protocol defined follow up period, however, if the investigator learns of any SAEs that occur after study participation has concluded and the event is deemed relevant to the use of study drug, he/she should promptly document and report the event to Gilead DSPH.

- All AEs and SAEs will be recorded in the eCRF database within the timelines outlined in the eCRF completion guideline.

- SAEs will be reported using an electronic SAE (eSAE) system.

7.3.2.1. Electronic Serious Adverse Event (eSAE) Reporting Process

- Site personnel record all SAE data in the eCRF database and from there transmit the SAE information to Gilead DSPH within 24 hours of the investigator's knowledge of the event. Detailed instructions can be found in the eCRF completion guidelines.

- If for any reason it is not possible to record the SAE information electronically, ie, the eCRF database is not functioning, record the SAE on the paper serious adverse event reporting form and submit within 24 hours to:

Gilead DSPH: Fax: PPD
Email: PPD

- As soon as it is possible to do so, any SAE reported via paper must be transcribed into the eCRF database according to instructions in the eCRF completion guidelines.

- If an SAE has been reported via a paper form because the eCRF database has been locked, no further action is necessary.

- For fatal or life-threatening events, copies of hospital case reports, autopsy reports, and other documents are also to be submitted by e-mail or fax when requested and applicable. Transmission of such documents should occur without personal subject identification, maintaining the traceability of a document to the subject identifiers.

- Additional information may be requested to ensure the timely completion of accurate safety reports.

- Any medications necessary for treatment of the SAE must be recorded onto the concomitant medication section of the subject's eCRF and the event description section of the SAE form.

7.4. Gilead Reporting Requirements

Depending on relevant local legislation or regulations, including the applicable US FDA Code of Federal Regulations, the EU Clinical Trials Directive (2001/20/EC) and relevant updates, and other country-specific legislation or regulations, Gilead may be required to expedite to worldwide regulatory agencies reports of SAEs, SADRs, or suspected unexpected serious adverse reactions (SUSARs). In accordance with the EU Clinical Trials Directive (2001/20/EC), Gilead or a specified designee will notify worldwide regulatory agencies and the relevant IEC in concerned Member States of applicable SUSARs as outlined in current regulations.

Assessment of expectedness for SAEs will be determined by Gilead using reference safety information specified in the IB or relevant local label as applicable.

All investigators will receive a safety letter notifying them of relevant SUSAR reports associated with any study drug. The investigator should notify the IRB or IEC of SUSAR reports as soon as is practical, where this is required by local regulatory agencies, and in accordance with the local institutional policy.

7.5. Clinical Laboratory Abnormalities and Other Abnormal Assessments as Adverse Events or Serious Adverse Events

Laboratory abnormalities are usually not recorded as AEs or SAEs. However, laboratory abnormalities (eg, clinical chemistry, hematology, and urinalysis) that require medical or surgical intervention or lead to interruption or discontinuation of study drug (GS-5745/placebo) or mFOLFOX6 must be recorded as an AE, as well as an SAE, if applicable. In addition, laboratory or other abnormal assessments (eg, electrocardiogram, X-rays, vital signs) that are associated with signs and/or symptoms must be recorded as an AE or SAE if they meet the definition of an AE or SAE as described in Sections 7.1.1 and 7.1.2, respectively. If the laboratory abnormality is part of a syndrome, record the syndrome or diagnosis (ie, anemia) not the laboratory result (ie, decreased hemoglobin).

Severity should be recorded and graded according to the NCI CTCAE (version 4.03).

For AEs associated with laboratory abnormalities, the event should be graded on the basis of the clinical severity in the context of the underlying conditions; this may or may not be in agreement with the grading of the laboratory abnormality.

7.6. Toxicity Management

Treatment-emergent toxicities will be noted by the Investigator and brought to the attention of the Gilead Sciences Medical Monitor or designee. Whether or not considered treatment-related, all subjects experiencing AEs must be monitored periodically until symptoms subside, any abnormal laboratory values have resolved or returned to baseline levels or they are considered irreversible, or until there is a satisfactory explanation for the changes observed.

Grade 3 or 4 clinically significant laboratory abnormalities should be confirmed by repeat testing as soon as practical to do so, and preferably within 3 calendar days after receipt of the original test results. Laboratory abnormalities (eg, thiamine deficiency) identified at screening/baseline and during study participation should be treated at the investigators discretion.

Any questions regarding toxicity management should be directed to the Gilead Sciences Medical Monitor or designee.

7.7. Special Situations Reports

7.7.1. Definitions of Special Situations

Special situation reports include all reports of medication error, abuse, misuse, overdose, reports of adverse events associated with product complaints, and pregnancy reports regardless of an associated AE.

Medication error is any unintentional error in the prescribing, dispensing, or administration of a medicinal product while in the control of the health care provider, subject, or consumer.

Abuse is defined as persistent or sporadic intentional excessive use of a medicinal product by a subject.

Misuse is defined as any intentional and inappropriate use of a medicinal product that is not in accordance with the protocol instructions or the local prescribing information.

An overdose is defined as an accidental or intentional administration of a quantity of a medicinal product given per administration or cumulatively which is above the maximum recommended dose as per protocol or in the product labelling (as it applies to the daily dose of the subject in question). In cases of a discrepancy in drug accountability, overdose will be established only when it is clear that the subject has taken the excess dose(s). Overdose cannot be established when the subject cannot account for the discrepancy except in cases in which the investigator has reason to suspect that the subject has taken the additional dose(s).

Product complaint is defined as complaints arising from potential deviations in the manufacture, packaging, or distribution of the medicinal product.

7.7.2. Instructions for Reporting Special Situations

7.7.2.1. Instructions for Reporting Pregnancies

The investigator should report pregnancies in female study subjects that are identified after initiation of study medication and throughout the study, including the post study drug follow-up period, to Gilead DSPH using the pregnancy report form within 24 hours of becoming aware of the pregnancy.

Refer to the eCRF completion guidelines for full instructions on the mechanism of pregnancy reporting.

The pregnancy itself is not considered an AE nor is an induced elective abortion to terminate a pregnancy without medical reasons.

Any premature termination of pregnancy (eg, a spontaneous abortion, an induced therapeutic abortion due to complications or other medical reasons) must be reported within 24 hours as an SAE. The underlying medical reason for this procedure should be recorded as the AE term.

A spontaneous abortion is always considered to be an SAE and will be reported. Furthermore, any SAE occurring as an adverse pregnancy outcome post study must be reported to Gilead DSPH.

The subject should receive appropriate monitoring and care until the conclusion of the pregnancy. The outcome should be reported to or Gilead DSPH using the pregnancy outcome report form. If the end of the pregnancy occurs after the study has been completed, the outcome should be reported directly to Gilead DSPH.

Pregnancies of female partners of male study subjects exposed to Gilead or other study drugs must also be reported and relevant information should be submitted to Gilead DSPH using the pregnancy and pregnancy outcome forms within 24 hours. Monitoring of the subject should continue until the conclusion of the pregnancy. If the end of the pregnancy occurs after the study has been completed, the outcome should be reported directly to Gilead DSPH.

Gilead DSPH contact information is as follows: Fax: PPD
Email: PPD

Refer to Appendix 3 for Pregnancy Precautions, Definition for Female of Childbearing Potential, and Contraceptive Requirements.

7.7.2.2. Emergency Unblinding

In the event of a medical emergency where breaking of the blind is being considered by the treating physician, the investigator may break the blind using the IWRS. It is recommended that the Gilead Medical Monitor be contacted before the investigator breaks the blind. Please note that the treatment assignment should not be unblinded if unblinding will not affect the way the subject would be treated. In the event of a medical emergency, where breaking of the blind is required per the medical judgment of the investigator, the Gilead Medical Monitor must be contacted as soon as possible after the unblinding. The unblinding must be clearly justified and explained by a comment in the source documentation, along with the date on which the code was broken and the identity of the person authorizing the unblinding.

Blinding of study treatment is critical to the integrity of this clinical trial and therefore, if a subject's treatment assignment is disclosed to the investigator, the subject will have study treatment discontinued. All subjects will be followed until study completion unless consent to do so is specifically withdrawn by the subject.

Gilead DSPH may independently unblind cases for expedited reporting of SUSARs.

7.7.2.3. Reporting Other Special Situations

All other special situation reports must be reported on the special situations report form and forwarded to Gilead DSPH within 24 hours of the investigator becoming aware of the situation. These reports must consist of situations that involve study IMP and/or Gilead concomitant medications, but do not apply to non-Gilead concomitant medications.

Special situations involving non-Gilead concomitant medications does not need to be reported on the special situations report form; however, for special situations that result in AEs due to a non-Gilead concomitant medication, the AE should be reported on the AE form.

Any inappropriate use of concomitant medications prohibited by this protocol should not be reported as "misuse," but may be more appropriately documented as a protocol deviation.

Refer to the eCRF completion guidelines for full instructions on the mechanism of special situations reporting.

All clinical sequelae in relation to these special situation reports will be reported as AEs or SAEs at the same time using the AE eCRF and/or the SAE report form. Details of the symptoms and signs, clinical management, and outcome will be reported, when available.

8. STATISTICAL CONSIDERATIONS

8.1. Analysis Objectives

8.1.1. Primary Objective

The primary objective is to compare the efficacy of GS-5745 versus placebo in combination with mFOLFOX6 as measured by OS.

8.1.2. Secondary Objectives

The secondary objectives are:

- To compare the efficacy of GS-5745 versus placebo in combination with mFOLFOX6 as measured by PFS
- To compare the efficacy of GS-5745 versus placebo in combination with mFOLFOX6 as measured by ORR per RECIST v1.1
- To compare the safety profile of GS-5745 versus placebo in combination with mFOLFOX6

8.1.3. Exploratory Objectives

8.2. Analysis Endpoints

8.2.1. Primary Endpoint

The primary endpoint of the study is OS as defined in Section 2.4.

8.2.2. Secondary and Exploratory Endpoints

The secondary and exploratory endpoints of the study are defined in Section 2.5.

8.3. Analysis Conventions

8.3.1. Analysis Sets

8.3.1.1. Intent-to-Treat (ITT) Analysis Set

The ITT analysis set includes data from all randomized subjects. Study drug assignment will be designated according to randomization.

This analysis set will be used in the analyses of subject characteristics, OS, PFS, ORR and DCR. The analysis of OS based on the ITT analysis set will be considered the primary analysis of the study.

Subjects in the ITT analysis set who do not have sufficient baseline or on-study tumor status information to be adequately assessed for response status (ie, those with best overall responses of NE or ND) will be included in the denominators in calculations of response rates and disease control rates.

8.3.1.2. Safety Analysis Set

The safety analysis set will include data from all subjects who receive ≥1 dose of study treatment, with treatment assignments designated according to the actual treatment received.

This analysis set will be used in the analysis of safety variables as well as study treatment administration. All data collected up to the last dose of GS-5745/placebo or mFOLFOX6 or 5-FU and LV, whichever is later, plus 55 days will be included in the safety summaries.

8.3.1.3. Pharmacodynamic and Pharmacokinetic (PK) Analysis Sets

The pharmacodynamics and PK analysis sets will include data from subjects in the safety analysis set who have the necessary baseline and on-study measurements to provide interpretable results for the specific parameters of interest.

8.4. Data Handling Conventions

8.4.1. General Methods

By-subject listings will be created for important variables from each eCRF module.

Summary tables for continuous variables will contain the following statistics:

N (number in analysis set), n (number with data), mean, standard deviation (std), 95% confidence intervals (CIs) on the mean, median, minimum, and maximum. Summary tables for categorical variables will include: N, n, percentage, and 95% CIs on the percentage. Unless otherwise indicated, 95% CIs for binary variables will be calculated using the binomial distribution (exact method) and will be 2-sided. Data will be described and summarized by relevant treatment arm, analysis set, and time point. As appropriate, changes from baseline to

each subsequent time point will be described and summarized by treatment arm. Similarly, as appropriate, the best change from baseline during the study will also be described and summarized by treatment arm. Graphical techniques (eg, waterfall plots, Kaplan-Meier curves, etc.) may be used when such methods are appropriate and informative.

The baseline value used in each analysis will be the last (most recent) pre-treatment value.

Subjects with discrepancies between the stratification factor values at randomization and the actual values as documented on data review will be categorized in the analyses according to the actual values. In the situation that there is insufficient information in a stratum (ie, if there are < 6 subjects or there are no events in a stratum), that stratum will be pooled with the smallest adjacent stratum for stratified analyses; the smallest stratum is defined as that stratum having the fewest number of subjects or the fewest number of events in case the former is a tie and the adjacent stratum is defined as a stratum having 2 factors of the 3 at the same level. Data from all sites will be pooled for all analyses. Analyses will be based upon the observed data unless methods for handling missing data are specified. If there is a significant degree of non-normality, analyses may be performed on log-transformed data or nonparametric tests may be applied, as appropriate.

Unless otherwise specified, all analyses will be 1-sided at the 0.025 level of significance.

The following censoring conventions will be applied to tumor control endpoints:

- OS: Data from surviving subjects will be censored at the last time that the subject was known to be alive.
- PFS: Data from surviving, non-progressing subjects will be censored at the earliest of the time of initiation of antitumor treatment other than the study treatment or the last time that lack of definitive progression was objectively documented. Data from subjects who have disease progression or die after ≥ 2 consecutive missing tumor assessments will be censored at the last time prior to the missing assessments that lack of definitive disease progression was objectively documented.
- DOR: Data from surviving, non-progressing subjects will be censored at the earliest of the time of initiation of antitumor treatment other than the study treatment or the last time that lack of definitive disease progression was objectively documented. Data from subjects who have disease progression or die after ≥ 2 consecutive missing tumor assessments will be censored at the last time prior to the missing assessments that lack of definitive disease progression was objectively documented.

8.4.2. Demographic Data and Baseline Characteristics

Demographic summaries will include sex, race/ethnicity, randomization stratification group, and age.

Baseline characteristics will include a summary of body weight, height, and body mass index.

Demographic and baseline characteristics will be summarized using standard descriptive methods.

8.5. Efficacy Analysis

8.5.1. Primary Analysis

The null hypothesis is that addition of GS-5745 to mFOLFOX6 will not improve OS in first-line patients with advanced gastric cancer. The alternative hypothesis is that addition of GS-5745 to mFOLFOX6 will improve OS.

The Kaplan-Meier (KM) method and logrank test stratified by ECOG status (0 or 1), geographic region (Latin America or other participating countries), and primary tumor site (gastric or GEJ) will be used to compare the 2 treatment groups. A Cox proportional hazard model with the same stratification factors will be used to estimate the hazard ratio and corresponding 95% confidence interval (CI).

8.5.2. Secondary and Exploratory Analyses

CCI

[redacted]

8.5.3. Control of Type I Error Rate in Efficacy Analyses

In the efficacy analyses, the following procedures will be implemented to preserve the overall type I error rate across the primary and secondary endpoints of the study, and at interim and final analysis, at a 1-sided significance level of 0.025.

The primary endpoint analysis will serve as the gatekeeper for the secondary endpoint analyses, ie, the primary efficacy endpoint must be met before the secondary efficacy endpoints can be tested. The secondary endpoints included in this sequential testing approach are (1) PFS and (2) ORR.

If the primary hypothesis is rejected, the 2 secondary endpoints will be sequentially tested at the 1-sided 0.016 significance level in the order listed above. If a null hypothesis is not rejected, formal sequential testing will be stopped and only nominal significance will be cited for the remaining secondary endpoints. Analyses and p-values will be reported for all the efficacy endpoints, including the primary endpoint, the secondary endpoints, and all of the exploratory endpoints.

8.6. Safety Analysis

All safety data collected on or after the date that GS-5745/placebo was first administered up to the date of last dose of GS-5745/placebo or mFOLFOX6 or 5-FU and LV (whichever is later) plus 55 days will be summarized by treatment group (according to the treatment received). Data for the pre-treatment and post-treatment follow-up period will be included in data listings.

In general, count and percent of subjects will summarize categorical and ordinal data. Mean, standard deviation, minimum, quartiles, median, and maximum will summarize continuous data.

8.6.1. Extent of Exposure

A subject's extent of exposure to GS-5745 will be generated from the study drug administration eCRF page. Exposure data will be summarized by treatment group.

8.6.2. Adverse Events

Clinical and laboratory adverse events will be coded using the Medical Dictionary for Regulatory Activities (MedDRA). System Organ Class (SOC), High-Level Group Term (HLGT), High-Level Term (HLT), Preferred Term (PT), and Lower-Level Term (LLT) will be attached to the clinical database.

Events will be summarized on the basis of the date of onset for the event. A treatment-emergent adverse event is defined as any adverse event with onset date on or after the date of first dose of study drug up to 55 days after permanent study drug discontinuation or any adverse events leading to premature study drug discontinuation.

Summaries (number and percentage of subjects) of treatment-emergent adverse events (by SOC, HLT, and PT) will be provided by treatment group.

8.6.3. Laboratory Evaluations

Selected laboratory data (using conventional units) will be summarized using only observed data. Data and change from baseline at all scheduled time points will be summarized.

Graded laboratory abnormalities will be defined using the NCI CTCAE (version 4.03). Maximum post-baseline grade will be summarized by count and percent of subjects with each grade.

Incidence of treatment-emergent laboratory abnormalities, defined as values that increase at least 1 toxicity grade from baseline at any time post-baseline, will be summarized by treatment group. If baseline data are missing, then any graded abnormality (ie, at least a Grade 1) will be considered treatment-emergent.

Laboratory abnormalities that occur before the first dose of study drug will be included in data listings.

8.6.4. Other Safety Evaluations

Similar general approaches to the AE and clinical laboratory data will be utilized to summarize other safety measures.

8.7. Pharmacokinetic Analysis

The plasma concentration data of GS-5745, oxaliplatin, and 5-FU (and metabolite, if applicable) will be summarized by nominal sampling time using descriptive statistics (eg, sample size, arithmetic mean, geometric mean, coefficient of variation (%) standard deviation, median, minimum, and maximum). PK parameters (C_{max}, T_{max}, C_{last}, T_{last}, C_{tau}, λ_z, AUC_{last}, AUC_{tau}, and $t_{½}$, as applicable) will be listed and summarized using descriptive statistics. Plasma concentrations over time may also be plotted in semi-logarithmic and linear formats as mean ± standard deviation, and median (Q1, Q3). Exposures (C_{max} and AUC) of oxaliplatin and 5-FU with or without co-administration of GS-5745 will be compared to evaluate if GS-5745 treatment alters the PK of oxaliplatin/5-FU.

Exposure-response analysis may be explored as appropriate.

The number and percentage of positive or negative anti-GS-5745 antibody values at each specified timepoint will be summarized. The effect of anti-GS-5745 antibody on GS-5745 PK, safety, and efficacy may be evaluated.

8.8. Biomarker Analysis

8.8.1. Pharmacodynamic and Exploratory Biomarker Analysis Analysis

CCI [redacted]

8.9. Patient-Reported Outcomes

The values and change from baseline for EQ-5D-3L index scores and the visual analog scale (VAS) will be summarized by descriptive statistics by visit for each treatment arm.

The EORTC QLQ-C30 and QLQ-STO22 questionnaire will be scored according to EORTC guidelines, such that all scales from 0 to 100, with higher functioning and global scores representing better QoL and higher symptom scores representing greater symptom burden. The values and change from baseline for QLQ-C30 and QLQ-STO22 scores will be summarized by descriptive statistics by visit for each treatment arm.

The time to QLQ-C30 and QLQ-STO22 deterioration is defined as the time from randomization to the first deterioration of ≥ 10 points from baseline. If no deterioration was observed, censoring occurs at the date of the last assessment. The hazard ratio and the 95% CI comparing the treatment and placebo arm will be provided using the Cox proportional hazard model.

The QLQ-C30 and QLQ-STO22 analysis will also characterize each post baseline assessment as Improved if the improvement from baseline is ≥ 10 points, and Stable if the change from baseline is within +/-10 points, exclusive. The proportion of subjects in each treatment arm with Improved/Stable scores will be summarized at each time point and overall. The odds ratio and the 95% CI comparing the treatment and placebo arm will be provided.

8.10. Interim Analyses

Two formal interim analyses (detailed in Appendix 6) are planned: the futility interim will be performed at approximately 33.3% information and the efficacy interim at 66.7% information. The final analysis will occur when 286 OS events have been observed.

The Lan-DeMets approach with O'Brien-Fleming type alpha spending function will be used to control the type I error rate for testing. The stopping boundaries at each efficacy analysis time are provided in Table 8-1.

Table 8-1. Stopping Boundaries for Efficacy Analyses

Efficacy Analysis	Events (%)	Stopping Boundary	
		Z scale	One-sided p-value scale
Interim	191 (66.7%)	2.509	0.006
Final	286 (100%)	1.993	0.023

Given significance of OS at either the interim or final analysis, the secondary endpoints of PFS and ORR will sequentially tested at the 1-sided 0.016 significance level.

If the DMC sees substantial evidence of benefit of the GS-5745 + mFOLFOX6 combination, (ie, OS is significantly better in Arm A compared to Arm B), Gilead personnel who are not involved with the study may perform an unblinded review of the integrated interim data, and may decide to stop the study and offer GS-5745 treatment to those subjects who are still enrolled in Arm B.

8.11. Sample Size

Assuming a median OS time for the mFOLFOX6 + placebo group of 11.5 months, 286 events are needed to detect a hazard ratio (HR) of 0.70 with 85% power at a 1-sided significance level of 0.025 using a log-rank test, given 1 efficacy interim at 66.7% information. With an accrual period of 18 months, minimum follow-up of 18 months, and a 10% annual dropout rate, a total sample size of 430 subjects (215 subjects per treatment group) is needed to observe the required 286 events within the given time frame.

As the targeted number of deaths is large (~286), if the null hypothesis of the primary end point of OS is rejected, it will convincingly demonstrate clinical treatment effect and would provide a narrow confidence interval. Based on the number of deaths and the assumed treatment effect on OS in the protocol, the expected 95% confidence interval on HR of OS between the 2 treatment groups is approximately (0.627, 0.996). At the time of final analysis and assuming OS is significant, PFS will be tested at one-sided type I error of 0.016. Assuming that the hazard ratio for PFS is 0.7, which is expected to be on par or better than the treatment effect in OS, and that median PFS for the control is 9 months, a sample size of 430 subjects (322 PFS events) will provide 85% power under the planned study enrollment (18 months), follow-up duration (18 month), and drop-out rate (annually 10%).

8.12. Data Monitoring Committee

An independent DMC will review the data for safety and efficacy. Safety review by the DMC will be performed when the first 60 subjects have completed 4 treatment cycles. Thereafter, review of safety data will be performed at regular intervals as described in the DMC charter. In addition, the DMC will meet after approximately 33.3% and approximately 66.7% of the expected number of events have occurred to review the results from the futility and efficacy interim analysis, respectively.

The DMC's specific activities will be defined by a mutually agreed charter, which will define the DMC's membership, conduct and meeting schedule.

While the DMC will be asked to advise Gilead regarding future conduct of the study, including possible early study termination, Gilead retains final decision-making authority on all aspects of the study.

9. RESPONSIBILITIES

9.1. Investigator Responsibilities

9.1.1. Good Clinical Practice

The investigator will ensure that this study is conducted in accordance with the principles of the Declaration of Helsinki (as amended in Edinburgh, Tokyo, Venice, Hong Kong, and South Africa), International Conference on Harmonisation (ICH) guidelines, or with the laws and regulations of the country in which the research is conducted, whichever affords the greater protection to the study subject. These standards are consistent with the European Union Clinical Trials Directive 2001/20/EC and Good Clinical Practice Directive 2005/28/EC.

The investigator will ensure adherence to the basic principles of Good Clinical Practice, as outlined in 21 CFR 312, subpart D, "Responsibilities of Sponsors and Investigators," 21 CFR, part 50, 1998, and 21 CFR, part 56, 1998.

The investigator and all applicable subinvestigators will comply with 21 CFR, Part 54, 1998, providing documentation of their financial interest or arrangements with Gilead, or proprietary interests in the investigational drug under study. This documentation must be provided prior to the investigator's (and any subinvestigator's) participation in the study. The investigator and subinvestigator agree to notify Gilead of any change in reportable interests during the study and for 1 year following completion of the study. Study completion is defined as the date when the last subject completes the protocol-defined activities.

9.1.2. Institutional Review Board (IRB)/Independent Ethics Committee (IEC) Review and Approval

The investigator (or sponsor as appropriate according to local regulations) will submit this protocol, informed consent form, and any accompanying material to be provided to the subject (such as advertisements, subject information sheets, or descriptions of the study used to obtain informed consent) to an IRB/IEC. The investigator will not begin any study subject activities until approval from the IRB/IEC has been documented and provided as a letter to the investigator.

Before implementation, the investigator will submit to and receive documented approval from the IRB/IEC any modifications made to the protocol or any accompanying material to be provided to the subject after initial IRB/IECapproval, with the exception of those necessary to reduce immediate risk to study subjects.

9.1.3. Informed Consent

The investigator is responsible for obtaining written informed consent from each individual participating in this study after adequate explanation of the aims, methods, objectives, and potential hazards of the study and before undertaking any study-related procedures.

The investigator must use the most current IRB/IEC approved consent form for documenting written informed consent. Each informed consent (or assent as applicable) will be appropriately signed and dated by the subject or the subject's legally authorized representative and the person conducting the consent discussion, and also by an impartial witness if required by IRB/IEC local requirements. The consent form will inform subjects about genomic testing and sample retention, and their right to receive clinically relevant genomic analysis results.

9.1.4. Confidentiality

The investigator must assure that subjects' anonymity will be strictly maintained and that their identities are protected from unauthorized parties. Only subject initials, date of birth, another unique identifier (as allowed by local law) and an identification code will be recorded on any form or biological sample submitted to the Sponsor, IRB/IEC, or laboratory. Laboratory specimens must be labeled in such a way as to protect subject identity while allowing the results to be recorded to the proper subject. Refer to specific laboratory instructions.
NOTE: The investigator must keep a screening log showing codes, names, and addresses for all subjects screened and for all subjects enrolled in the trial. Subject data will be processed in accordance with all applicable regulations.

The investigator agrees that all information received from Gilead, including but not limited to the investigator brochure, this protocol, eCRF, the study drug, and any other study information, remain the sole and exclusive property of Gilead during the conduct of the study and thereafter. This information is not to be disclosed to any third party (except employees or agents directly involved in the conduct of the study or as required by law) without prior written consent from Gilead. The investigator further agrees to take all reasonable precautions to prevent the disclosure by any employee or agent of the study site to any third party or otherwise into the public domain.

9.1.5. Study Files and Retention of Records

The investigator must maintain adequate and accurate records to enable the conduct of the study to be fully documented and the study data to be subsequently verified. These documents should be classified into at least the following 2 categories: (1) investigator's study file, and (2) subject clinical source documents.

The investigator's study file will contain the protocol/amendments, IRB/IEC and governmental approval with correspondence, informed consent, drug records, staff curriculum vitae and authorization forms, and other appropriate documents and correspondence.

The required source data should include sequential notes containing at least the following information for each subject:

- Subject identification (name, date of birth, gender);
- Documentation that subject meets eligibility criteria, ie, history, physical examination, and confirmation of diagnosis (to support inclusion and exclusion criteria);

- Documentation of the reason(s) a consented subject is not randomized
- Participation in study (including study number);
- Study discussed and date of informed consent;
- Dates of all visits;
- Documentation that protocol specific procedures were performed;
- Results of efficacy parameters, as required by the protocol;
- Start and end date (including dose regimen) of IMP, including dates of dispensing and return;
- Record of all adverse events and other safety parameters (start and end date, and including causality and severity);
- Concomitant medication (including start and end date, dose if relevant; dose changes);
- Date of study completion and reason for early discontinuation, if it occurs.

All clinical study documents must be retained by the investigator until at least 2 years or according to local laws, whichever is longer, after the last approval of a marketing application in an ICH region (ie, United States, Europe, or Japan) and until there are no pending or planned marketing applications in an ICH region; or, if no application is filed or if the application is not approved for such indication, until 2 years after the investigation is discontinued and regulatory authorities have been notified. Investigators may be required to retain documents longer if specified by regulatory requirements, by local regulations, or by an agreement with Gilead. The investigator must notify Gilead before destroying any clinical study records.

Should the investigator wish to assign the study records to another party or move them to another location, Gilead must be notified in advance.

If the investigator cannot provide for this archiving requirement at the study site for any or all of the documents, special arrangements must be made between the investigator and Gilead to store these records securely away from the site so that they can be returned sealed to the investigator in case of an inspection. When source documents are required for the continued care of the subject, appropriate copies should be made for storage away from the site.

9.1.6. Case Report Forms

For each subject consented, eCRFs will be completed by an authorized study staff member whose training for this function is documented according to study procedures. eCRFs should be completed on the day of the subject visit to enable the sponsor to perform central monitoring of safety data. Subsequent to data entry, a study monitor will perform source data verification within the electronic data capture (EDC) system. Original entries as well as any changes to data

fields will be stored in the audit trail of the system. Prior to database lock (or any interim time points as described in the Clinical Data Management Plan), the investigator will use his/her log in credentials to confirm that the forms have been reviewed, and that the entries accurately reflect the information in the source documents. The eCRF capture the data required per the protocol schedule of events and procedures. System-generated or manual queries will be issued to the investigative site staff as data discrepancies are identified by the monitor or internal Gilead staff, who routinely review the data for completeness, correctness, and consistency. The site coordinator is responsible for responding to the queries in a timely manner, within the system, either by confirming the data as correct or updating the original entry, and providing the reason for the update (eg, data entry error). At the conclusion of the trial, Gilead will provide the site with a read-only archive copy of the data entered by that site. This archive must be stored in accordance with the records retention requirements.

9.1.7. Investigational Medicinal Product & mFOLFOX6 Accountability and Return

Used and unused study drug supplies and mFOLFOX6, should be destroyed on site if the site has an appropriate standard operating procedure (SOP) for drug destruction as determined by Gilead. The site may destroy used (empty or partially empty) and unused study drug supplies in accordance with that site's approved SOP. A copy of the site's approved SOP will be obtained for central files.

The study monitor will evaluate each study center's study drug disposal procedures and provide appropriate instruction for destruction of unused study drug supplies on site. The investigator must maintain accurate records for all study drug destroyed at the site. Records must show the identification and quantity of each unit destroyed, the method of destruction, and the person who disposed of the study drug. Upon study completion, copies of the study drug accountability records must be filed at the site. Another copy will be returned to Gilead.

If destruction of study drug on site is not possible, arrangements will be made between the site and Gilead Sciences (or Gilead Sciences' representative) for return of unused study drug supplies. The monitor will provide further instructions for the return.

The study monitor will review study drug supplies and associated records at study monitoring visits.

9.1.8. Inspections

The investigator will make available all source documents and other records for this trial to Gilead's appointed study monitors, to IRB/IEC, or to regulatory authority or health authority inspectors.

9.1.9. Protocol Compliance

The investigator is responsible for ensuring the study is conducted in accordance with the procedures and evaluations described in this protocol.

9.2. Sponsor Responsibilities

9.2.1. Protocol Modifications

Protocol modifications, except those intended to reduce immediate risk to study subjects, may be made only by Gilead. The investigator must submit all protocol modifications to the IRB/IEC in accordance with local requirements and receive documented IRB/IECapproval before modifications can be implemented.

9.2.2. Study Report and Publications

A clinical study report (CSR) will be prepared and provided to the regulatory agency (ies) Gilead will ensure that the report meets the standards set out in the ICH Guideline for Structure and Content of Clinical Study Reports (ICH E3). Note that an abbreviated report may be prepared in certain cases.

Investigators in this study may communicate, orally present, or publish in scientific journals or other scholarly media only after the following conditions have been met:

- The results of the study in their entirety have been publicly disclosed by or with the consent of Gilead in an abstract, manuscript, or presentation form or the study has been completed at all study sites for at least 2 years
- The investigator will submit to Gilead any proposed publication or presentation along with the respective scientific journal or presentation forum at least 30 days before submission of the publication or presentation
- No such communication, presentation, or publication will include Gilead's confidential information (see Section 9.1.4)
- The investigator will comply with Gilead's request to delete references to its confidential information (other than the study results) in any paper or presentation and agrees to withhold publication or presentation for an additional 60 days in order to obtain patent protection if deemed necessary

9.3. Joint Investigator/Sponsor Responsibilities

9.3.1. Payment Reporting

Investigators and their study staff may be asked to provide services performed under this protocol, eg, attendance at Investigator's Meetings. If required under the applicable statutory and regulatory requirements, Gilead will capture and disclose to Federal and State agencies any expenses paid or reimbursed for such services, including any clinical trial payments, meal, travel expenses or reimbursements, consulting fees, and any other transfer of value.

9.3.2. Access to Information for Monitoring

In accordance with regulations and guidelines, the study monitor must have direct access to the investigator's source documentation in order to verify the accuracy of the data recorded in the eCRF.

The monitor is responsible for routine review of the eCRF at regular intervals throughout the study to verify adherence to the protocol and the completeness, consistency, and accuracy of the data being entered on them. The monitor should have access to any subject records needed to verify the entries on the eCRF. The investigator agrees to cooperate with the monitor to ensure that any problems detected through any type of monitoring (central, on site) are resolved.

9.3.3. Access to Information for Auditing or Inspections

Representatives of regulatory authorities or of Gilead may conduct inspections or audits of the clinical study. If the investigator is notified of an inspection by a regulatory authority the investigator agrees to notify the Gilead medical monitor immediately. The investigator agrees to provide to representatives of a regulatory agency or Gilead access to records, facilities, and personnel for the effective conduct of any inspection or audit.

9.3.4. Study Discontinuation

Both the sponsor and the investigator reserve the right to terminate the study at any time. Should this be necessary, both parties will arrange discontinuation procedures and notify the appropriate regulatory authority(ies), IRBs, and IECs. In terminating the study, Gilead and the investigator will assure that adequate consideration is given to the protection of the subjects' interests.

10. REFERENCES

Agrawal S, Anderson P, Durbeej M, van Rooijen N, Ivars F, Opdenakker G, et al. Dystroglycan is selectively cleaved at the parenchymal basement membrane at sites of leukocyte extravasation in experimental autoimmune encephalomyelitis. J Exp Med 2006;203 (4):1007-19.

Castaneda FE, Walia B, Vijay-Kumar M, Patel NR, Roser S, Kolachala VL, et al. Targeted deletion of metalloproteinase 9 attenuates experimental colitis in mice: central role of epithelial-derived MMP. Gastroenterology 2005;129 (6):1991-2008.

Chen EI, Kridel SJ, Howard EW, Li W, Godzik A, Smith JW. A unique substrate recognition profile for matrix metalloproteinase-2. J Biol Chem 2002;277 (6):4485-91.

Dubois B, Masure S, Hurtenbach U, Paemen L, Heremans H, van den Oord J, et al. Resistance of young gelatinase B-deficient mice to experimental autoimmune encephalomyelitis and necrotizing tail lesions. J Clin Invest 1999;104 (11):1507-15.

Farina AR, Mackay AR. Gelatinase B/MMP-9 in Tumour Pathogenesis and Progression. Cancers 2014;6 (1):240-96.

Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11. Available at http://globocan.iarc.fr. Accessed 09 July 2014. International Agency for Research on Cancer 2013.

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

Garg P, Vijay-Kumar M, Wang L, Gewirtz AT, Merlin D, Sitaraman SV. Matrix metalloproteinase-9-mediated tissue injury overrides the protective effect of matrix metalloproteinase-2 during colitis. American journal of physiology. Gastrointestinal and liver physiology 2009;296 (2):G175-84.

Hijova E. Matrix metalloproteinases: their biological functions and clinical implications. Bratislavske lekarske listy 2005;106 (3):127-32.

Hu J, Van den Steen PE, Sang QX, Opdenakker G. Matrix metalloproteinase inhibitors as therapy for inflammatory and vascular diseases. Nat Rev Drug Discov 2007;6 (6):480-98.

Hung HM, Wang SJ, O'Neill R. Statistical considerations for testing multiple endpoints in group sequential or adaptive clinical trials. J Biopharm Stat 2007;17 (6):1201-10.

Imai K, Yokohama Y, Nakanishi I, Ohuchi E, Fujii Y, Nakai N, et al. Matrix metalloproteinase 7 (matrilysin) from human rectal carcinoma cells. Activation of the precursor, interaction with other matrix metalloproteinases and enzymic properties. J Biol Chem 1995;270 (12):6691-7.

Itoh T, Matsuda H, Tanioka M, Kuwabara K, Itohara S, Suzuki R. The role of matrix metalloproteinase-2 and matrix metalloproteinase-9 in antibody-induced arthritis. J Immunol 2002;169 (5):2643-7.

Itoh T, Tanioka M, Matsuda H, Nishimoto H, Yoshioka T, Suzuki R, et al. Experimental metastasis is suppressed in MMP-9-deficient mice. Clinical & experimental metastasis 1999;17 (2):177-81.

Kang JH, Lee SI, Lim do H, Park KW, Oh SY, Kwon HC, et al. Salvage chemotherapy for pretreated gastric cancer: a randomized phase III trial comparing chemotherapy plus best supportive care with best supportive care alone. J Clin Oncol 2012;30 (13):1513-8.

Kridel SJ, Chen E, Kotra LP, Howard EW, Mobashery S, Smith JW. Substrate hydrolysis by matrix metalloproteinase-9. J Biol Chem 2001;276 (23):20572-8.

Lan KKG, Hu P, Proschan M. A Conditional Power Approach to the Evaluation of Predictive Power. Statistics in biopharmaceutical research 2012;1 (2):131-6.

Lausch E, Keppler R, Hilbert K, Cormier-Daire V, Nikkel S, Nishimura G, et al. Mutations in MMP9 and MMP13 determine the mode of inheritance and the clinical spectrum of metaphyseal anadysplasia. American journal of human genetics 2009;85 (2):168-78.

Li H, Mittal A, Makonchuk DY, Bhatnagar S, Kumar A. Matrix metalloproteinase-9 inhibition ameliorates pathogenesis and improves skeletal muscle regeneration in muscular dystrophy. Human molecular genetics 2009;18 (14):2584-98.

Miyazaki D, Nakamura A, Fukushima K, Yoshida K, Takeda S, Ikeda S. Matrix metalloproteinase-2 ablation in dystrophin-deficient mdx muscles reduces angiogenesis resulting in impaired growth of regenerated muscle fibers. Human molecular genetics 2011;20 (9):1787-99.

Mott JD, Werb Z. Regulation of matrix biology by matrix metalloproteinases. Current opinion in cell biology 2004;16 (5):558-64.

Naito Y, Yoshikawa T. Role of matrix metalloproteinases in inflammatory bowel disease. Molecular aspects of medicine 2005;26 (4-5):379-90.

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

Olson MW, Gervasi DC, Mobashery S, Fridman R. Kinetic analysis of the binding of human matrix metalloproteinase-2 and -9 to tissue inhibitor of metalloproteinase (TIMP)-1 and TIMP-2. J Biol Chem 1997;272 (47):29975-83.

Opdenakker G, Nelissen I, Van Damme J. Functional roles and therapeutic targeting of gelatinase B and chemokines in multiple sclerosis. Lancet neurology 2003;2 (12):747-56.

Santana A, Medina C, Paz-Cabrera MC, Diaz-Gonzalez F, Farre E, Salas A, et al. Attenuation of dextran sodium sulphate induced colitis in matrix metalloproteinase-9 deficient mice. World J Gastroenterol 2006;12 (40):6464-72.

Tamhane AC, Mehta CR, Liu L. Testing a primary and a secondary endpoint in a group sequential design. Biometrics 2010;66 (4):1174-84.

Thuss-Patience PC, Kretzschmar A, Bichev D, Deist T, Hinke A, Breithaupt K, et al. Survival advantage for irinotecan versus best supportive care as second-line chemotherapy in gastric cancer--a randomised phase III study of the Arbeitsgemeinschaft Internistische Onkologie (AIO). Eur J Cancer 2011;47 (15):2306-14.

Van Wart HE, Birkedal-Hansen H. The cysteine switch: a principle of regulation of metalloproteinase activity with potential applicability to the entire matrix metalloproteinase gene family. Proc Natl Acad Sci U S A 1990;87 (14):5578-82.

Vempati P, Karagiannis ED, Popel AS. A biochemical model of matrix metalloproteinase 9 activation and inhibition. J Biol Chem 2007;282 (52):37585-96.

Vu TH, Shipley JM, Bergers G, Berger JE, Helms JA, Hanahan D, et al. MMP-9/gelatinase B is a key regulator of growth plate angiogenesis and apoptosis of hypertrophic chondrocytes. Cell 1998;93 (3):411-22.

Waddell T, Verheij M, Allum W, Cunningham D, Cervantes A, Arnold D. Gastric cancer: ESMO-ESSO-ESTRO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2013;24 Suppl 6:vi57-63.

11. APPENDICES

Appendix 1. Investigator Signature Page

GILEAD SCIENCES, INC.
333 LAKESIDE DRIVE
FOSTER CITY, CA 94404

STUDY ACKNOWLEDGEMENT

A Phase 3 Randomized, Double-Blind, Placebo-Controlled Study to Evaluate the Efficacy and Safety of GS-5745 Combined with mFOLFOX6 as First Line Treatment in Patients with Advanced Gastric or Gastroesophageal Junction Adenocarcinoma

GS-US-296-1080, Amendment 6, 06 March 2017

This protocol has been approved by Gilead Sciences, Inc. The following signature documents this approval.

PPD

Medical Monitor

PPD

06-MAR-2017
Date

INVESTIGATOR STATEMENT

I have read the protocol, including all appendices, and I agree that it contains all necessary details for me and my staff to conduct this study as described. I will conduct this study as outlined herein and will make a reasonable effort to complete the study within the time designated.

I will provide all study personnel under my supervision copies of the protocol and access to all information provided by Gilead Sciences, Inc. I will discuss this material with them to ensure that they are fully informed about the drugs and the study.

Principal Investigator Name (Printed)

Signature

Date

Site Number

Appendix 2. Study Procedures Table

Study Phase	Screening	Randomization Cycle 1	Cycle 1	Each Cycle (28 days)		Cycle 2 (PK/Antibody/ Biomarker)	Cycle 3 (PK/Antibody/ Biomarker)	Cycle 4 & every 2 cycles (PK/Antibody/ Biomarker)	Cycle 5, 7 & Every 3 Cycles (PK/Antibody/ Biomarker)	EOT [r]	Every 8 weeks	Disease Progression	EOS[s]	30-day Safety Follow-up[t]	55-day Safety Follow-Up[u]	5-year Long Term Follow-Up
Cycle Day		**1[a]**	**15**	**1**	**15**	**1[h, i, j, k], 2[i], 8[i], 15[i]**	**1**	**1**	**1**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	
Window (day)	**-28**		**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**		**±7**		**±7**	**±7**	**±7**	**N/A**
Informed Consent	X															
Medical and Medication History	X															
Physical Examination[b]	X	X		X						X			X	X		
Vital Signs & Weight	X	X	X	X	X					X			X	X		
ECOG Performance Status[c]	X	X[c]		X						X			X	X		
12-lead ECG[v]	X	X	X	X	X					X			X			
Adverse events/ Concomitant medications[d]	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	
IXRS Registration	X	X	X	X	X					X			X			
Study Drug/Chemotherapy Administration[e]		X	X	X	X											
QLQ-C30 & STO22		X		X						X			X	X		
EQ-5D-3L		X		X						X			X	X		
Hematology	X	X	X	X	X					X			X	X		
Chemistry	X	X	X	X	X					X			X	X		

Study Phase	Screening	Randomization Cycle 1	Cycle 1	Each Cycle (28 days)		Cycle 2 (PK/Antibody/ Biomarker)	Cycle 3 (PK/Antibody/ Biomarker)	Cycle 4 & every 2 cycles (PK/Antibody/ Biomarker)	Cycle 5, 7 & Every 3 Cycles (PK/Antibody/ Biomarker)	EOT [r]	Every 8 weeks	Disease Progression	EOS[s]	30-day Safety Follow-up[t]	55-day Safety Follow-Up[u]	5-year Long Term Follow-Up
Cycle Day		**1[a]**	**15**	**1**	**15**	**1[h, i, j, k], 2[i], 8[i], 15[i]**	**1**	**1**	**1**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	
Window (day)	**-28**		**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**		**±7**		**±7**	**±7**	**±7**	**N/A**
Coagulation	X									X			X			
Urinalysis	X	X		X						X			X	X		
Pregnancy Test[f]	X	X		X						X			X	X		
HER2 Testing[g]	X															
GS-5745 PK[h]		X				X	X		X	X			X			
CCI																
Anti-GS-5745 Antibody[j]		X				X	X		X	X			X	X		
Blood biomarkers[k]		X	X			X		X								
Pharmacodynamic biomarker and immune monitoring[l]		X					X									
Biomarker sample for circulating tumor DNA[m]		X					X					X				
CCI																
Collect archival tumor tissue for biomarker testing[o]		X														

Study Phase	Screening	Randomization Cycle 1	Cycle 1	Each Cycle (28 days)		Cycle 2 (PK/Antibody/ Biomarker)	Cycle 3 (PK/Antibody/ Biomarker)	Cycle 4 & every 2 cycles (PK/Antibody/ Biomarker)	Cycle 5, 7 & Every 3 Cycles (PK/Antibody/ Biomarker)	EOT [r]	Every 8 weeks	Disease Progression	EOS[s]	30-day Safety Follow-up[t]	55-day Safety Follow-Up[u]	5-year Long Term Follow-Up
Cycle Day		**1[a]**	**15**	**1**	**15**	**1[h, i, j, k], 2[i], 8[i], 15[i]**	**1**	**1**	**1**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	**N/A**	
Window (day)	**-28**		**±3**	**±3**	**±3**	**±3**	**±3**	**±3**	**±3**		**±7**		**±7**	**±7**	**±7**	**N/A**
CT or MRI & Treatment Response Assesment[p]	X										X		X			
Core Biopsy[q]												X				
Overall Survival and Other Antitumor Therapy																X (Every 3 months)

a Cycle 1 Day 1 (C1D1) must occur within 3 days following randomization.
b Complete physical examination (PE) to be performed at Screening, EOT and EOS. A modified PE capturing changes from prior exams will be performed at subsequent visits. Height is required at Screening only.
c ECOG performance status on Cycle 1 Day 1 may be waived if has been conducted during screening within 4 days of Cycle 1 Day 1.
d Adverse events will be assessed at each clinic visit from Screening up to and including the 55-day Safety Follow-up visit. Concomitant medications will be recorded at each clinic visit from Screening up to and including the 55-day Safety Follow-up visit or EOS visit whichever is later.
e Study drug (GS-5745/placebo) will be dosed on Days 1 and 15 of each cycle over 30 (± 5) min. Chemotherapy (mFOLFOX6) will be administered after study drug and will consist of l-LV 200mg/m^2 or dl-LV 400mg/m^2 and oxaliplatin 85 mg/m^2 followed by bolus 5-FU 400mg/m^2 and a 46 hour infusion of 5-FU 2400 mg/m^2. After 6 cycles of mFOLFOX6, the oxaliplatin component will be discontinued. Leucovorin and 5-FU will be continued as maintenance therapy, in combination with GS-5745/placebo, in the absence of disease progression or toxicity warranting discontinuation of therapy.
f If applicable (females of child bearing potential). Serum pregnancy testing will be conducted at Screening. Urine pregnancy testing will be conducted pre-dose on Day 1 of each cycle, at EOT, and EOS.
g Tumor tissue will be tested for HER2 status, if unknown, with an approved IHC and ISH kit.
h Plasma samples will be collected for GS-5745 PK at 30(± 15) min after the end of infusion on Day 1 of Cycle 1, prior to dosing and 30(± 15) min after the end of infusion on Day 1 of cycles 2, 3, 5, 7, and every 3 cycles thereafter, at the EOT and EOS visits.
i CCI [redacted]

j Serum samples for anti-GS-5745 antibody will be collected prior to dosing on Day 1 of cycles 1, 2, 3, 5, 7, and every 3 cycles thereafter, EOT, EOS, and 30-day Safety Follow Up visits.
k Blood biomarkers will be collected prior to dosing on Cycle 1 Day 1, Cycle 1 Day 15, and Day 1 of Cycles 2 and 4 and every 2 cycles thereafter
l Samples for pharmacodynamic biomarkers and immune-monitoring will be collected prior to dosing on Day 1 of cycles 1and 3 only
m Biomarker Sample for circulating tumor DNA will be collected predose on Day 1 of cycles 1 and 3 and at disease progression.
n CCI [redacted]
o Archival tumor tissue block will be collected and shipped to central laboratory for sectioning after Day 1 of Cycle 1. If an archival block was submitted for HER2 testing, and the subject is randomized, the block will be stored for biomarker analysis. However, if slides were submitted for HER2 testing, additional slides will be required after Day 1 of Cycle 1 for biomarker analysis.
p Tumor evaluation by CT or MRI will be performed during screening and approximately every 8 weeks (starting from Cycle 1 Day 1) regardless of cycle number or dose interruption. Scan at EOS visit is not necessary if restaging scan is performed within the prior 8 weeks. Treatment response assessment will be per RECIST v1.1.
q For subjects who stop study treatment in the absence of disease progression (eg. experienced unexpected toxicity) and remain on study for follow up for progression-free survival, tumor evaluation by CT or MRI should continue approximately every 8 weeks until disease progression or initiation of non-study specific anti-neoplastic therapy in the absence of progression, whichever occurs earlier.
r A core biopsy will be performed after disease progression, if medically feasible. The sample should be collected by the last clinic visit on study ie EOS or 30 Day Safety Follow up visit.
s End of treatment (EOT) assessments will be completed only by subjects who discontinue all treatment prior to disease progression. These assessments should be completed as soon as possible after the decision is made. Every attempt should be made to keep the subject in the study and continue to perform tumor evaluation by CT or MRI every 8 weeks until disease progression.
t End of study (EOS) assessments will be completed when the subject meets at least 1 of the criteria for study discontinuation (Section 3.6).
u The 30-day safety follow up visit will be performed following the last dose of GS-5745/placebo, or all components mFOLFOX6, whichever is discontinued later.
v The 55-day safety follow up phone visit will be performed following the last dose of GS-5745/placebo, or all components mFOLFOX6, whichever is discontinued later.
w A single 12-lead electrocardiogram (ECG) will be collected at Screening, Day 1 of each cycle, and at the EOT and EOS visits, per standard practice. Additional ECGs following administration of oxaliplatin will be collected at visits when oxaliplatin is administered after the infusion is completed (i.e., Days 1 and 15 of each cycle, up to a maximum of six 28-day cycles) as required per country label guidance regarding cardiac monitoring.

Appendix 3. Pregnancy Precautions, Definition for Female of Childbearing Potential, and Contraceptive Requirements

1) Pregnancy and Contraception Requirements for Males and Females of Childbearing Potential

The risks of treatment with GS-5745 during pregnancy have not been evaluated in humans. The potential for genotoxicity and embryofetal toxicity is considered to be low based on nonclinical toxicological studies. In both the rat and rabbit definitive embryofetal developmental toxicity studies, there were no GS-5745 related effects on embryofetal survival and growth and no fetal anomalies. In a fertility study in male and female rats, no test article-related effects on reproductive performance and intrauterine survival were observed at any dosage level. Please refer to the latest version of the investigator's brochure for additional information on the study drug of GS-5745. Please refer to the regional prescribing information for information on the potential risks of treatment with mFOLFOX6.

2) Definition of Female of Childbearing Potential

For the purposes of this study, a female subject of childbearing potential is a woman who has not had a hysterectomy, bilateral oophorectomy, or medically documented ovarian failure. This definition includes pubertal females regardless of whether or not she has had a menses (premenarchal, Tanner Stage 3) and perimenopausal women who have had a spontaneous menses in the last 12 months. A woman who has had a tubal sterilization is considered to be of childbearing potential.

- Women ≤ 54 years of age with amenorrhea of any duration will be considered to be of childbearing potential unless they have had a hysterectomy, bilateral oophorectomy, or medically documented ovarian failure.

- Women > 54 years of age with cessation (for ≥12 months) of previously occurring menses due to ovarian failure will not be considered to be of childbearing potential.

3) Contraceptive Requirements for Females

Female subjects of childbearing potential must agree to use protocol specified method(s) of contraception from the screening/randomization visit throughout the study period, 90 days following the last dose of study drug (GS-5745/placebo) and for 4 months after the last dose of oxaliplatin or 6 months after the last dose of 5-FU whichever occurs later, unless the subject chooses continuous heterosexual abstinence as a lifestyle choice. The investigator should counsel subjects on the protocol specified method(s) for avoiding pregnancy during the study. These methods are recommended due to the low failure rate (ie, less than 1% per year). See the protocol specified contraceptive methods listed below.

Female study subjects who are not heterosexually active must have periodic confirmation of continued abstinence from heterosexual intercourse and regular pregnancy testing while taking study drug. The investigator should counsel subjects on the protocol specified method(s) for avoiding pregnancy in case the subject chooses to engage in heterosexual intercourse.

Female subjects of childbearing potential must have a negative serum pregnancy test at screening and a negative urine pregnancy test at Day 1 of each cycle prior to receiving the dose of study drug. Lactating females must discontinue nursing before study drug administration.

Consistent and correct use of 1 of the following methods of birth control listed below:

- Intrauterine device (IUD) with a failure rate of <1% per year
- Tubal sterilization
- Essure micro-insert system (provided confirmation of success 3 months after procedure). This is not yet approved in Japan.
- Vasectomy in the male partner (provided that the partner is the sole sexual partner and had confirmation of surgical success 3 months after procedure)

Female subjects must also refrain from egg donation and in vitro fertilization during treatment and until at least 90 days following the last dose of GS-5745 and for 4 months after the last dose of oxaliplatin or 6 months following the last dose of 5-FU whichever occurs later.

4) Contraceptive Requirements for Males

Male subjects must agree to use condoms and avoid sperm donation from the screening/randomization visit throughout the study period, 90 days after administration of the last dose of study drug (GS-5745/placebo) and for 6 months after the last dose of oxaliplatin or 6 months following the last dose of 5-FU whichever occurs later.

5) Contraception Requirements for Oxaliplatin

Female subjects of childbearing potential must use a protocol recommended method of contraception from the screening visit throughout the study treatment period and for 4 months after the last dose of oxaliplatin unless the subject chooses continuous heterosexual abstinence as a lifestyle-choice. Male subjects of reproductive potential, must use a protocol-recommended method of contraception and to refrain from sperm donation from the start of study drug, throughout the study treatment period, for 6 months following the last dose of oxaliplatin.

6) Unacceptable Birth Control Methods

Birth control methods that are unacceptable include periodic abstinence (e.g., calendar, ovulation, symptothermal, post-ovulation methods), withdrawal (coitus interruptus), spermicides only, and lactational amenorrhea method (LAM). Female condom and male condom should not be used together.

7) Procedures to be Followed in the Event of Pregnancy

Subjects should be instructed to notify the investigator if they become pregnant at any time during the study, and if they become pregnant within 90 days following the last dose the study drug (GS-5745/placebo) and within 4 months (or 6 months for the partner of male subjects) of

the last dose of oxaliplatin, or 6 months after the last dose of 5-FU whichever occurs later. Subjects who become pregnant or who suspect that they are pregnant during the study must report the information to the investigator and discontinue study drug immediately. The investigator should report all pregnancies to the CRO Safety Department using the pregnancy report form within 24 hours of becoming aware of the pregnancy. The investigator should counsel the subject regarding the possible effects of prior study drug exposure on the fetus and the need to inform the study site of the outcome of the pregnancy. Subjects whose partner has become pregnant or suspects she is pregnant during the study must report the information to the investigator. Instructions for reporting partner pregnancy, pregnancy and pregnancy outcome are outlined in Section 7.7.2.1.

Appendix 4. ECOG Performance Status

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

Reference for ECOG {Oken 1982}

Appendix 5. Dose Modification Tables for mFOLFOX6

Appendix Table 1. Recommended Dose Modifications for Oxaliplatin+5-Fluorouracil/Leucovorin[a]

NCI CTCAE v. 4.03 System Organ Class[b]	Adverse Event	Dose Level for Subsequent Cycles Based on Interval Adverse Events	At Time of Retreatment
All Adverse Events < 1		Maintain dose level	Maintain dose level
Blood and lymphatic system disorders:	***Hemolytic uremic syndrome*** (HUS)[c]		
	> Grade 3	Discontinue oxaliplatin	Discontinue oxaliplatin
Investigations:	***Neutrophil count decreased***		If ANC < 1500 at start of cycle, hold and check weekly then treat based on interval adverse event. If ANC < 1500 after 4 weeks, discontinue therapy.
	Grade 1 (ANC < LLN - 1500/mm^3)	Maintain dose level	
	Grade 2 (ANC < 1500 - 1000/mm^3)	Maintain dose level	
	Grade 3 (ANC < 1000 - 500/mm^3)	Omit bolus 5-FU and decrease 1 oxaliplatin dose level	
	Grade 4 (ANC < 500/mm^3)	Omit bolus 5-FU and decrease both infusion 5-FU and oxaliplatin 1 dose level	
	Platelet count decreased		If PLT < 75,000 at start of cycle, hold and check weekly then treat based on interval adverse event. If PLT < 75,000 after 4 weeks, discontinue therapy.
	Grade 1 (PLT < LLN - 75,000/mm3)	Maintain dose level	
	Grade 2 (PLT < 75,000 - 50,000/mm3)	Maintain dose level	
	Grade 3 (PLT < 50,000 - 25,000/mm3)	Omit bolus 5-FU and decrease 1 oxaliplatin dose level	
	Grade 4 (PLT < 25,000/mm3)	Omit bolus 5-FU and decrease 2 oxaliplatin dose levels	

NCI CTCAE v. 4.03 System Organ Class[b]	Adverse Event	Dose Level for Subsequent Cycles Based on Interval Adverse Events	At Time of Retreatment
Gastrointestinal disorders:	***Diarrhea***		If Grade ≥ 2 at start of cycle, hold and check weekly then treat based on interval adverse event. If Grade ≥ 2 after 4 weeks, discontinue therapy.
	Grade 1, 2	Maintain dose level	
	Grade 3	Decrease one 5-FU dose level	
	Grade 4	Decrease both 5-FU and oxaliplatin 1 dose level	
	Mucositis oral		
	Grade 1, 2	Maintain dose level	
	Grade 3	Decrease one 5-FU dose level	
	Grade 4	Decrease one 5-FU dose level	
	Vomiting		
	Grades 1, 2	Maintain dose level	
	Grade 3	Decrease1 oxaliplatin dose level	
	Grade 4	Decrease both 5-FU and oxaliplatin 1 dose level	
Metabolism and nutrition disorders:	***Hypomagnesemia***	Note: Dose reduction is not required for hypomagnesemia unless symptoms are present. If Grade ≥ 2 after 4 weeks, discontinue therapy.	
Neurology:	Do not use CTCAE.	See Appendix Table 2 for adverse event scale and oxaliplatin dose modifications.	
Respiratory, thoracic, and mediastinal disorders:	***Cough*** ≥ Grade 3 ***Dyspnea*** ≥ Grade 3 ***Hypoxia*** ≥ Grade 3 ***Pneumonitis*** ≥ Grade 3	Hold oxaliplatin until interstitial lung disease is ruled out. If interstitial lung disease, oxaliplatin should be permanently discontinued.	
Other non-hematologic adverse events[d, e]:	Grades 1, 2	Maintain dose level	
	Grades 3, 4	Decrease offending agent 1 dose level	

a The dose of leucovorin will not be adjusted due to adverse event. It should remain at 400 mg/m^2 dl-leucovorin or 200 mg/m^2 of l-leucovorin for all courses. Leucovorin will be given immediately prior to each 5-FU dose; thus, if 5-FU is delayed, leucovorin will be delayed. Leucovorin doses may be adjusted per institutional guidelines in the event of a supply shortage.

b For ≤ NCI CTCAE v. 4.03 Grade 2 toxicity not described, maintain dose level of agent.

c Recommended evaluation of suspected HUS: Evaluation should include CBC differential, platelets, PT, PTT, fibrinogen, FDP, Anti thrombin III, Von Willebrand factor, anti-nuclear antibody, rheumatoid factor, Compliment Cascade C3, C4, and CH50, anti-platelet antibodies, platelet-associated IgG, and circulating immune complexes. Renal evaluation should include creatinine, BUN, and urinalysis with microscopic examination. Other laboratory and hematologic evaluations as appropriate should also be obtained, including peripheral blood smear and free hemoglobin.

d Exceptions: fatigue, anorexia, nausea/vomiting if can be controlled by antiemetics, and viral infections.

e Dose modifications for other non-hematologic adverse events at the start of subsequent courses of therapy, and at time of retreatment are also based on NCI CTCAE v. 4.03 criteria.

Appendix Table 2. Oxaliplatin[a] Dose Modifications for Non-CTCAE Neurologic Adverse Events

Adverse Events	Duration of Adverse Event		Persistent[b] Between Cycles
	1 - 7 Days	> 7 Days	
Paresthesias/Dysesthesias			
Paresthesias/dysesthesias[c] of short duration that resolve and do not interfere with function (Grade 1)	No change	No change	No change
Paresthesias/dysesthesias[c] interfering with function, but not activities of daily living (ADL) (Grade 2)	No change	No change	Decrease 1 oxaliplatin dose level
Paresthesias/dysesthesias[c] with pain or with functional impairment that also interfere with ADL (Grade 3)	1st time: Decrease 1 oxaliplatin dose level 2nd time: Decrease 1 oxaliplatin dose level	1st time: Decrease 1 oxaliplatin dose level 2nd time: Decrease 1 oxaliplatin dose level	Discontinue
Persistent paresthesias/dysesthesias that are disabling or life-threatening (Grade 4)	Discontinue	Discontinue	Discontinue
Laryngeal Dysesthesias (investigator discretion used for grading):			
Grade 0 = none; Grade 1 = mild	No change	Increase duration of infusion to 6 hours	Increase duration of infusion to 6 hours
Grade 2 = moderate (Also recommended is administration of benzodiazepine and patient education. Management of patient if ≥ Grade 2 laryngeal dysesthesias occurs while treatment is being administered.)	Stop oxaliplatin infusion. Administer benzodiazepine and give patient reassurance. At the discretion of the investigator, the infusion can be restarted at 1/3 the original rate of infusion.		
Grade 3 = severe			

a If oxaliplatin is discontinued, continue other study agents unless adverse events preclude their continuation.
b Not resolved by the beginning of the next cycle.
c May be cold-induced.

Appendix 6. Details for Testing Efficacy Endpoints

This appendix provides details for testing multiple efficacy endpoints at interim and final analyses.

The final analysis will occur when 286 overall survival (OS) events have been observed. Two formal interim analyses are planned: the futility interim analysis and the efficacy interim analysis will be performed when approximately 33.3% and 66.7% of OS events have occurred, respectively.

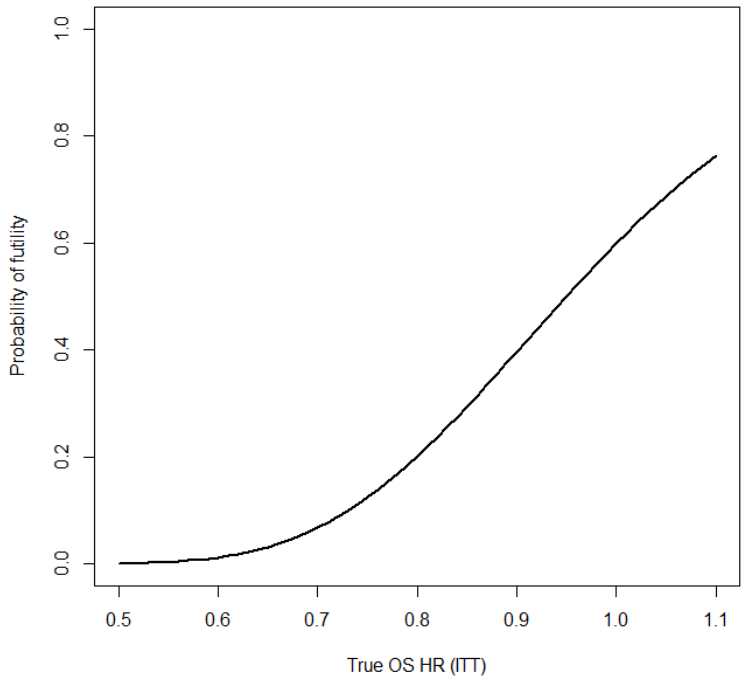
Futility Interim Analysis

The end point of the futility interim analysis is OS. The futility analysis will occur approximately after 95 OS events (33.3% of the information relative to the final analysis) have been observed in the ITT population. The analysis will perform a stratified log-rank test for OS in the ITT population. A non-binding futility rule will be implemented. Based on the analysis results, the DMC may recommend terminating the study for lack of efficacy if the predictive power (PP) on OS in ITT population is < 14%. Otherwise, the study will be continued.

Predictive power is defined as a weighted average of the conditional power; the weighting function is determined by the sampling distribution of the observed hazard ratio based on the data at the interim analysis {Lan 2012}. When exactly 95 OS events are observed at the interim, this futility rule of PP <14% corresponds to observing an OS hazard ratio (HR) > 0.95 in the ITT population.

Appendix Figure 1 presents the operating characteristics of futility rule at the interim of 95 OS events with futility boundary of 0.95. When the true hazard ratio is 0.7, the probability of terminating the study early is 7%. On the other hand, when the true hazard ratio is 1, indicating lack of efficacy, the probability of terminating the study early is 60%.

Appendix Figure 1. Probability of Early Stop due to Futility



Efficacy Interim and Final Analysis

In the efficacy interim and final analysis, the primary and secondary endpoints will be tested sequentially in the following gatekeeping order: the primary OS endpoint, then the secondary PFS endpoint, and finally the secondary ORR endpoint. Specifically:

- At the efficacy interim analysis, the OS event point will be tested at one-sided alpha level of 0.006. If the OS endpoint is not rejected, the study will continue to final analysis. If the OS endpoint is rejected, the DMC would recommend early stop for efficacy. Furthermore, the PFS endpoint will be tested at one-sided alpha level of 0.016. Only if the PFS endpoint is rejected, the ORR endpoint will be tested at one-sided alpha level of 0.016.

- At the final analysis, the OS event point will be tested at one-sided alpha level of 0.023. If the OS endpoint is rejected, the PFS endpoint will be tested at one-sided alpha level of 0.016. If the PFS endpoint is also rejected, the ORR endpoint will be tested at one-sided alpha level of 0.016.

This gatekeeping testing strategy is summarized in the table below. Overall, this testing strategy employs O'Brian-Fleming boundary for the primary OS endpoint and the Pocock type boundary for the secondary PFS/ORR endpoints. The testing strategy controls the overall one-sided family-wise type I error to be at 0.025, equivalent to two sided 0.05 by appropriately adjusting for multiplicity in the efficacy interim and final analyses.

Appendix Table 3. Gate-keeping Decision Boundaries for Efficacy Analysis

Efficacy Analysis	Event (%)	One-sided Decision Boundary		
		OS	PFS (After OS boundary is crossed)	ORR (After PFS boundary is crossed)
Interim	191 (66.7%)	0.006	0.016	0.016
Final	286 (100%)	0.023	0.016	0.016

In the following scenarios we will provide detailed justification for this testing strategy. During efficacy interim and final analysis, type I error could occur in the following non-overlapping scenarios:

1) Reject H_{01} when H_{01} is true

2) Reject H_{01} and H_{02} when H_{11} and H_{02} is true

3) Reject H_{01}, H_{02} and H_{03} when H_{11}, H_{12}, and H_{03} is true

where H_{ij} (i=0: null, i=1: alternative) are the hypothesis for the jth endpoint (1 = primary OS, 2 = secondary PFS, 3 = secondary ORR).

Scenario 1

The type I error rate is controlled by the O'Brien-Fleming (OF) boundaries on the primary OS endpoint.

Scenario 2

We herein justify that the use of 1-sided significance level of 0.016 for the testing of secondary PFS endpoint at both the interim and final analyses will control the Type I error rate at 0.025 one-sided. We follow the methods in Hung et al (2007) {Hung 2007} and Tamhane et al (2010) {Tamhane 2010}.

Per Hung et al {Hung 2007}, the one-sided type I error rate in Scenario 2 in the hierarchical testing strategy is:

$$\alpha_2 = \Pr(T_{11} > C_{11}, T_{21} > C_{21} \mid H_{02}) + \Pr(T_{11} \le C_{11}, T_{12} > C_{12}, T_{22} > C_{22} \mid H_{02}) \qquad (1)$$

where T_{jk} and C_{jk} are the test statistic and critical value, respectively, for the *j*th endpoint (1 = primary OS, 2 = secondary PFS) and the *k*th analysis (1=interim, 2=final). Accordingly, (1) can also be written as:

$$\alpha_2 = \Pr(Z_{11} > z_{11}, Z_{21} > z_{21} \mid H_{02}) + \Pr(Z_{11} \le z_{11}, Z_{12} > z_{12}, Z_{22} > z_{22} \mid H_{02}) \qquad (2)$$

where Z_{jk} and z_{jk} are the corresponding test statistics and critical values expressed in the setting of the standard normal distribution (ie, after standardization). Equation (2) can be evaluated analytically and we prove in the following that in the scenario 2, under the proposed the testing strategy, the one-sided type I error rate is controlled at 0.025.

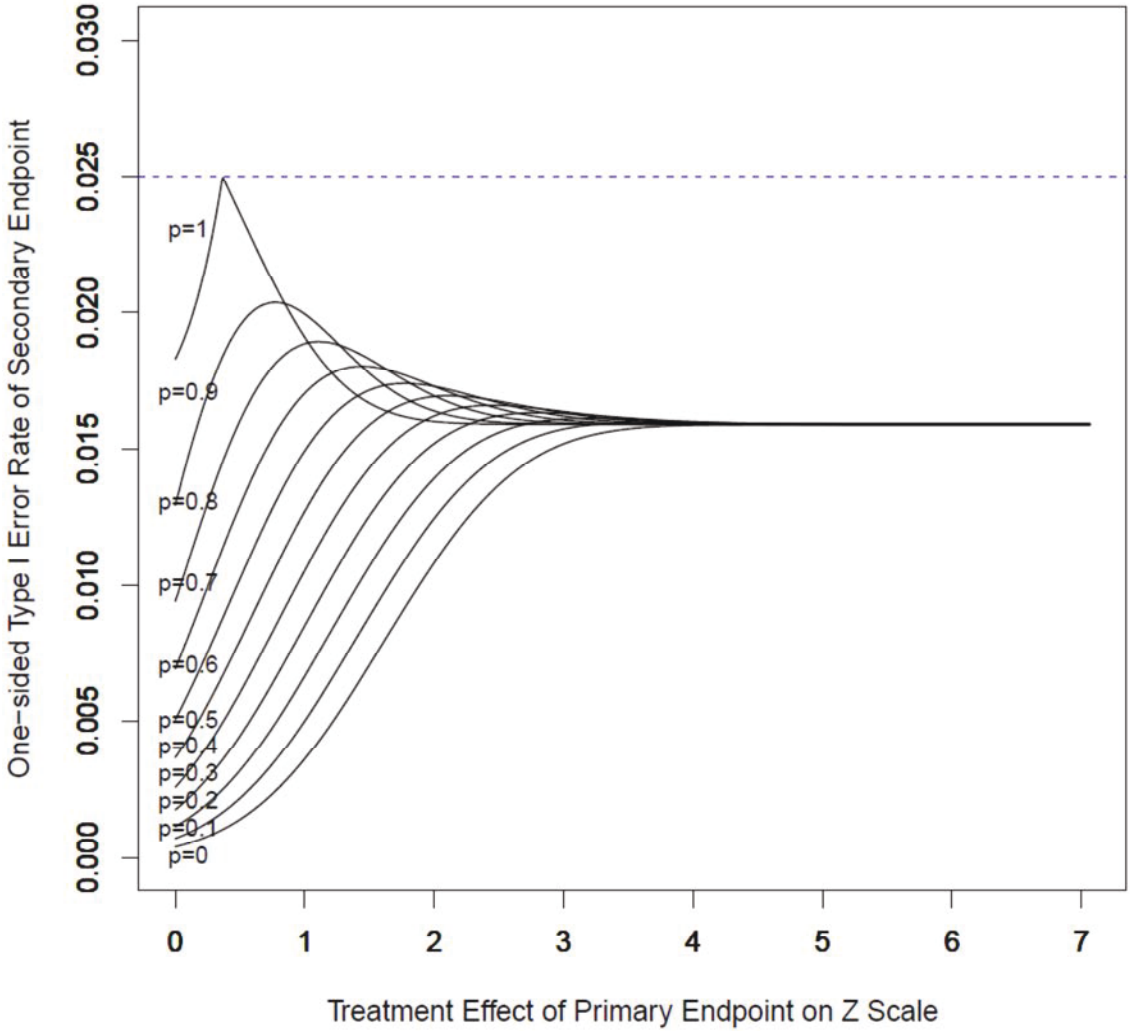
It can be derived that $\widetilde{Z} = (Z_{11}, Z_{21}, Z_{12}, Z_{22})$ follows the multivariate normal distribution below:

$$\begin{pmatrix} Z_{11} \\ Z_{21} \\ Z_{12} \\ Z_{22} \end{pmatrix} \sim MVN\left(\begin{pmatrix} \mu_{11} \\ \mu_{21} \\ \mu_{12} \\ \mu_{22} \end{pmatrix}, \begin{pmatrix} 1 & \rho & \sqrt{t} & \rho\sqrt{t} \\ \rho & 1 & \rho\sqrt{t} & \sqrt{t} \\ \sqrt{t} & \rho\sqrt{t} & 1 & \rho \\ \rho\sqrt{t} & \sqrt{t} & \rho & 1 \end{pmatrix} \right) \qquad (3)$$

where (μ_{11}, μ_{12}) and (μ_{21}, μ_{22}) are the normalized mean treatment effects on the primary and secondary endpoints, respectively, for data up to the interim and final analysis. The mean treatment effects between the interim and final analysis can be related by the information fraction t: $\mu_{j1} = \mu_{j2}/\sqrt{t}$ where $j = 1, 2$. Under the assumption of Scenario 2, the mean treatment effect on the primary OS endpoint (μ_{11}, μ_{12}) is positive representing the alternative hypothesis, and the mean treatment effect on secondary PFS endpoint (μ_{21}, μ_{22}) is set to (0, 0) under the null hypothesis. ρ denotes the correlation between the primary and the secondary test statistics at each analysis stage, ranging from 0 to 1, assuming a non-negative correlation. $t = 0.667$ denotes the planned information fraction at the interim efficacy analysis. The one-sided OF boundaries for the primary endpoint are given by $z_{11} = z_{0.006}$ and $z_{12} = z_{0.023}$. Thus, equation (2) can be evaluated using the pmvnorm function in the R package mvtnorm based on the mean and covariance structure specified in equation (3).

Appendix Figure 2 shows the analytically calculated type I error rate in the second scenario, which is the error rate for the secondary endpoint. It demonstrates that the use of a 1-sided significance level of 0.016 (equivalent to 2-sided level 0.032) for the testing of secondary endpoints at both the interim and final analysis controls the overall type I error rate at the 1-sided level 0.025 regardless of the correlation between the primary and secondary test statistics and the magnitude of treatment effect of the primary endpoint.

Appendix Figure 2. Type I Error Rate of Secondary Endpoint in a Group Sequential Design with OF Boundaries for Primary Endpoint and 1-sided Significance Level of 0.016 for the Secondary Endpoints with 1 Interim and 1 Final Analysis



Scenario 3

The one-sided type I error rate of Scenario 3 is bounded by the following inequality:

$$\begin{aligned} &P(Reject\ H_{01}, Reject\ H_{02}, Reject\ H_{03} | H_{11}, H_{12}, H_{03}) \\ &\leq P(Reject\ H_{01}, Reject\ H_{03} | H_{11}, H_{12}, H_{03}) \\ &= P(Reject\ H_{01}, Reject\ H_{03} | H_{11}, H_{03}) \end{aligned}$$

Applying the same argument in Scenario 2 (with PFS replaced by ORR), the type I error rate can be protected by using one-sided significance level of 0.016 for the testing of secondary ORR endpoint at both the interim and final analyses.

Appendix 7. Revised RECIST Guideline (version 1.1)

EUROPEAN JOURNAL OF CANCER 45 (2009) 228 247

available at www.sciencedirect.com

journal homepage: www.ejconline.com

New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1)

E.A. Eisenhauer[a,*], P. Therasse[b], J. Bogaerts[c], L.H. Schwartz[d], D. Sargent[e], R. Ford[f], J. Dancey[g], S. Arbuck[h], S. Gwyther[i], M. Mooney[g], L. Rubinstein[g], L. Shankar[g], L. Dodd[g], R. Kaplan[j], D. Lacombe[c], J. Verweij[k]

[a]National Cancer Institute of Canada Clinical Trials Group, 10 Stuart Street, Queen's University, Kingston, ON, Canada
[b]GlaxoSmithKline Biologicals, Rixensart, Belgium
[c]European Organisation for Research and Treatment of Cancer, Data Centre, Brussels, Belgium
[d]Memorial Sloan Kettering Cancer Center, New York, NY, USA
[e]Mayo Clinic, Rochester, MN, USA
[f]RadPharm, Princeton, NJ, USA
[g]Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, MD, USA
[h]Schering Plough, Kenilworth, NJ, USA
[i]East Surrey Hospital, Redhill, Surrey, UK
[j]National Cancer Research Network, Leeds, UK
[k]Erasmus University Medical Center, Rotterdam, The Netherlands

ARTICLE INFO

Article history:
Received 17 October 2008
Accepted 29 October 2008

Keywords:
Response criteria
Solid tumours
Guidelines

ABSTRACT

Background: Assessment of the change in tumour burden is an important feature of the clinical evaluation of cancer therapeutics: both tumour shrinkage (objective response) and disease progression are useful endpoints in clinical trials. Since RECIST was published in 2000, many investigators, cooperative groups, industry and government authorities have adopted these criteria in the assessment of treatment outcomes. However, a number of questions and issues have arisen which have led to the development of a revised RECIST guideline (version 1.1). Evidence for changes, summarised in separate papers in this special issue, has come from assessment of a large data warehouse (>6500 patients), simulation studies and literature reviews.

Highlights of revised RECIST 1.1: Major changes include: *Number of lesions to be assessed*: based on evidence from numerous trial databases merged into a data warehouse for analysis pur poses, the number of lesions required to assess tumour burden for response determination has been reduced from a maximum of 10 to a maximum of five total (and from five to two per organ, maximum). *Assessment of pathological lymph* nodes is now incorporated: nodes with a *short* axis of ⩾15 mm are considered measurable and assessable as target lesions. The short axis measurement should be included in the sum of lesions in calculation of tumour response. Nodes that shrink to <10 mm short axis are considered normal. *Confirma tion of response* is required for trials with response primary endpoint but is no longer required in randomised studies since the control arm serves as appropriate means of inter pretation of data. *Disease progression* is clarified in several aspects: in addition to the previ ous definition of progression in target disease of 20% increase in sum, a 5 mm absolute increase is now required as well to guard against over calling PD when the total sum is very

* *Corresponding author:* Tel.: PPD ; fax: PPD .
E mail address: PPD (E A. Eisenhauer).
0959 8049/$ see front matter
doi:10.1016/j.ejca.2008.10.026

small. Furthermore, there is guidance offered on what constitutes 'unequivocal progres sion' of non measurable/non target disease, a source of confusion in the original RECIST guideline. Finally, a section on detection of new lesions, including the interpretation of FDG PET scan assessment is included. *Imaging guidance*: the revised RECIST includes a new imaging appendix with updated recommendations on the optimal anatomical assess ment of lesions.

Future work: A key question considered by the RECIST Working Group in developing RECIST 1.1 was whether it was appropriate to move from anatomic unidimensional assessment of tumour burden to either volumetric anatomical assessment or to functional assessment with PET or MRI. It was concluded that, at present, there is not sufficient standardisation or evidence to abandon anatomical assessment of tumour burden. The only exception to this is in the use of FDG PET imaging as an adjunct to determination of progression. As is detailed in the final paper in this special issue, the use of these promising newer approaches requires appropriate clinical validation studies.

1. Background

1.1. History of RECIST criteria

Assessment of the change in tumour burden is an important feature of the clinical evaluation of cancer therapeutics. Both tumour shrinkage (objective response) and time to the devel opment of disease progression are important endpoints in cancer clinical trials. The use of tumour regression as the endpoint for phase II trials screening new agents for evi dence of anti tumour effect is supported by years of evi dence suggesting that, for many solid tumours, agents which produce tumour shrinkage in a proportion of patients have a reasonable (albeit imperfect) chance of subsequently demonstrating an improvement in overall survival or other time to event measures in randomised phase III studies (re viewed in [1 4]). At the current time objective response car ries with it a body of evidence greater than for any other biomarker supporting its utility as a measure of promising treatment effect in phase II screening trials. Furthermore, at both the phase II and phase III stage of drug development, clinical trials in advanced disease settings are increasingly utilising time to progression (or progression free survival) as an endpoint upon which efficacy conclusions are drawn, which is also based on anatomical measurement of tumour size.

However, both of these tumour endpoints, objective re sponse and time to disease progression, are useful only if based on widely accepted and readily applied standard crite ria based on anatomical tumour burden. In 1981 the World Health Organisation (WHO) first published tumour response criteria, mainly for use in trials where tumour response was the primary endpoint. The WHO criteria introduced the con cept of an overall assessment of tumour burden by summing the products of bidimensional lesion measurements and determined response to therapy by evaluation of change from baseline while on treatment.[5] However, in the decades that followed their publication, cooperative groups and pharma ceutical companies that used the WHO criteria often 'modi fied' them to accommodate new technologies or to address areas that were unclear in the original document. This led to confusion in interpretation of trial results[6] and in fact, the application of varying response criteria was shown to lead to very different conclusions about the efficacy of the same regimen.[7] In response to these problems, an International Working Party was formed in the mid 1990s to standardise and simplify response criteria. New criteria, known as RECIST (Response Evaluation Criteria in Solid Tumours), were pub lished in 2000.[8] Key features of the original RECIST include definitions of minimum size of measurable lesions, instruc tions on how many lesions to follow (up to 10; a maximum five per organ site), and the use of unidimensional, rather than bidimensional, measures for overall evaluation of tu mour burden. These criteria have subsequently been widely adopted by academic institutions, cooperative groups, and industry for trials where the primary endpoints are objective response or progression. In addition, regulatory authorities accept RECIST as an appropriate guideline for these assessments.

1.2. Why update RECIST?

Since RECIST was published in 2000, many investigators have confirmed in prospective analyses the validity of substituting unidimensional for bidimensional (and even three dimen sional) based criteria (reviewed in [9]). With rare exceptions (e.g. mesothelioma), the use of unidimensional criteria seems to perform well in solid tumour phase II studies.

However, a number of questions and issues have arisen which merit answers and further clarity. Amongst these are whether fewer than 10 lesions can be assessed without affecting the overall assigned response for patients (or the conclusion about activity in trials); how to apply RECIST in randomised phase III trials where progression, not response, is the primary endpoint particularly if not all patients have measurable disease; whether or how to utilise newer imag ing technologies such as FDG PET and MRI; how to handle assessment of lymph nodes; whether response confirmation is truly needed; and, not least, the applicability of RECIST in trials of targeted non cytotoxic drugs. This revision of the RECIST guidelines includes updates that touch on all these points.

1.3. *Process of RECIST 1.1 development*

The RECIST Working Group, consisting of clinicians with expertise in early drug development from academic research organisations, government and industry, together with imaging specialists and statisticians, has met regularly to set the agenda for an update to RECIST, determine the evidence needed to justify the various changes made, and to review emerging evidence. A critical aspect of the revision process was to create a database of prospectively documented solid tumour measurement data obtained from industry and academic group trials. This database, assembled at the EORTC Data Centre under the leadership of Jan Bogaerts and Patrick Therasse (co authors of this guideline), consists of >6500 patients with >18,000 target lesions and was utilised to investigate the impact of a variety of questions (e.g. number of target lesions required, the need for response confirmation, and lymph node measurement rules) on response and progression free survival outcomes. The results of this work, which after evaluation by the RECIST Working Group led to most of the changes in this revised guideline, are reported in detail in a separate paper in this special issue.[10] Larry Schwartz and Robert Ford (also co authors of this guideline) also provided key databases from which inferences have been made that inform these revisions.[11]

The publication of this revised guideline is believed to be timely since it incorporates changes to simplify, optimise and standardise the assessment of tumour burden in clinical trials. A summary of key changes is found in Appendix I. Because the fundamental approach to assessment remains grounded in the anatomical, rather than functional, assessment of disease, we have elected to name this version RECIST 1.1, rather than 2.0.

1.4. *What about volumetric or functional assessment?*

This raises the question, frequently posed, about whether it is 'time' to move from anatomic unidimensional assessment of tumour burden to either volumetric anatomical assessment *or* to functional assessment (e.g. dynamic contrast enhanced MRI or CT or (18)F fluorodeoxyglucose positron emission tomographic (FDG PET) techniques assessing tumour metabolism). As can be seen, the Working Group and particularly those involved in imaging research, did not believe that there is at present sufficient standardisation and widespread availability to recommend adoption of these alternative assessment methods. The only exception to this is in the use of FDG PET imaging as an adjunct to determination of progression, as described later in this guideline. As detailed in paper in this special issue[12], we believe that the use of these promising newer approaches (which could either *add to* or *substitute for* anatomical assessment as described in RECIST) requires appropriate and rigorous clinical validation studies. This paper by Sargent et al. illustrates the type of data that will be needed to be able to define 'endpoints' for these modalities and how to determine where and when such criteria/modalities can be used to improve the reliability with which truly active new agents are identified and truly inactive new agents are discarded in comparison to RECIST criteria in phase II screening trials. The RECIST Working Group looks forward to such data emerging in the next few years to allow the appropriate changes to the next iteration of the RECIST criteria.

2. Purpose of this guideline

This guideline describes a standard approach to solid tumour measurement and definitions for objective assessment of change in tumour size for use in adult and paediatric cancer clinical trials. It is expected these criteria will be useful in all trials where objective response is the primary study endpoint, as well as in trials where assessment of stable disease, tumour progression or time to progression analyses are undertaken, since all of these outcome measures are based on an assessment of anatomical tumour burden and its change on study. There are no assumptions in this paper about the proportion of patients meeting the criteria for any of these endpoints which will signal that an agent or treatment regimen is active: those definitions are dependent on type of cancer in which a trial is being undertaken and the specific agent(s) under study. Protocols must include appropriate statistical sections which define the efficacy parameters upon which the trial sample size and decision criteria are based. In addition to providing definitions and criteria for assessment of tumour response, this guideline also makes recommendations regarding standard reporting of the results of trials that utilise tumour response as an endpoint.

While these guidelines may be applied in malignant brain tumour studies, there are also separate criteria published for response assessment in that setting.[13] This guideline is *not* intended for use for studies of malignant lymphoma since international guidelines for response assessment in lymphoma are published separately.[14]

Finally, many oncologists in their daily clinical practice follow their patients' malignant disease by means of repeated imaging studies and make decisions about continued therapy on the basis of both objective and symptomatic criteria. It is not intended that these RECIST guidelines play a role in that decision making, except if determined appropriate by the treating oncologist.

3. Measurability of tumour at baseline

3.1. *Definitions*

At baseline, tumour lesions/lymph nodes will be categorised measurable or non measurable as follows:

3.1.1. *Measurable*

Tumour lesions: Must be accurately measured in at least one dimension (*longest* diameter in the plane of measurement is to be recorded) with a *minimum* size of:

- 10 mm by CT scan (CT scan slice thickness no greater than 5 mm; see Appendix II on imaging guidance).
- 10 mm caliper measurement by clinical exam (lesions which cannot be accurately measured with calipers should be recorded as non measurable).
- 20 mm by chest X ray.

Malignant lymph nodes: To be considered pathologically en larged *and* measurable, a lymph node must be ⩾15 mm in *short* axis when assessed by CT scan (CT scan slice thickness recommended to be no greater than 5 mm). At baseline and in follow up, only the *short* axis will be measured and followed (see Schwartz et al. in this Special Issue[15]). See also notes be low on 'Baseline documentation of target and non target le sions' for information on lymph node measurement.

3.1.2. *Non-measurable*

All other lesions, including small lesions (longest diameter <10 mm or pathological lymph nodes with ⩾10 to <15 mm short axis) as well as truly non measurable lesions. Lesions considered truly non measurable include: leptomeningeal dis ease, ascites, pleural or pericardial effusion, inflammatory breast disease, lymphangitic involvement of skin or lung, abdominal masses/abdominal organomegaly identified by physical exam that is not measurable by reproducible imaging techniques.

3.1.3. *Special considerations regarding lesion measurability*

Bone lesions, cystic lesions, and lesions previously treated with local therapy require particular comment:

Bone lesions:.

- Bone scan, PET scan or plain films are not considered ade quate imaging techniques to measure bone lesions. How ever, these techniques can be used to confirm the presence or disappearance of bone lesions.
- Lytic bone lesions or mixed lytic blastic lesions, with *identi fiable soft tissue components*, that can be evaluated by cross sectional imaging techniques such as CT or MRI can be con sidered as measurable lesions if the *soft tissue component* meets the definition of measurability described above.
- Blastic bone lesions are non measurable.

Cystic lesions:.

- Lesions that meet the criteria for radiographically defined simple cysts should not be considered as malignant lesions (neither measurable nor non measurable) since they are, by definition, simple cysts.
- 'Cystic lesions' thought to represent cystic metastases can be considered as measurable lesions, if they meet the defi nition of measurability described above. However, if non cystic lesions are present in the same patient, these are pre ferred for selection as target lesions.

Lesions with prior local treatment:

- Tumour lesions situated in a previously irradiated area, or in an area subjected to other loco regional therapy, are usu ally not considered measurable unless there has been dem onstrated progression in the lesion. Study protocols should detail the conditions under which such lesions would be considered measurable.

3.2. *Specifications by methods of measurements*

3.2.1. *Measurement of lesions*

All measurements should be recorded in metric notation, using calipers if clinically assessed. All baseline evaluations should be performed as close as possible to the treatment start and never more than 4 weeks before the beginning of the treatment.

3.2.2. *Method of assessment*

The same method of assessment and the same technique should be used to characterise each identified and reported lesion at baseline and during follow up. Imaging based evalu ation should always be done rather than clinical examination unless the lesion(s) being followed cannot be imaged but are assessable by clinical exam.

Clinical lesions: Clinical lesions will only be considered mea surable when they are superficial and ⩾10 mm diameter as assessed using calipers (e.g. skin nodules). For the case of skin lesions, documentation by colour photography including a ru ler to estimate the size of the lesion is suggested. As noted above, when lesions can be evaluated by both clinical exam and imaging, imaging evaluation should be undertaken since it is more objective and may also be reviewed at the end of the study.

Chest X ray: Chest CT is preferred over chest X ray, particu larly when progression is an important endpoint, since CT is more sensitive than X ray, particularly in identifying new le sions. However, lesions on chest X ray may be considered measurable if they are clearly defined and surrounded by aer ated lung. See Appendix II for more details.

CT, MRI: CT is the best currently available and reproducible method to measure lesions selected for response assessment. This guideline has defined measurability of lesions on CT scan based on the assumption that CT slice thickness is 5 mm or less. As is described in Appendix II, when CT scans have slice thickness greater than 5 mm, the minimum size for a measurable lesion should be twice the slice thickness. MRI is also acceptable in certain situations (e.g. for body scans). More details concerning the use of both CT and MRI for assessment of objective tumour response evaluation are provided in Appendix II.

Ultrasound: Ultrasound is not useful in assessment of lesion size and should not be used as a method of measurement. Ultrasound examinations cannot be reproduced in their en tirety for independent review at a later date and, because they are operator dependent, it cannot be guaranteed that the same technique and measurements will be taken from one assessment to the next (described in greater detail in Appendix II). If new lesions are identified by ultrasound in the course of the study, confirmation by CT or MRI is ad vised. If there is concern about radiation exposure at CT, MRI may be used instead of CT in selected instances.

Endoscopy, laparoscopy: The utilisation of these techniques for objective tumour evaluation is not advised. However, they can be useful to confirm complete pathological response when biopsies are obtained or to determine relapse in trials where recurrence following complete response or surgical resection is an endpoint.

Tumour markers: Tumour markers *alone* cannot be used to as sess *objective* tumour response. If markers are initially above

the upper normal limit, however, they must normalise for a patient to be considered in complete response. Because tumour markers are disease specific, instructions for their measurement should be incorporated into protocols on a disease specific basis. Specific guidelines for both CA 125 response (in recurrent ovarian cancer) and PSA response (in recurrent prostate cancer), have been published.[16–18] In addition, the Gynecologic Cancer Intergroup has developed CA125 progression criteria which are to be integrated with objective tumour assessment for use in first line trials in ovarian cancer.[19]

Cytology, histology: These techniques can be used to differentiate between PR and CR in rare cases if required by protocol (for example, residual lesions in tumour types such as germ cell tumours, where known residual benign tumours can remain). When effusions are known to be a potential adverse effect of treatment (e.g. with certain taxane compounds or angiogenesis inhibitors), the cytological confirmation of the neoplastic origin of any effusion that appears or worsens during treatment can be considered if the measurable tumour has met criteria for response or stable disease in order to differentiate between response (or stable disease) and progressive disease.

4. Tumour response evaluation

4.1. *Assessment of overall tumour burden and measurable disease*

To assess objective response or future progression, it is necessary to estimate the *overall tumour burden at baseline* and use this as a comparator for subsequent measurements. Only patients with measurable disease at baseline should be included in protocols where objective tumour response is the primary endpoint. Measurable disease is defined by the presence of at least one measurable lesion (as detailed above in Section 3). In studies where the primary endpoint is tumour progression (either time to progression or proportion with progression at a fixed date), the protocol must specify if entry is restricted to those with measurable disease or whether patients having non measurable disease only are also eligible.

4.2. *Baseline documentation of 'target' and 'non-target' lesions*

When more than one measurable lesion is present at baseline all lesions up to a maximum of five lesions total (and a maximum of two lesions per organ) representative of all involved organs should be identified as *target lesions* and will be recorded and measured at baseline (this means in instances where patients have only one or two organ sites involved a *maximum* of two and four lesions respectively will be recorded). For evidence to support the selection of only five target lesions, see analyses on a large prospective database in the article by Bogaerts et al.[10].

Target lesions should be selected on the basis of their size (lesions with the longest diameter), be representative of all involved organs, but in addition should be those that lend themselves to *reproducible repeated measurements*. It may be the case that, on occasion, the largest lesion does not lend itself to reproducible measurement in which circumstance the next largest lesion which can be measured reproducibly should be selected. To illustrate this point see the example in Fig. 3 of Appendix II.

Lymph nodes merit special mention since they are normal anatomical structures which may be visible by imaging even if not involved by tumour. As noted in Section 3, pathological nodes which are defined as measurable and may be identified as target lesions must meet the criterion of a short axis of $\geqslant$15 mm by CT scan. Only the *short* axis of these nodes will contribute to the baseline sum. The short axis of the node is the diameter normally used by radiologists to judge if a node is involved by solid tumour. Nodal size is normally reported as two dimensions in the plane in which the image is obtained (for CT scan this is almost always the axial plane; for MRI the plane of acquisition may be axial, saggital or coronal). The smaller of these measures is the short axis. For example, an abdominal node which is reported as being 20 mm $\times$ 30 mm has a short axis of 20 mm and qualifies as a malignant, measurable node. In this example, 20 mm should be recorded as the node measurement (See also the example in Fig. 4 in Appendix II). All other pathological nodes (those with short axis $\geqslant$10 mm but <15 mm) should be considered non target lesions. Nodes that have a short axis <10 mm are considered non pathological and should not be recorded or followed.

A *sum of the diameters* (longest for non nodal lesions, short axis for nodal lesions) for all target lesions will be calculated and reported as the *baseline sum diameters*. If lymph nodes are to be included in the sum, then as noted above, only the *short* axis is added into the sum. The baseline sum diameters will be used as reference to further characterise any objective tumour regression in the measurable dimension of the disease.

All other lesions (or sites of disease) including pathological lymph nodes should be identified as *non target lesions* and should also be recorded at baseline. Measurements are not required and these lesions should be followed as 'present', 'absent', or in rare cases 'unequivocal progression' (more details to follow). In addition, it is possible to record multiple non target lesions involving the same organ as a single item on the case record form (e.g. 'multiple enlarged pelvic lymph nodes' or 'multiple liver metastases').

4.3. *Response criteria*

This section provides the definitions of the criteria used to determine objective tumour response for target lesions.

4.3.1. *Evaluation of target lesions*

Complete Response (CR): Disappearance of all target lesions. Any pathological lymph nodes (whether target or non target) must have reduction in short axis to <10 mm.

Partial Response (PR): At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameters.

Progressive Disease (PD): At least a 20% increase in the sum of diameters of target lesions, taking as reference the *smallest sum on study* (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase of at least 5 mm. (*Note:* the appearance of one or more new lesions is also considered progression).

Stable Disease (SD): Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameters while on study.

4.3.2. Special notes on the assessment of target lesions

Lymph nodes. Lymph nodes identified as target lesions should always have the actual short axis measurement recorded (measured in the same anatomical plane as the baseline examination), even if the nodes regress to below 10 mm on study. This means that when lymph nodes are included as target lesions, the 'sum' of lesions may not be zero even if complete response criteria are met, since a normal lymph node is defined as having a short axis of <10 mm. Case report forms or other data collection methods may therefore be designed to have target nodal lesions recorded in a separate section where, in order to qualify for CR, each node must achieve a short axis <10 mm. For PR, SD and PD, the actual short axis measurement of the nodes is to be included in the sum of target lesions.

Target lesions that become 'too small to measure'. While on study, all lesions (nodal and non nodal) recorded at baseline should have their actual measurements recorded at each subsequent evaluation, even when very small (e.g. 2 mm). However, sometimes lesions or lymph nodes which are recorded as target lesions at baseline become so faint on CT scan that the radiologist may not feel comfortable assigning an exact measure and may report them as being 'too small to measure'. When this occurs it is important that a value be recorded on the case report form. If it is the opinion of the radiologist that the lesion has likely disappeared, the measurement should be recorded as 0 mm. If the lesion is believed to be present and is faintly seen but too small to measure, a default value of 5 mm should be assigned (*Note:* It is less likely that this rule will be used for lymph nodes since they usually have a definable size when normal and are frequently surrounded by fat such as in the retroperitoneum; however, if a lymph node is believed to be present and is faintly seen but too small to measure, a default value of 5 mm should be assigned in this circumstance as well). This default value is derived from the 5 mm CT slice thickness (but should not be changed with varying CT slice thickness). The measurement of these lesions is potentially non reproducible, therefore providing this default value will prevent false responses or progressions based upon measurement error. To reiterate, however, if the radiologist is able to provide an actual measure, that should be recorded, even if it is below 5 mm.

Lesions that split or coalesce on treatment. As noted in Appendix II, when non nodal lesions 'fragment', the longest diameters of the fragmented portions should be added together to calculate the target lesion sum. Similarly, as lesions coalesce, a plane between them may be maintained that would aid in obtaining maximal diameter measurements of each individual lesion. If the lesions have truly coalesced such that they are no longer separable, the vector of the longest diameter in this instance should be the maximal longest diameter for the 'coalesced lesion'.

4.3.3. Evaluation of non-target lesions

This section provides the definitions of the criteria used to determine the tumour response for the group of non target lesions. While some non target lesions may actually be measurable, they need not be measured and instead should be assessed only *qualitatively* at the time points specified in the protocol.

Complete Response (CR): Disappearance of all non target lesions and normalisation of tumour marker level. All lymph nodes must be non pathological in size (<10 mm short axis).

Non CR/Non PD: Persistence of one or more non target lesion(s) and/or maintenance of tumour marker level above the normal limits.

Progressive Disease (PD): *Unequivocal progression* (see comments below) of existing non target lesions. (*Note:* the appearance of one or more new lesions is also considered progression).

4.3.4. Special notes on assessment of progression of non-target disease

The concept of progression of non target disease requires additional explanation as follows:

When the patient also has measurable disease. In this setting, to achieve 'unequivocal progression' on the basis of the non target disease, there must be an overall level of substantial worsening in non target disease such that, even in presence of SD or PR in target disease, the overall tumour burden has increased sufficiently to merit discontinuation of therapy (see examples in Appendix II and further details below). A modest 'increase' in the size of one or more non target lesions is usually not sufficient to quality for unequivocal progression status. The designation of overall progression *solely* on the basis of change in non target disease in the face of SD or PR of target disease will therefore be extremely rare.

When the patient has only non-measurable disease. This circumstance arises in some phase III trials when it is not a criterion of study entry to have measurable disease. The same general concepts apply here as noted above, however, in this instance there is no measurable disease assessment to factor into the interpretation of an increase in non measurable disease burden. Because worsening in non target disease cannot be easily quantified (by definition: if all lesions are truly non measurable) a useful test that can be applied when assessing patients for unequivocal progression is to consider if the increase in overall disease burden based on the change in non measurable disease is comparable in magnitude to the increase that would be required to declare PD for measurable disease: i.e. an increase in tumour burden representing an additional 73% increase in 'volume' (which is equivalent to a 20% increase diameter in a measurable lesion). Examples include an increase in a pleural effusion from 'trace' to 'large', an increase in lymphangitic

disease from localised to widespread, or may be described in protocols as 'sufficient to require a change in therapy'. Some illustrative examples are shown in Figs. 5 and 6 in Appendix II. If 'unequivocal progression' is seen, the patient should be considered to have had overall PD at that point. While it would be ideal to have objective criteria to apply to non measurable disease, the very nature of that disease makes it impossible to do so, therefore the increase must be substantial.

4.3.5. *New lesions*

The appearance of new malignant lesions denotes disease progression; therefore, some comments on detection of new lesions are important. There are no specific criteria for the identification of new radiographic lesions; however, the finding of a new lesion should be unequivocal: i.e. not attributable to differences in scanning technique, change in imaging modality or findings thought to represent something other than tumour (for example, some 'new' bone lesions may be simply healing or flare of pre existing lesions). This is particularly important when the patient's baseline lesions show partial or complete response. For example, necrosis of a liver lesion may be reported on a CT scan report as a 'new' cystic lesion, which it is not.

A lesion identified on a follow up study in an anatomical location that was *not* scanned at baseline is considered a new lesion and will indicate disease progression. An example of this is the patient who has visceral disease at baseline and while on study has a CT or MRI brain ordered which reveals metastases. The patient's brain metastases are considered to be evidence of PD even if he/she did not have brain imaging at baseline.

If a new lesion is equivocal, for example because of its small size, continued therapy and follow up evaluation will clarify if it represents truly new disease. If repeat scans confirm there is definitely a new lesion, then progression should be declared using the date of the initial scan.

While FDG PET response assessments need additional study, it is sometimes reasonable to incorporate the use of FDG PET scanning to complement CT scanning in assessment of progression (particularly possible 'new' disease). New lesions on the basis of FDG PET imaging can be identified according to the following algorithm:

a. Negative FDG PET at baseline, with a positive[1] FDG PET at follow up is a sign of PD based on a new lesion.
b. No FDG PET at baseline and a positive FDG PET at follow up:
   If the positive FDG PET at follow up corresponds to a new site of disease confirmed by CT, this is PD.
   If the positive FDG PET at follow up is not confirmed as a new site of disease on CT, additional follow up CT scans are needed to determine if there is truly progression occurring at that site (if so, the date of PD will be the date of the initial abnormal FDG PET scan).
   If the positive FDG PET at follow up corresponds to a pre existing site of disease on CT that is not progressing on the basis of the anatomic images, this is not PD.

[1] A 'positive' FDG PET scan lesion means one which is FDG avid with an uptake greater than twice that of the surrounding tissue on the attenuation corrected image.

4.4. *Evaluation of best overall response*

The best overall response is the best response recorded from the start of the study treatment until the end of treatment taking into account any requirement for confirmation. On occasion a response may not be documented until after the end of therapy so protocols should be clear if post treatment assessments are to be considered in determination of best overall response. Protocols must specify how any new therapy introduced before progression will affect best response designation. The patient's best overall response assignment will depend on the findings of both target and non target disease and will also take into consideration the appearance of new lesions. Furthermore, depending on the nature of the study and the protocol requirements, it may also require confirmatory measurement (see Section 4.6). Specifically, in non randomised trials where response is the primary endpoint, confirmation of PR or CR is needed to deem either one the 'best overall response'. This is described further below.

4.4.1. *Time point response*

It is assumed that at each protocol specified time point, a response assessment occurs. Table 1 on the next page provides a summary of the overall response status calculation at each time point for patients who have measurable disease at baseline.

When patients have non measurable (therefore non target) disease only, Table 2 is to be used.

4.4.2. *Missing assessments and inevaluable designation*

When no imaging/measurement is done at all at a particular time point, the patient is not evaluable (NE) at that time point. If only a subset of lesion measurements are made at an assessment, usually the case is also considered NE at that time point, unless a convincing argument can be made that the contribution of the individual missing lesion(s) would not change the assigned time point response. This would be most likely to happen in the case of PD. For example, if a patient had a baseline sum of 50 mm with three measured lesions and at follow up only two lesions were assessed, but those gave a sum of 80 mm, the patient will have achieved PD status, regardless of the contribution of the missing lesion.

4.4.3. *Best overall response: all time points*

The *best overall response* is determined once all the data for the patient is known.

Best response determination in trials where confirmation of complete or partial response IS NOT required: Best response in these trials is defined as the best response across all time points (for example, a patient who has SD at first assessment, PR at second assessment, and PD on last assessment has a best overall response of PR). When SD is believed to be best response, it must also meet the protocol specified minimum time from baseline. If the minimum time is not met when SD is otherwise the best time point response, the patient's best response depends on the subsequent assessments. For example, a patient who has SD at first assessment, PD at second and does not meet minimum duration for SD, will have a best response of PD. The same patient lost to follow up after the first SD assessment would be considered inevaluable.

Table 1 – Time point response: patients with target (+/– non-target) disease.

Target lesions	Non target lesions	New lesions	Overall response
CR	CR	No	CR
CR	Non CR/non PD	No	PR
CR	Not evaluated	No	PR
PR	Non PD or not all evaluated	No	PR
SD	Non PD or not all evaluated	No	SD
Not all evaluated	Non PD	No	NE
PD	Any	Yes or No	PD
Any	PD	Yes or No	PD
Any	Any	Yes	PD

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and NE = inevaluable.

Table 2 – Time point response: patients with non-target disease only.

Non target lesions	New lesions	Overall response
CR	No	CR
Non CR/non PD	No	Non CR/non PD[a]
Not all evaluated	No	NE
Unequivocal PD	Yes or No	PD
Any	Yes	PD

CR = complete response, PD = progressive disease, and NE = inevaluable.

a 'Non CR/non PD' is preferred over 'stable disease' for non target disease since SD is increasingly used as endpoint for assessment of efficacy in some trials so to assign this category when no lesions can be measured is not advised.

Best response determination in trials where confirmation of complete or partial response IS required: Complete or partial responses may be claimed only if the criteria for each are met at a subsequent time point as specified in the protocol (generally 4 weeks later). In this circumstance, the best overall response can be interpreted as in Table 3.

4.4.4. *Special notes on response assessment*

When nodal disease is included in the sum of target lesions and the nodes decrease to 'normal' size (<10 mm), they may still have a measurement reported on scans. This measurement should be recorded even though the nodes are normal in order not to overstate progression should it be based on increase in size of the nodes. As noted earlier, this means that patients with CR may not have a total sum of 'zero' on the case report form (CRF).

In trials where confirmation of response is required, repeated 'NE' time point assessments may complicate best response determination. The analysis plan for the trial must address how missing data/assessments will be addressed in determination of response and progression. For example, in most trials it is reasonable to consider a patient with time point responses of PR NE PR as a confirmed response.

Patients with a global deterioration of health status requiring discontinuation of treatment without objective evidence of disease progression at that time should be reported as 'symptomatic deterioration'. Every effort should be made to document objective progression even after discontinuation of treatment. Symptomatic deterioration is *not* a descriptor of an objective response: it is a reason for stopping study therapy. The objective response status of such patients is to be determined by evaluation of target and non target disease as shown in Tables 1 3.

Conditions that define 'early progression, early death and inevaluability' are study specific and should be clearly described in each protocol (depending on treatment duration, treatment periodicity).

In some circumstances it may be difficult to distinguish residual disease from normal tissue. When the evaluation of complete response depends upon this determination, it is recommended that the residual lesion be investigated (fine

Table 3 – Best overall response when confirmation of CR and PR required.

Overall response First time point	Overall response Subsequent time point	BEST overall response
CR	CR	CR
CR	PR	SD, PD or PR[a]
CR	SD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
CR	NE	SD provided minimum criteria for SD duration met, otherwise NE
PR	CR	PR
PR	PR	PR
PR	SD	SD
PR	PD	SD provided minimum criteria for SD duration met, otherwise, PD
PR	NE	SD provided minimum criteria for SD duration met, otherwise NE
NE	NE	NE

CR = complete response, PR = partial response, SD = stable disease, PD = progressive disease, and NE = inevaluable.

a If a CR is truly met at first time point, then any disease seen at a subsequent time point, even disease meeting PR criteria relative to baseline, makes the disease PD at that point (since disease must have reappeared after CR). Best response would depend on whether minimum duration for SD was met. However, sometimes 'CR' may be claimed when subsequent scans suggest small lesions were likely still present and in fact the patient had PR, not CR at the first time point. Under these circumstances, the original CR should be changed to PR and the best response is PR.

needle aspirate/biopsy) before assigning a status of complete response. FDG PET may be used to upgrade a response to a CR in a manner similar to a biopsy in cases where a residual radiographic abnormality is thought to represent fibrosis or scarring. The use of FDG PET in this circumstance should be prospectively described in the protocol and supported by dis ease specific medical literature for the indication. However, it must be acknowledged that both approaches may lead to false positive CR due to limitations of FDG PET and biopsy res olution/sensitivity.

For equivocal findings of progression (e.g. very small and uncertain new lesions; cystic changes or necrosis in existing lesions), treatment may continue until the next scheduled assessment. If at the next scheduled assessment, progression is confirmed, the date of progression should be the earlier date when progression was suspected.

4.5. *Frequency of tumour re-evaluation*

Frequency of tumour re evaluation while on treatment should be protocol specific and adapted to the type and sche dule of treatment. However, in the context of phase II studies where the beneficial effect of therapy is not known, follow up every 6 8 weeks (timed to coincide with the end of a cycle) is reasonable. Smaller or greater time intervals than these could be justified in specific regimens or circumstances. The proto col should specify which organ sites are to be evaluated at baseline (usually those most likely to be involved with meta static disease for the tumour type under study) and how often evaluations are repeated. Normally, all target and non target sites are evaluated at each assessment. In selected circum stances certain non target organs may be evaluated less fre quently. For example, bone scans may need to be repeated only when complete response is identified in target disease or when progression in bone is suspected.

After the end of the treatment, the need for repetitive tu mour evaluations depends on whether the trial has as a goal the response rate or the time to an event (progression/death). If 'time to an event' (e.g. time to progression, disease free survival, progression free survival) is the main endpoint of the study, then routine scheduled re evaluation of protocol specified sites of disease is warranted. In randomised com parative trials in particular, the scheduled assessments should be performed as identified on a calendar schedule (for example: every 6 8 weeks on treatment or every 3 4 months after treatment) and should not be affected by delays in therapy, drug holidays or any other events that might lead to imbalance in a treatment arm in the timing of disease assessment.

4.6. *Confirmatory measurement/duration of response*

4.6.1. *Confirmation*

In non randomised trials where response is the primary end point, confirmation of PR and CR is required to ensure re sponses identified are not the result of measurement error. This will also permit appropriate interpretation of results in the context of historical data where response has traditionally required confirmation in such trials (see the paper by Bogaerts et al. in this Special Issue[10]). However, in all other circum stances, i.e. in randomised trials (phase II or III) or studies where stable disease or progression are the primary endpoints, confirmation of response is not required since it will not add va lue to the interpretation of trial results. However, elimination of the requirement for response confirmation may increase the importance of central review to protect against bias, in partic ular in studies which are not blinded.

In the case of SD, measurements must have met the SD criteria at least once after study entry at a minimum interval (in general not less than 6 8 weeks) that is defined in the study protocol.

4.6.2. *Duration of overall response*

The duration of overall response is measured from the time measurement criteria are first met for CR/PR (whichever is first recorded) until the first date that recurrent or progressive dis ease is objectively documented (taking as reference for progres sive disease the smallest measurements recorded on study).

The duration of overall complete response is measured from the time measurement criteria are first met for CR until the first date that recurrent disease is objectively documented.

4.6.3. *Duration of stable disease*

Stable disease is measured from the start of the treatment (in randomised trials, from date of randomisation) until the crite ria for progression are met, taking as reference the *smallest sum on study* (if the baseline sum is the smallest, this is the reference for calculation of PD).

The clinical relevance of the duration of stable disease var ies in different studies and diseases. If the proportion of pa tients achieving stable disease for a minimum period of time is an endpoint of importance in a particular trial, the protocol should specify the minimal time interval required between two measurements for determination of stable disease.

Note: The duration of response and stable disease as well as the progression free survival are influenced by the frequency of follow up after baseline evaluation. It is not in the scope of this guideline to define a standard follow up frequency. The fre quency should take into account many parameters including disease types and stages, treatment periodicity and standard practice. However, these limitations of the precision of the measured endpoint should be taken into account if compari sons between trials are to be made.

4.7. *Progression-free survival/proportion progression-free*

4.7.1. *Phase II trials*

This guideline is focused primarily on the use of objective re sponse endpoints for phase II trials. In some circumstances, 're sponse rate' may not be the optimal method to assess the potential anticancer activity of new agents/regimens. In such cases 'progression free survival' (PFS) or the 'proportion pro gression free' at landmark time points, might be considered appropriate alternatives to provide an initial signal of biologic effect of new agents. It is clear, however, that in an uncontrolled trial, these measures are subject to criticism since an appar ently promising observation may be related to biological factors such as patient selection and not the impact of the intervention. Thus, phase II screening trials utilising these endpoints are best designed with a randomised control. Exceptions may exist

where the behaviour patterns of certain cancers are so consis tent (and usually consistently poor), that a non randomised trial is justifiable (see for example van Glabbeke et al.[20]). How ever, in these cases it will be essential to document with care the basis for estimating the expected PFS or proportion progres sion free in the absence of a treatment effect.

4.7.2. *Phase III trials*

Phase III trials in advanced cancers are increasingly designed to evaluate progression free survival or time to progression as the primary outcome of interest. Assessment of progression is relatively straightforward if the protocol requires all pa tients to have measurable disease. However, restricting entry to this subset of patients is subject to criticism: it may result in a trial where the results are less likely to be generalisable if, in the disease under study, a substantial proportion of pa tients would be excluded. Moreover, the restriction to entry will slow recruitment to the study. Increasingly, therefore, tri als allow entry of both patients with measurable disease as well as those with non measurable disease only. In this cir cumstance, care must be taken to explicitly describe the find ings which would qualify for progressive disease for those patients *without* measurable lesions. Furthermore, in this set ting, protocols must indicate if the maximum number of re corded target lesions for those patients with measurable disease may be relaxed from five to three (based on the data found in Bogaerts et al.[10] and Moskowitz et al.[11]). As found in the 'special notes on assessment of progression', these guide lines offer recommendations for assessment of progression in this setting. Furthermore, if available, validated tumour mar ker measures of progression (as has been proposed for ovarian cancer) may be useful to integrate into the definition of pro gression. Centralised blinded review of imaging studies or of source imaging reports to verify 'unequivocal progression' may be needed if important drug development or drug ap proval decisions are to be based on the study outcome. Finally, as noted earlier, because the date of progression is subject to ascertainment bias, timing of investigations in study arms should be the same. The article by Dancey et al. in this special issue[21] provides a more detailed discussion of the assessment of progression in randomised trials.

4.8. *Independent review of response and progression*

For trials where *objective response* (CR + PR) is the primary end point, and in particular where key drug development deci sions are based on the observation of a minimum number of responders, it is recommended that all claimed responses be reviewed by an expert(s) independent of the study. If the study is a randomised trial, ideally reviewers should be blinded to treatment assignment. Simultaneous review of the patients' files and radiological images is the best approach.

Independent review of progression presents some more complex issues: for example, there are statistical problems with the use of central review based progression time in place of investigator based progression time due to the poten tial introduction of informative censoring when the former precedes the latter. An overview of these factors and other lessons learned from independent review is provided in an article by Ford et al. in this special issue.[22]

4.9. *Reporting best response results*

4.9.1. *Phase II trials*

When response is the primary endpoint, and thus all patients must have measurable disease to enter the trial, all patients included in the study must be accounted for in the report of the results, even if there are major protocol treatment devia tions or if they are not evaluable. Each patient will be assigned one of the following categories:

1. Complete response
2. Partial response
3. Stable disease
4. Progression
5. Inevaluable for response: specify reasons (for example: early death, malignant disease; early death, toxicity; tumour assessments not repeated/incomplete; other (specify)).

Normally, all *eligible* patients should be included in the denominator for the calculation of the response rate for phase II trials (in some protocols it will be appropriate to include all treated patients). It is generally preferred that 95% two sided confidence limits are given for the calculated response rate. Trial conclusions should be based on the response rate for all eligible (or all treated) patients and should *not* be based on a selected 'evaluable' subset.

4.9.2. *Phase III trials*

Response evaluation in phase III trials may be an indicator of the relative anti tumour activity of the treatments eval uated and is almost always a secondary endpoint. Ob served differences in response rate may not predict the clinically relevant therapeutic benefit for the population studied. If objective response is selected as a primary end point for a phase III study (only in circumstances where a direct relationship between objective tumour response and a clinically relevant therapeutic benefit can be unambigu ously demonstrated for the population studied), the same criteria as those applying to phase II trials should be used and all patients entered should have at least one measur able lesion.

In those many cases where response is a secondary end point and not all trial patients have measurable disease, the method for reporting overall best response rates must be pre specified in the protocol. In practice, response rate may be reported using either an 'intent to treat' analysis (all ran domised patients in the denominator) or an analysis where only the subset of patients with measurable disease at baseline are included. The protocol should clearly specify how response results will be reported, including any subset analyses that are planned.

The original version of RECIST suggested that in phase III trials one could write protocols using a 'relaxed' interpreta tion of the RECIST guidelines (for example, reducing the num ber of lesions measured) but this should no longer be done since these revised guidelines have been amended in such a way that it is clear how these criteria should be applied for all trials in which anatomical assessment of tumour response or progression are endpoints.

Appendix I. Summary of major changes RECIST 1.0 to RECIST 1.1

	RECIST 1.0	RECIST 1.1	Rationale	Reference in special issue (if applicable)
Minimum size measurable lesions	CT: 10 mm spiral 20 mm non spiral	CT 10 mm; delete reference to spiral scan	Most scans used have 5 mm or less slice thickness Clearer to give instruction based on slice interval if it is greater than 5 mm	
	Clinical: 20 mm	Clinical: 10 mm (must be measurable with calipers)	Caliper measurement will make this reliable	
	Lymph node: not mentioned	CT: ⩾15 mm short axis for target ⩾10 <15 mm for non target <10 mm is non pathological	Since nodes are normal structure need to define pathological enlargement. Short axis is most sensitive	Schwartz et al.[15]
Special considerations on lesion measurability		Notes included on bone lesions, cystic lesions	Clarify frequently asked questions	
Overall tumour burden	10 lesions (5 per organ)	5 lesions (2 per organ)	Data warehouse analysis shows no loss of information if lesion number reduced from 10 to 5. A maximum of 2 lesions per organ yields sufficient representation per disease site	Bogaerts et al.[10]
Response criteria target disease	CR lymph node not mentioned	CR lymph nodes must be <10 mm short axis	In keeping with normal size of nodes	Schwartz et al.[15]
	PD 20% increase over smallest sum on study or new lesions	PD 20% increase over smallest sum on study (including baseline if that is smallest) and at least 5 mm increase or new lesions	Clarification that if baseline measurement is smaller than any on study measurement, it is reference against which PD is assessed 5 mm absolute increase to guard against over calling PD when total sum is very small and 20% increase is within measurement error	
Response criteria non target disease	'unequivocal progression' considered as PD	More detailed description of 'unequivocal progression' to indicate that it should not normally trump target disease status. It must be representative of overall disease status change, not a single lesion increase	Confusion with RECIST 1.0 where some were considering PD if 'increase' in any non target lesion, even when target disease is stable or responding	
New lesions		New section on New lesions	To provide guidance on when a lesion is considered new (and thus PD)	
Overall response	Table integrated target and non target lesions	Two tables: one integrating target and non target and the other of non target only	To account for the fact that RECIST criteria are now being used in trials where PFS is the endpoint and not all patients have measurable (target) disease at baseline	Dancey et al.[21]

		Special notes: How to assess and measure lymph nodes CR in face of residual tissue Discussion of 'equivocal' progression	Frequently asked questions on these topics	
Confirmatory measure	For CR and PR: criteria must be met again 4 weeks after initial documentation	Retain this requirement ONLY for non randomised trials with primary endpoint of response	Data warehouse shows that response rates rise when confirmation is eliminated, but the only circumstance where this is important is in trials where there is no concurrent comparative control and where this measure is the primary endpoint	Bogaerts et al.[10]
Progression free survival	General comments only	More specific comments on use of PFS (or proportion progression free) as phase II endpoint Greater detail on PFS assessment in phase III trials	Increasing use of PFS in phase III trials requires guidance on assessment of PD in patients with non measurable disease	Dancey et al.[21]
Reporting of response results	9 categories suggested for reporting phase II results	Divided into phase II and phase III 9 categories collapsed into 5 In phase III, guidance given about reporting response	Simplifies reporting and clarifies how to report phase II and III data consistently	
Response in phase III trials	More relaxed guidelines possible if protocol specified	This section removed and referenced in section above: no need to have different criteria for phase II and III	Simplification of response assessment by reducing number of lesions and eliminating need for confirmation in randomised studies where response is not the primary endpoint makes separate 'rules' unnecessary	
Imaging appendix	Appendix I	Appendix II: updated with detailed guidance on use of MRI, PET/CT Other practical guidance included	Evolving use of newer modalities addressed. Enhanced guidance in response to frequent questions and from radiology review experience	
New appendices		Appendix I: comparison of RECIST 1.0 and 1.1 Appendix III: frequently asked questions		

Conflict of interest statement

None declared.

Acknowledgements

The RECIST Working Group would like to thank the following organisations which made data bases available to us in order to perform the analyses which informed decisions about changes to this version of the criteria: Amgen; AstraZeneca; Breast Cancer International Research Group (BCIRG); Bristol Myers Squibb; European Organisation for Research and Treatment of Cancer (EORTC) Breast Cancer Group and Gastrointestinal Group; Erasmus University Medical Center, Rotterdam, The Netherlands; Genentech; Pfizer; RadPharm; Roche; Sanofi Aventis.

We would also like to thank the following individuals from academic, government, and pharmaceutical organisations for providing helpful comments on an earlier draft of these revised guidelines: Ohad Amit, Phil Murphy, Teri Crofts and Janet Begun, GlaxoSmithKline, USA; Laurence H. Baker, Southwest Oncology Group, USA; Karla Ballman, Mayo Clinic, USA; Charles Baum, Darrel Cohen, and Mary Ashford Collier, Pfizer, USA; Gary J. Becker, American Board of Radiology, Tucson, USA; Jean Yves Blay, University Claude Pertrand, Lyon France; Renzo Canetta, Bristol Myers Squibb, USA; David Chang, Amgen Inc., USA; Sandra Chica, Perceptive Informations Inc. (PAREXEL), USA; Martin Edelman, University of Maryland Greenbaum Cancer Centre, USA; Gwendolyn Fyfe, Genentech, USA; Bruce Giantonio, Eastern Cooperative Oncology Group, USA; Gary Gordon, Abbott Pharmaceuticals, USA; Ronald Gottlieb, Roswell Park Cancer Institute, USA; Simon Kao, University of Iowa College of Medicine, USA; Wasaburo Koizumi, Kitasato University, Japan; Alessandro Riva, Novartis Pharmaceuticals, USA; Wayne Rackhoff, Ortho Biotech Oncology Research and Development, USA; Nagahiro Saijo, President Japanese Society of Medical Oncology, Japan; Mitchell Schnall American College of Radiology Imaging Network, USA; Yoshik Shimamura, PAREXEL International Inc., Japan; Rajeshwari Sridhara, Centre for Drug Evaluation and Research, Food and Drug Administration, USA; Andrew Stone, Alan Barge, AstraZeneca, United Kingdom; Orhan Suleiman, Centre for Drug Evaluation and Research, Food and Drug Administration, USA; Daniel C. Sullivan, Duke University Medical Centre, USA; Masakazu Toi, Kyoto University, Japan; Cindy Welsh, Centre for Drug Evaluation and Research, Food and Drug Administration, USA.

Finally, the RECIST Working Group would like to thank individuals who were not permanent members of the group (which are all acknowledged as co authors) but who attended working group meetings from time to time and made contributions to the total process over the past 7 years: Richard Pazdur, Food and Drug Administration, USA; Francesco Pignatti, European Medicines Agency, London, UK.

Appendix II. Specifications for standard anatomical radiological imaging

These protocols for image acquisition of computed tomography (CT) and magnetic resonance imaging (MRI) are recommendations intended for patients on clinical trials where RECIST assessment will be performed. Standardisation of imaging requirements and image acquisition parameters is ideal to allow for optimal comparability of subjects within a study and results between studies. These recommendations are designed to balance optimised image acquisition protocols with techniques that should be feasible to perform globally at imaging facilities in all types of radiology practices. These guidelines are not applicable to functional imaging techniques or volumetric assessment of tumour size.

Scanner quality control is highly recommended and should follow standard manufacturer and facility maintenance schedules using commercial phantoms. It is likely that for RECIST unidimensional measurements this will be adequate to produce reproducible measurements. Imaging quality control for CT includes an analysis of image noise and uniformity and CT number as well as spatial resolution. The frequency of quality control analysis is also variable and should focus on clinically relevant scanning parameters. Dose analysis is always important and the use of imaging should follow the ALARA principle, 'As Low As Reasonably Achievable', which refers to making every reasonable effort to maintain radiation exposures as far below the dose limits as possible.

Specific notes

Chest X ray measurement of lesions surrounded by pulmonary parenchyma is feasible, but not preferable as the measurement represents a summation of densities. Furthermore, there is poor identification of new lesions within the chest on X ray as compared with CT. Therefore, measurements of pulmonary parenchymal lesions as well as mediastinal disease are optimally performed with CT of the chest. MRI of the chest should only be performed in extenuating circumstances. Even if IV contrast cannot be administered (for example, in the situation of allergy to contrast), a non contrast CT of the chest is still preferred over MRI or chest X ray.

CT scans: CT scans of the chest, abdomen, and pelvis should be contiguous throughout all the anatomic region of interest. As a general rule, the minimum size of a measurable lesion at baseline should be no less than double the slice thickness and also have a minimum size of 10 mm (see below for minimum size when scanners have a slice thickness more than 5 mm). While the precise physics of lesion size and partial volume averaging is complex, lesions smaller than 10 mm may be difficult to accurately and reproducibly measure. While this rule is applicable to baseline scans, as lesions potentially decrease in size at follow up CT studies, they should still be measured. Lesions which are reported as 'too small to measure' should be assigned a default measurement of 5 mm if they are still visible.

The most critical CT image acquisition parameters for optimal tumour evaluation using RECIST are *anatomic coverage, contrast administration, slice thickness, and reconstruction interval*.

a. *Anatomic coverage*: Optimal anatomic coverage for most solid tumours is the chest, abdomen and pelvis. Coverage should encompass all areas of known predilection for metastases in the disease under evaluation and

should additionally investigate areas that may be involved based on signs and symptoms of individual patients. Because a lesion later identified in a body part not scanned at baseline would be considered as a new lesion representing disease progression, careful consid eration should be given to the extent of imaging coverage at baseline and at subsequent follow up time points. This will enable better consistency not only of tumour measurements but also identification of new disease.

b. *IV contrast administration*: Optimal visualisation and measurement of metastases in solid tumours requires consistent administration (dose and rate) of IV contrast as well as timing of scanning. Typically, most abdomi nal imaging is performed during the portal venous phase and (optimally) about the same time frame after injection on each examination (see Fig. 1 for impact of different phase of IV contrast on lesion measurement). Most solid tumours may be scanned with a single phase after administration of contrast. While triphasic CT scans are sometimes performed on other types of vascular tumours to improve lesion conspicuity, for consistency and uniformity, we would recommend tri phasic CT for hepatocellular and neuroendocrine tumours for which this scanning protocol is generally standard of care, and the improved temporal resolution of the triphasic scan will enhance the radiologists' abil ity to consistently and reproducibly measure these lesions. The precise dose and rate of IV contrast is dependent upon the CT scanning equipment, CT acqui sition protocol, the type of contrast used, the available venous access and the medical condition of the patient. Therefore, the method of administration of intravenous contrast agents is variable. Rather than try to institute rigid rules regarding methods for administering contrast agents and the volume injected, it is appropriate to suggest that an adequate volume of a suitable contrast agent should be given so that the metastases are demonstrated to best effect and a *con sistent method* is used on subsequent examinations for any given patient (ideally, this would be specified in the protocol or for an institution). It is very important that the same technique be used at baseline and on fol low up examinations for a given patient. This will greatly enhance the reproducibility of the tumour mea surements. If prior to enrolment it is known a patient is not able to undergo CT scans with IV contrast due to allergy or renal insufficiency, the decision as to whether a non contrast CT or MRI (with or without IV contrast) should be used to evaluate the subject at baseline and follow up should be guided by the tumour type under investigation and the anatomic location of the disease. For patients who develop contraindica tions to contrast after baseline contrast CT is done, the decision as to whether non contrast CT or MRI (enhanced or non enhanced) should be performed should also be based on the tumour type, anatomic location of the disease and should be optimised to allow for comparison to the prior studies if possible. Each case should be discussed with the radiologist to determine if substitution of these other approaches is possible and, if not, the patient should be considered not evaluable from that point forward. Care must be taken in measurement of target lesions on a different modality and interpretation of non target disease or new lesions, since the same lesion may appear to have a different size using a new modality (see Fig. 2 for a comparison of CT and MRI of the same lesion). Oral contrast is recommended to help visualise and differ entiate structures in the abdomen.

c. *Slice thickness and reconstruction interval*: RECIST measure ments may be performed at most clinically obtained slice thicknesses. It is recommended that CT scans be performed at 5 mm contiguous slice thickness or less and indeed this guideline presumes a minimum 5 mm thickness in recommendations for measurable lesion definition. Indeed, variations in slice thickness can have an impact on lesion measurement and on detection of new lesions. However, consideration should also be given for minimising radiation exposure. With these parameters, a minimum 10 mm lesion is considered measurable at baseline. Occasionally, institutions may perform medically acceptable scans at slice thicknesses *greater than* 5 mm. If this occurs, the minimum size of measurable lesions at baseline should be *twice the slice*

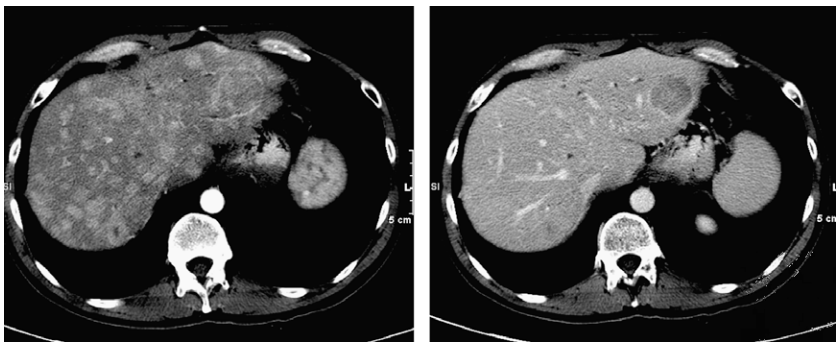


Fig. 1 – Difference in measurement/visualisation with different phases of IV contrast administration. Hypervascular metastases imaged in the arterial phase (left) and the portal venous phase (right). Note that the number of lesions visible differs greatly between the two phases of contrast administration as does any potential lesion measurement. Consistent CT scan acquisition, including phase of contrast administration, is important for optimal and reproducible tumour

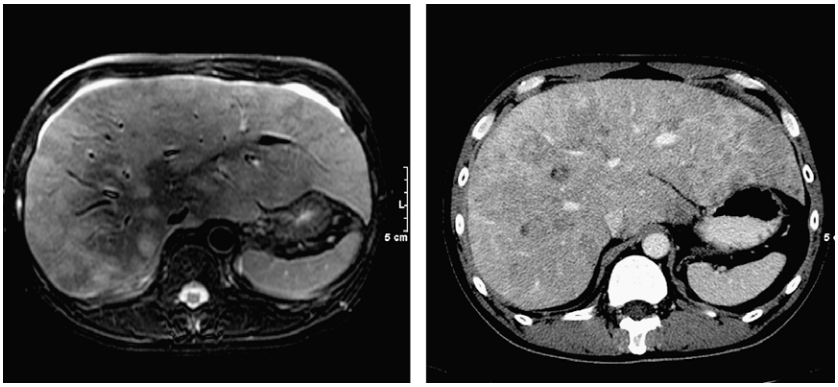


Fig. 2 – CT versus MRI of same lesions showing apparent 'progression' due only to differing method of measurement.

thickness of the baseline scans. Most contemporary CT scanners are multidetector which have many imaging options for these acquisition parameters.[23] The equip ment vendor and scanning manual should be reviewed if there are any specific system questions.

d. *Alternative contrast agents*: There are a number of other, new contrast agents, some organ specific.[24] They may be used as part of patient care for instance, in liver lesion assessment, or lymph node characterisation[25], but *should not as yet be used in clinical trials*.

FDG PET has gained acceptance as a valuable tool for detecting, staging and restaging several malignancies. Criteria for incorporating (or substituting) FDG PET into anatomical assessment of tumour response in phase II trials are not yet available, though much research is ongoing. Nevertheless, FDG PET is being used in many drug development trials both as a tool to assess therapeutic efficacy and also in assessment of progression. If FDG PET scans are included in a protocol, by consensus, an FDG uptake period of 60 min prior to imaging has been decided as the most appropriate for imaging of pa tients with malignancy.[26] Whole body acquisition is impor tant since this allows for sampling of all areas of interest and can assess if new lesions have appeared thus determining the possibility of interval progression of disease. Images from the base of the skull to the level of the mid thigh should be ob tained 60 min post injection. PET camera specifications are variable and manufacturer specific, so every attempt should be made to use the same scanner, or the same model scanner, for serial scans on the same patient. Whole body acquisitions can be performed in either 2 or 3 dimensional mode with attenuation correction, but the method chosen should be con sistent across all patients and serial scans in the clinical trial.

PET/CT scans: Combined modality scanning such as with PET CT is increasingly used in clinical care, and is a modal ity/technology that is in rapid evolution; therefore, the recom mendations in this paper may change rather quickly with time. At present, low dose or attenuation correction CT por tions of a combined PET CT are of limited use in anatomically based efficacy assessments and it is therefore suggested that they should not be substituted for dedicated diagnostic con trast enhanced CT scans for anatomically based RECIST mea surements. However, if a site can document that the CT performed as part of a PET CT is of identical diagnostic qual ity to a diagnostic CT (with IV and oral contrast) then the CT portion of the PET CT can be used for RECIST measurements. Note, however, that the PET portion of the CT introduces addi tional data which may bias an investigator if it is not routinely or serially performed.

Ultrasound examinations should not be used in clinical trials to measure tumour regression or progression of lesions be cause the examination is necessarily subjective and operator dependent. The reasons for this are several: Entire examina tions cannot be reproduced for independent review at a later date, and it must be assumed, whether or not it is the case, that the hard copy films available represent a true and accu rate reflection of events. Furthermore, if, for example, the only measurable lesion is in the para aortic region of the abdomen and if gas in the bowel overlies the lesion, the lesion will not be detected because the ultrasound beam cannot penetrate the gas. Accordingly, the disease staging (or restag ing for treatment evaluation) for this patient will not be accurate.

While evaluation of lesions by *physical examination* is also of limited reproducibility, it is permitted when lesions are superficial, at least 10 mm size, and can be assessed using calipers. In general, it is preferred if patients on clinical trials have at least one lesion that is measurable by CT. Other skin or palpable lesions may be measured on physical examina tion and be considered target lesions.

Use of *MRI* remains a complex issue. MRI has excellent contrast, spatial and temporal resolution; however, there are many image acquisition variables involved in MRI, which greatly impact image quality, lesion conspicuity and mea surement. Furthermore, the availability of MRI is variable globally. As with CT, if an MRI is performed, the technical specifications of the scanning sequences used should be optimised for the evaluation of the type and site of disease. Furthermore, as with CT, the modality used at follow up should be the same as was used at baseline and the lesions should be measured/assessed on the same pulse sequence. Generally, axial imaging of the abdomen and pelvis with T1 and T2 weighted imaging along with gadolinium enhanced imaging should be performed. The field of view, matrix, number of excitations, phase encode steps, use of fat sup pression and fast sequences should be optimised for the spe

cific body part being imaged as well as the scanner utilised. It is beyond the scope of this document or appendix to pre scribe specific MRI pulse sequence parameters for all scan ners, body parts and diseases. Ideally, the same type of scanner should be used and the image acquisition protocol should be followed as closely as possible to prior scans. Body scans should be performed with breath hold scanning tech niques if possible.

Selection of target lesions: In general, the *largest* lesions rep resentative of involved organs (up to a maximum of two per organ and five total) are selected to follow as target lesions. However, in some cases, the largest lesions may not be easily measured and are not suitable for follow up because of their configuration. In these cases, identification of the largest *most reproducible* lesions is advised. Fig. 3 provides an illustrative example where the largest lesion is not the most reproducible and another lesion is better to select and follow:

Measurement of lesions

The longest diameter of selected lesions should be measured in the plane in which the images were acquired. For body CT, this is the axial plane. In the event isotropic reconstructions are performed, measurements can be made on these recon structed images; however, it should be cautioned that not all radiology sites are capable of producing isotropic recon structions. This could lead to the undesirable situation of measurements in the axial plane at one assessment point and in a different plane at a subsequent assessment. There are some tumours, for instance paraspinal lesions, which are better measured in the coronal or sagittal plane. It would be acceptable to measure these lesions in these planes if the reconstructions in those planes were isotropic or the images were acquired with MRI in those planes. Using the same plane of evaluation, the maximal diameter of each target lesion should always be measured at subsequent follow up time points even if this results in measuring the lesion at a differ ent slice level or in a different orientation or vector compared with the baseline study. Software tools that calculate the maximal diameter for a perimeter of a tumour may be em ployed and may even reduce variability.

The only exception to the longest diameter rule is lymph node measurement. Because malignant nodes are identified by the length of their short axis, this is the guide used to determine not only whether they are pathological but is also the dimension measured for adding into the sum of target le sions. Fig. 4 illustrates this point: the large arrow identifies a malignant node: the shorter perpendicular axis is $\geqslant$15 mm and will be recorded. Close by (small arrow) there is a normal node: note here the long axis is greater than 10 mm but the short axis is well below 10 mm. This node should be consid ered non pathological.

If a lesion disappears and reappears at a subsequent time point it should continue to be measured. However, the pa tient's response at the point in time when the lesion reap pears will depend upon the status of his/her other lesions. For example, if the patient's tumour had reached a CR status and the lesion reappeared, then the patient would be consid ered PD at the time of reappearance. In contrast, if the tumour status was a PR or SD and one lesion which had disappeared then reappears, its maximal diameter should be added to the sum of the remaining lesions for a calculated response: in other words, the reappearance of an apparently 'disappeared' single lesion amongst many which remain is not in itself en

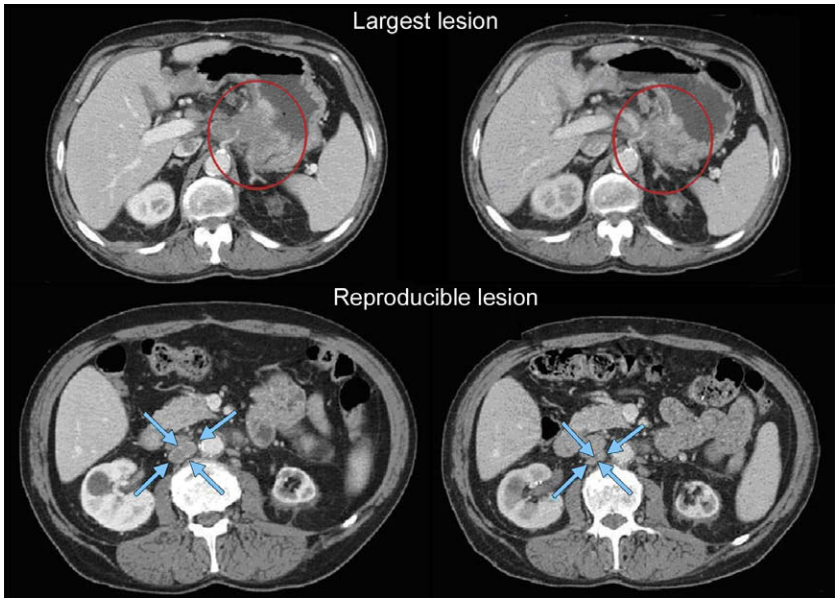


Fig. 3 – Largest lesion may not be most reproducible: most reproducible should be selected as target. In this example, the primary gastric lesion (circled at baseline and at follow-up in the top two images) may be able to be measured with thin section volumetric CT with the same degree of gastric distention at baseline and follow-up. However, this is potentially challenging to reproduce in a multicentre trial and if attempted should be done with careful imaging input and analysis. The most reproducible lesion is a lymph node (circled at baseline and at follow-up in the bottom two images).

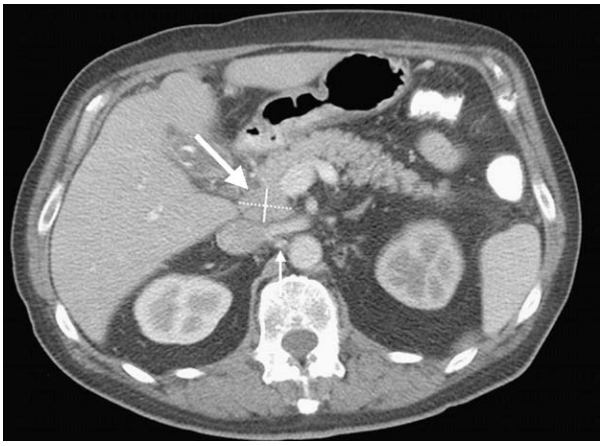


Fig. 4 – Lymph node assessment: large arrow illustrates a pathological node with the short axis shown as a solid line which should be measured and followed. Small arrow illustrates a non-pathological node which has a short axis <10 mm.

ough to qualify for PD: that requires the sum of all lesions to meet the PD criteria. The rationale for such a categorisation is based upon the realisation that most lesions do not actually 'disappear' but are not visualised because they are beyond the resolving power of the imaging modality employed.

The identification of the precise boundary definition of a lesion may be difficult especially when the lesion is embedded in an organ with a similar contrast such as the liver, pancreas, kidney, adrenal or spleen. Additionally, peritumoural oedema may surround a lesion and may be difficult to distinguish on certain modalities between this oedema and actual tumour. In fact, pathologically, the presence of tumour cells within the oedema region is variable. Therefore, it is most critical that the measurements be obtained in a reproducible manner from baseline and all subsequent follow up time points. This is also a strong reason to consistently utilise the same imaging modality.

When lesions 'fragment', the individual lesion diameters should be added together to calculate the target lesion sum. Similarly, as lesions coalesce, a plane between them may be maintained that would aid in obtaining maximal diameter measurements of each individual lesion. If the lesions have truly coalesced such that they are no longer separable, the vector of the longest diameter in this instance should be the maximal longest diameter for the 'merged lesion'.

Progression of non-target lesions

To achieve 'unequivocal progression' there must be an *overall* level of substantial worsening in non target disease that is of a magnitude that, even in the presence of SD or PR in target disease, the treating physician would feel it important to change therapy. Examples of unequivocal progression are shown in Figs. 5 and 6.

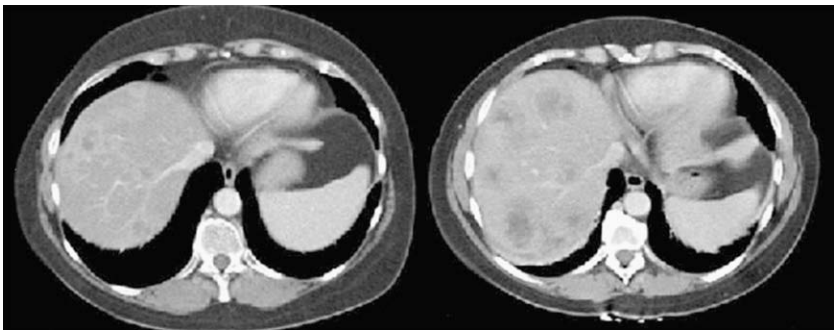


Fig. 5 – Example of unequivocal progression in non-target lesions in liver.

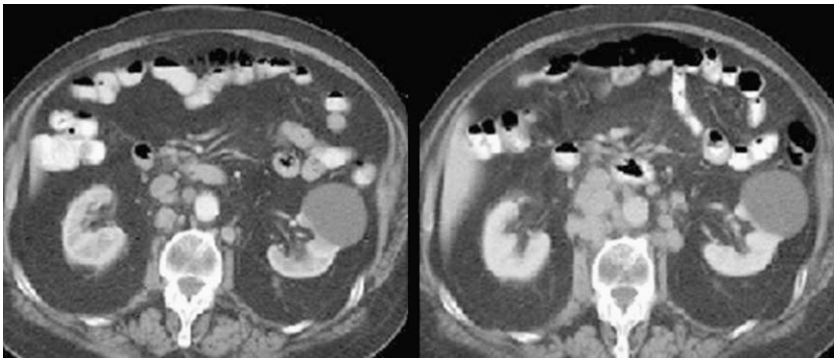


Fig. 6 – Example of unequivocal progression in non-target lesion (nodes).

Appendix III. Frequently asked questions

Question	Answer
What should be done if several unique lesions at baseline become confluent at a follow up evaluation?	Measure the longest diameter of the confluent mass and record to add into the sum of the longest diameters
How large does a new lesion have to be to count as progression? Does any small subcentimetre lesion qualify, or should the lesion be at least measurable?	New lesions do not need to meet 'measurability criteria' to be considered valid. If it is clear on previous images (with the same technique) that a lesion was absent then its definitive appearance implies progression. If there is any doubt (because of the techniques or conditions) then it is suggested that treatment continue until next scheduled assessment when, generally, all should be clear. Either it gets bigger and the date of progression is the date of the first suspicion, or it disappears and one may then consider it an artefact with the support of the radiologists
How should one lesion be measured if on subsequent exams it is split into two?	Measure the longest diameter of each lesion and add this into the sum
Does the definition of progression depend on the status of all target lesions or only one?	As per the RECIST 1.1 guideline, progression requires a 20% increase in the sum of diameters of all target lesions AND a minimum absolute increase of 5 mm in the sum
Are RECIST criteria accepted by regulatory agencies?	Many cooperative groups and members of pharma were involved in preparing RECIST 1.0 and have adopted them. The FDA was consulted in their development and supports their use, though they don't require it. The European and Canadian regulatory authorities also participated and the RECIST criteria are now integrated in the European note for guidance for the development of anticancer agents. Many pharmaceutical companies are also using them. RECIST 1.1 was similarly widely distributed before publication
What is the criterion for a measurable lesion if the CT slice thickness is >5 mm?	RECIST 1.1 recommends that CT scans have a maximum slice thickness of 5 mm and the minimum size for a measurable lesion is twice that: 10 mm (even if slice thickness is <5 mm). If scanners with slice thickness >5 mm are used, the minimum lesion size must have a longest diameter twice the actual slice thickness
What should we record when target lesions become so small they are below the 10 mm 'measurable' size?	Target lesion measurability is defined at baseline. Thereafter, actual measurements, even if <10 mm, should be recorded. If lesions become very small, some radiologists indicate they are 'too small to measure'. This guideline advises that when this occurs, if the lesion is actually still present, a default measurement of 5 mm should be applied. If in fact the radiologist believes the lesion has gone, a default measurement of 0 mm should be recorded
If a patient has several lesions which have decreased in size to meet PR criteria and one has actually disappeared, does that patient have PD if the 'disappeared' lesion reappears?	Unless the sum meets the PD criteria, the reappearance of a lesion in the setting of PR (or SD) is not PD. The lesion should simply be added into the sum. If the patients had had a CR, clearly reappearance of an absent lesion would qualify for PD
When measuring the longest diameter of target lesions in response to treatment, is the same axis that was used initially used subsequently, even if there is a shape change to the lesion that may have produced a new longest diameter?	The longest diameter of the lesion should always be measured even if the actual axis is different from the one used to measure the lesion initially (or at different time point during follow up) The only exception to this is lymph nodes: as per RECIST 1.1 the short axis should always be followed and as in the case of target lesions, the vector of the short axis may change on follow up
Target lesions have been selected at baseline and followed but then one of these target lesions then becomes non evaluable (i.e. different technique used) What is the effect this has on the other target lesions and the overall response?	What may be done in such cases is one of the following: (a) If the patient is still being treated, call the centre to be sure that future evaluations are done with the baseline technique so at least SOME courses are fully evaluable (b) If that is not possible, check if there IS a baseline exam by the same technique which was used to follow patients…in which case if you retrieve the baseline measures from that technique you retrieve the lesion evaluability (c) If neither (a) nor (b) is possible then it is a judgement call about whether you delete the lesion from all forms or consider the impact of the lesion overall is so important that its being non evaluable makes the overall response interpretation inevaluable without it. Such a decision should be discussed in a review panel It is NOT recommended that the lesion be included in baseline sums and then excluded from follow up sums since this biases in favour of a response

(*continued on next page*)

Appendix III – *continued*

Question	Answer
What if a single non target lesion cannot be reviewed, for whatever reason; does this negate the overall assessment?	Sometimes the major contribution of a single non target lesion may be in the setting of CR having otherwise been achieved: failure to examine one non target in that setting will leave you unable to claim CR. It is also possible that the non target lesion has undergone such substantial progression that it would override the target disease and render patient PD. However, this is very unlikely, especially if the rest of the measurable disease is stable or responding
A patient has a 32% decrease in sum cycle 2, a 28% decrease cycle 4 and a 33% decrease cycle 6. Does confirmation of PR have to take place in sequential scans or is a case like this confirmed PR?	It is not infrequent that tumour shrinkage hovers around the 30% mark. In this case, most would consider PR to have been confirmed looking at this overall case. Had there been two or three non PR observations between the two time point PR responses, the most conservative approach would be to consider this case SD
In the setting of a breast cancer neoadjuvant study, would mammography not be used to assess lesions? Is CT preferred in this setting?	Neither CT nor mammography are optimal in this setting. MRI is the preferred modality to follow breast lesions in a neoadjuvant setting
A patient has a lesion measurable by clinical exam and by CT scan. Which should be followed?	CT scan. Always follow by imaging if that option exists since it can be reviewed and verified
A lesion which was solid at baseline has become necrotic in the centre. How should this be measured?	The longest diameter of the entire lesion should be followed. Eventually, necrotic lesions which are responding to treatment decrease in size. In reporting the results of trials, you may wish to report on this phenomenon if it is seen frequently since some agents (e.g. angiogenesis inhibitors) may produce this effect
If I am going to use MRI to follow disease, what is minimum size for measurability?	MRI may be substituted for contrast enhanced CT for some sites, but not lung. The minimum size for measurability is the same as for CT (10 mm) as long as the scans are performed with slice thickness of 5 mm and no gap. In the event the MRI is performed with thicker slices, the size of a measurable lesion at baseline should be two times the slice thickness. In the event there are inter slice gaps, this also needs to be considered in determining the size of measurable lesions at baseline
Can PET CT be used with RECIST?	At present, the low dose or attenuation correction CT portion of a combined PET CT is not always of optimal diagnostic CT quality for use with RECIST measurements. However, if your site has documented that the CT performed as part of a PET CT is of the same diagnostic quality as a diagnostic CT (with IV and oral contrast) then the PET CT can be used for RECIST measurements. Note, however, that the PET portion of the CT introduces additional data which may bias an investigator if it is not routinely or serially performed

REFERENCES

1. Paesmans M, Sculier JP, Libert P, et al. Response to chemotherapy has predictive value for further survival of patients with advanced non small cell lung cancer: 10 years experience of the European Lung Cancer Working Party. *Eur J Cancer* 1997;**33**:2326 32.
2. Buyse M, Thirion P, Carlson RW, et al. Relation between tumor response to first line chemotherapy and survival in advanced colorectal cancer: a meta analysis. Meta analysis group in Cancer. *Lancet* 2000;**356**:373 8.
3. Goffin J, Baral S, Tu D, et al. Objective responses in patients with malignant melanoma or renal cell cancer in early clinical studies do not predict regulatory approval. *Clin Cancer Res* 2005;**15**:5928 34.
4. El Maraghi RH, Eisenhauer EA. Review of phase II trial designs used in studies of molecular targeted agents: outcomes and predictors of success in phase III. *J Clin Oncol* 2008;**10**:1346 54.
5. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981;**47**:207 14.
6. Tonkin K, Tritchler D, Tannock I. Criteria of tumor response used in clinical trials of chemotherapy. *J Clin Oncol* 1985;**3**:870 5.
7. Baar J, Tannock I. Analyzing the same data in two ways: a demonstration model to illustrate the reporting and misreporting of clinical trials. *J Clin Oncol* 1989;**7**:969 78.
8. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors (RECIST Guidelines). *J Natl Cancer Inst* 2000;**92**:205 16.
9. Therasse P, Eisenhauer EA, Verweij J. RECIST revisited: a review of validation studies on tumour assessment. *Eur J Cancer* 2006;**42**:1031 9.
10. Bogaerts J, Ford R, Sargent D, et al. Individual patient data analysis to assess modifications to the RECIST criteria. *Eur J Cancer* 2009;**45**:248 60.
11. Moskowitz CS, Jia X, Schwartz LH, Gönen M. A simulation study to evaluate the impact of the number of lesions measured on response assessment. *Eur J Cancer* 2009;**45**:300 10.
12. Sargent D, Rubinstein L, Schwartz L, et al. Validation of novel imaging methodologies for use as cancer clinical trials end points. *Eur J Cancer* 2009;**45**:290 9.

13. Macdonald DR, Cascino TL, Schold Jr SC, Cairncross JG. Response criteria for phase II studies of supratentorial malignant glioma. *J Clin Oncol* 1990;**8**:1277 80.
14. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. *J Clin Oncol* 2007;**10**:579 86.
15. Schwartz LH, Bogaerts J, Ford R, et al. Evaluation of lymph nodes with RECIST 1.1. *Eur J Cancer* 2009;**45**:261 7.
16. Rustin GJ, Quinn M, Thigpen T, et al. Re: New guidelines to evaluate the response to treatment in solid tumors (ovarian cancer). *J Natl Cancer Inst* 2004;**96**:487 8.
17. Bubley GJ, Carducci M, Dahut W, et al. Eligibility and response guidelines for phase II clinical trials in androgen independent prostate cancer: recommendations from the Prostate Specific Antigen Working Group. *J Clin Oncol* 1999;**17**:3461 7.
18. Scher H, Halabi S, Tannock I, et al. Design and end points of clinical trials for patients with progressive prostate cancer and castrate levels of testosterone: recommendations of the Prostate Cancer Clinical Trials Working Group. *J Clin Oncol* 2008;**26**:1148 59.
19. Vergote I, Rustin GJ, Eisenhauer EA, et al. Re: new guidelines to evaluate the response to treatment in solid tumors [ovarian cancer]. Gynecologic Cancer Intergroup. *J Natl Cancer Inst* 2000;**92**:1534 5.
20. Van Glabbeke M, Verweij J, Judson I, Nielsen OS. EORTC Soft Tissue and Bone Sarcoma Group: Progression free rate as the principal end point for phase II trials in soft tissue sarcomas. *Eur J Cancer* 2002;**38**:543 9.
21. Dancey JE, Dodd LE, Ford R, et al. Recommendations for the assessment of progression in randomised cancer treatment trials. *Eur J Cancer* 2009;**45**:281 9.
22. Ford R, Schwartz L, Dancey J, et al. Lessons learned from independent central review. *Eur J Cancer* 2009;**45**: 268 74.
23. Catalano C, Francone M, Ascarelli A, Mangia M, Iacucci I, Passariello R. Optimizing radiation dose and image quality. *Eur Radiol* 2007;**17**(Suppl 6):F26 32.
24. Low RN. Abdominal MRI advances in the detection of liver tumours and characterization. *Lancet Oncol* 2007;**8**(6):525 35.
25. Barrett T, Choyke PL, Kobayashi H. Imaging of the lymphatic system: new horizons. *Contrast Media Mol Imaging* 2006;**1**(6):230 45.
26. Shankar LK, Hoffman JM, Bacharach S, et al. National Cancer Institute. Consensus recommendations for the use of 18F FDG PET as an indicator of therapeutic response in patients in National Cancer Institute Trials. *J Nucl Med* 2006;**47**(6):1059 66.