STATISTICAL ANALYSIS PLAN

STUDY TITLE: **A PHASE III, MULTICENTER, RANDOMIZED, OPEN-LABEL STUDY OF ATEZOLIZUMAB (ANTI-PD-L1 ANTIBODY) PLUS BEVACIZUMAB VERSUS ACTIVE SURVEILLANCE AS ADJUVANT THERAPY IN PATIENTS WITH HEPATOCELLULAR CARCINOMA AT HIGH RISK OF RECURRENCE AFTER SURGICAL RESECTION OR ABLATION**

STUDY NUMBER:	WO41535
STUDY NAME:	IMbrave050
VERSION NUMBER:	4
ROCHE COMPOUND(S):	Atezolizumab (RO5541267) Bevacizumab (RO4876646)
EUDRACT NUMBER:	2019-002 491-14
IND NUMBER:	135913
NCT NUMBER:	NCT04102098
PLAN PREPARED BY:	[redacted], Ph.D.

STATISTICAL ANALYSIS PLAN APPROVAL

SPONSOR:	F. Hoffmann-La Roche Ltd
LEGAL REGISTERED ADDRESS:	Grenzacherstrasse 124 4070 Basel, Switzerland
DATE FINAL:	See electronic date stamp above

CONFIDENTIAL

This is an F. Hoffmann-La Roche Ltd document that contains confidential information. Nothing herein is to be disclosed without written consent from F. Hoffmann-La Roche Ltd.

STATISTICAL ANALYSIS PLAN VERSION HISTORY

This SAP was developed based on Roche SAP model document (Version 2.0).

SAP Version	Approval Date	Based on Protocol (Version, Approval Date)
4	see electronic date stamp on title page	V 4.0, 30 November 2021
3	see electronic date stamp on title page	V 4.0, 30 November 2021
2	see electronic date stamp on title page	V 4.0, 30 November 2021
1	see electronic date stamp on title page	V 3.0, 03 December 2020

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE

Key changes to the SAP, along with the rationale(s) for each change, are summarized below.

Section	Description of Change	Rationale for Change
Section 5.4.2.2	The language to differentiate patients by their survival status in the censoring rule has been removed so that patients who have not experienced disease recurrence will be censored at the date of the last assessment for hepatocellular carcinoma (HCC) occurrence	The last assessment of HCC occurrence date is the last known event-free date for patients who have not experienced disease recurrence at the time of analysis regardless of their survival status.
Section 5.4.2.5	The language has been updated to clarify the analysis for time to extrahepatic spread (EHS) or macrovascular invasion will be conducted for the study period prior to crossover	Tumor assessment schedule after the first HCC recurrence is different for patients who crossed over. Patients who crossed over will continue to have scheduled tumor assessments, while patients in the experiment arm and patients in the active surveillance arm who did not cross over will discontinue the assessment of HCC occurrence after the first documented occurrence of HCC
Section 5.3.4.1	α-fetoprotein (AFP) has been removed from the subgroup analysis	AFP is not clinically relevant in the HCC adjuvant setting
Section 5.5	The definition of recurrence-free survival (RFS) after the first HCC recurrence for patients with no evidence of disease (NED) at the time of crossover has been updated to the time from first exposure to any dose of crossover treatment to the second documented HCC recurrence or death from any cause (whichever occurs first), as determined by the investigator. The definition of progression-free survival (PFS) after the first HCC recurrence for patients with measurable disease at the time of crossover has been updated to the time from first exposure to any dose of crossover treatment to the first documented occurrence of disease progression beyond the initial unresectable disease recurrence as determined by the investigator according to Response Evaluation Criteria in	The change is made due to variable time from first HCC recurrence to the first exposure in patients

	Solid Tumors (RECIST) v1.1, or death from any cause (whichever occurs first)	
Section 5.7.8	The language has been updated to clarify that the China subpopulation is of patients enrolled in mainland China, Taiwan and Hong Kong	To include all patients enrolled in mainland China, Taiwan and Hong Kong reflecting the latest common practice.

Additional minor changes have been made throughout to improve clarity and consistency.

TABLE OF CONTENTS

STATISTICAL ANALYSIS PLAN AMENDMENT RATIONALE 3

1. INTRODUCTION 10

1.1 Objectives and Endpoints 10

1.2 Study Design 13

1.2.1 Treatment Assignment and Blinding 17

1.2.2 Independent Review Facility 18

1.2.3 Data Monitoring 18

2. STATISTICAL HYPOTHESES 18

3. SAMPLE SIZE DETERMINATION 19

4. ANALYSIS SETS 20

4.1 Intent-to-Treat Population 20

4.2 PD-L1-high subgroup 20

4.3 Safety Population 20

4.4 Pharmacokinetic-Evaluable Population 21

4.5 Immunogenicity analysis population 21

5. STATISTICAL ANALYSES 21

5.1 General Consideration 21

5.2 Patient Disposition 21

5.3 Primary Endpoint Analysis 21

5.3.1 Definition of Primary Endpoint 21

5.3.2 Main Analytical Approach for Primary Endpoint 22

5.3.3 Sensitivity Analyses for Primary Endpoint 23

5.3.3.1 Impact of Missing Scheduled Tumor Assessments on Primary Endpoint IRF-Assessed RFS 23

5.3.3.2 Sensitivity Analyses on IRF-Assessed RFS for Patients in Any Treatment Arm Receiving Non-Protocol Therapy, Resection, or Ablation 23

5.3.3.3 Sensitivity Analyses on IRF-Assessed RFS for Missing Data 23

5.3.3.4 Sensitivity Analyses on IRF-Assessed RFS for Patients with Baseline Disease 24

5.3.4 Supplementary Analyses for Primary Endpoint 24
5.3.4.1 Subgroup Analyses for Primary Endpoint 24
5.4 Secondary Endpoints Analyses 24
5.4.1 Key Secondary Endpoint: Overall Survival 24
5.4.1.1 Sensitivity Analyses on OS for Patients in Control Arm Receiving Atezolizumab, Bevacizumab, Immunotherapy, Resection, or Ablation 25
5.4.2 Supportive Secondary Endpoint(s) 25
5.4.2.1 Investigator-Assessed Recurrence Free Survival 25
5.4.2.2 Time to Recurrence 25
5.4.2.3 IRF-Assessed and Investigator-Assessed RFS at 24-Month and 36-Month Landmark Analysis 25
5.4.2.4 Overall Survival 24-Month and 36-Month Landmark Analysis 26
5.4.2.5 Time to Extrahepatic Spread or Macrovascular Invasion 26
5.4.2.6 Recurrence-Free Survival among Patients with PD-L1-High Tumors 26
5.5 Exploratory Endpoints Analysis 26
5.6 Safety Analyses 27
5.6.1 Extent of Exposure 28
5.6.2 Adverse Events 28
5.6.3 Laboratory Data 28
5.6.4 Vital Signs 28
5.6.5 Electrocardiogram (ECG) 28
5.7 Other Analyses 29
5.7.1 Summaries of Conduct of Study 29
5.7.2 Summaries of Treatment Group Comparability 29
5.7.3 Pharmacokinetic Analyses 29
5.7.4 Immunogenicity Analyses 29
5.7.5 Biomarker Analyses 30
5.7.6 Health Utility Analyses 30
5.7.7 Analyses of Subgroups of Interest 30
5.7.8 Analyses of China Subpopulation 30
5.8 Interim Analyses 31

5.8.1 Planned Interim Analyses 31
6. SUPPORTING DOCUMENTATION 34
7. REFERENCES 36

LIST OF TABLES

Table 1 Objectives and Corresponding Endpoints 10
Table 2 High Risk Criteria for This Study 14
Table 3 Analysis Timing and Stopping Boundaries for Recurrence-Free Survival Analyses 31
Table 4 Projected Interim and Final OS Analysis Characteristics When Statistical Significance is Reached at the Interim Analysis for RFS 33
Table 5 Projected Interim and Final OS Analysis Characteristics When Statistical Significance is Reached at the Final Analysis for RFS 34

LIST OF FIGURES

Figure 1 Study Schema 15

LIST OF APPENDICES

Appendix 1 Changes to Protocol-Planned Analyses 35

LIST OF ABBREVIATIONS AND DEFINITIONS OF TERMS

Abbreviation or Term	Description
ADA	anti-drug antibody
CSR	Clinical Study Report
CR	complete response
CT	computed tomography
CI	confidence interval
ECOG	Eastern Cooperative Oncology Group
ECG	electrocardiogram
EC	Ethics Committee
EORTC	European Organisation for Research and Treatment of Cancer
EQ-5D-5L	EuroQol 5-Dimension, 5-Level Questionnaire
EHS	extrahepatic spread
HR	hazard ratio
HCC	hepatocellular carcinoma
iDCC	independent Data Coordinating Committee
iDMC	independent Data Monitoring Committee
IRF	Independent Review Facility
IRB	Institutional Review Board
ITT	intent-to-treat (population)
IxRS	interactive voice or web-based response system
IV	intravenous
IL42–EORTC QLQ-C30 (Reduced)	Item list 42–European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 (reduced version of the EORTC QLQ-C30)
MRI	magnetic resonance imaging
MedDRA	Medical Dictionary for Regulatory Activities
MWA	microwave ablation
MDD	minimum detectable difference
NCI CTCAE v 5.0	National Cancer Institute Common Terminology Criteria for Adverse Events, Version 5.0
NED	no evidence of disease
ORR	objective response rate
OS	overall survival
PR	partial response
PK	pharmacokinetic
Vp	portal vein

PD-L1	Programmed death-ligand 1
PFS	progression-free survival
RFA	radiofrequency ablation
RFS	recurrence-free survival
RECIST v1.1	Response Evaluation Criteria in Solid Tumors, Version 1.1
SAP	Statistical Analysis Plan
TTR	time to recurrence
TACE	transarterial chemoembolization
TAP	tumor area (PD-L1) positive

1. INTRODUCTION

This Statistical Analysis Plan (SAP) describes the analyses that are planned to be performed for the Clinical Study Report (CSR) of WO41535 (IMbrave050). Detailed background information of the study can be found in the study protocol (Version 4). The analyses described in this SAP will supersede those specified in Protocol IMbrave050 for the purposes of a regulatory filing.

1.1 OBJECTIVES AND ENDPOINTS

This is a Phase III, multicenter, randomized, open-label study of atezolizumab (anti-Programmed death-ligand 1 [PD-L1] monoclonal antibody) plus bevacizumab (anti-VEGF monoclonal antibody) versus active surveillance as adjuvant therapy in patients with hepatocellular carcinoma (HCC) at high risk of recurrence after surgical resection or ablation. Specific objectives and corresponding endpoints for the study are outlined in Table 1.

Table 1 Objectives and Corresponding Endpoints

Primary Efficacy Objective (Primary Objective)	Corresponding Endpoint
• To evaluate the efficacy of atezolizumab plus bevacizumab compared with active surveillance	• RFS after randomization, defined as the time from randomization to the first documented occurrence of intrahepatic or extrahepatic HCC as determined by an IRF, or death from any cause (whichever occurs first)
Secondary Efficacy Objective	**Corresponding Endpoints**
• To evaluate the efficacy of atezolizumab plus bevacizumab compared with active surveillance	• OS after randomization, defined as the time from randomization to death from any cause • RFS after randomization as determined by the investigator • TTR after randomization, defined as the time from randomization to first documented occurrence of intrahepatic or extrahepatic HCC, as determined by the investigator and by an IRF • IRF-assessed RFS and investigator-assessed RFS rate at 24 and 36 months after

Table 1 Objectives and Corresponding Endpoints (cont.)

Secondary Efficacy Objective	Corresponding Endpoints
• To evaluate the efficacy of atezolizumab plus bevacizumab compared with active surveillance (contd..)	randomization • OS rate at 24 months and 36 months, defined as the proportion of patients who have not experienced death from any cause at 24 and 36 months after randomization, respectively • Time to EHS or macrovascular invasion after randomization, defined as the time from randomization to the first appearance of EHS or macrovascular invasion, as determined by the investigator • RFS after randomization as determined by the investigator and by an IRF, among patients in the PD-L1-high subgroup
Exploratory Efficacy Objective	**Corresponding Endpoints**
• To evaluate the efficacy of atezolizumab plus bevacizumab compared with active surveillance	• Change from baseline in physical functioning, emotional functioning, role functioning, social functioning, and global health status/quality of life scores as assessed by the IL42–EORTC QLQ-C30 (Reduced) at specified timepoints (including post-treatment [Arm A] or surveillance [Arm B] and post-recurrence) • OS among patients in the PD-L1-high subgroup • For patients randomized to active surveillance who recur and cross over to treatment with atezolizumab plus bevacizumab and have NED at the time of crossover: – RFS after first HCC recurrence, defined as the time from first exposure to any dose of crossover treatment to the second documented HCC recurrence as determined by the investigator, or death from any cause (whichever occurs first). HCC recurrence is defined as occurrence of intrahepatic or extrahepatic HCC. • For patients randomized to active surveillance

Table 1 Objectives and Corresponding Endpoints (cont.)

Exploratory Efficacy Objective	Corresponding Endpoints
• To evaluate the efficacy of atezolizumab plus bevacizumab compared with active surveillance (contd..)	who recur and cross over to treatment with atezolizumab plus bevacizumab and who have measurable disease at the time of crossover: – PFS after first HCC recurrence, defined as the time from first exposure to any dose of crossover treatment to the first documented occurrence of disease progression beyond the initial unresectable disease recurrence as determined by the investigator according to RECIST v1.1 or death from any cause (whichever occurs first), – ORR, defined as the proportion of patients with a complete or partial response, as determined by the investigator according to RECIST v1.1
Safety Objective	**Corresponding Endpoints**
• To evaluate the safety of atezolizumab plus bevacizumab compared with active surveillance	• Incidence and severity of adverse events, with severity determined according to NCI CTCAE v5.0 • Change from baseline in targeted vital signs • Change from baseline in targeted clinical laboratory test results
Pharmacokinetic Objective	**Corresponding Endpoint**
• To characterize the PK profile of atezolizumab when given in combination with bevacizumab	• Serum concentration of atezolizumab at specified timepoints
Immunogenicity Objective	**Corresponding Endpoint**
• To evaluate the immune response to atezolizumab	• Prevalence of ADAs to atezolizumab at baseline and incidence of ADAs to atezolizumab during the study
Exploratory Immunogenicity Objective	**Corresponding Endpoint**
• To evaluate potential effects of ADAs	• Relationship between treatment-emergent ADA status and efficacy, safety, or PK endpoints

Table 1 Objectives and Corresponding Endpoints (cont.)

Exploratory Biomarker Objective	Corresponding Endpoints
• To identify and/or evaluate biomarkers that are predictive of response to study treatment (i.e., predictive biomarkers), are early surrogates of efficacy, are associated with progression to a more severe disease state (i.e., prognostic biomarkers), are associated with acquired resistance to study treatment, can provide evidence of study treatment activity (i.e., pharmacodynamic biomarkers), or can increase the knowledge and understanding of disease biology and drug safety	• Relationship between biomarkers in blood, serum, plasma, and tumor tissue and efficacy, safety, PK, immunogenicity, or other biomarker endpoints
Exploratory Health Status Utility Objective	**Corresponding Endpoint**
• To evaluate health status utility scores of patients treated with atezolizumab plus bevacizumab compared with active surveillance	• Change from baseline in EQ-5D-5L index-based and VAS scores at specified timepoints (including post-treatment [Arm A] or surveillance [Arm B] and post-recurrence)

ADA=anti-drug antibody; EHS=extrahepatic spread; EQ-5D-5L=EuroQol 5-Dimension, 5-Level Questionnaire; HCC=hepatocellular carcinoma; IL42–EORTC QLQ-C30(Reduced)=IL42-European Organisation for Research and Treatment of Cancer Quality of Life-Core 30 Questionnaire (Reduced); IRF=Independent Review Facility; NCI CTCAE v5.0=National Cancer Institute Common Terminology Criteria for Adverse Events, Version 5.0; NED=no evidence of disease; ORR=objective response rate; OS=overall survival; PD-L1=Programmed death-ligand 1; PFS=progression-free survival; PK=pharmacokinetic; RECIST v1.1=Response Evaluation Criteria in Solid Tumors, Version 1.1; RFS=recurrence-free survival; TTR=time to recurrence; VAS=visual analog scale.

1.2 STUDY DESIGN

This is a Phase III, global, multicenter, open-label, two-arm, randomized study designed to evaluate the efficacy and safety of adjuvant therapy with atezolizumab plus bevacizumab compared with active surveillance in patients with completely resected or ablated HCC who are at high risk for disease recurrence.

Patients who have undergone surgical resection may have received one cycle of adjuvant transarterial chemoembolization (TACE) prior to study entry (randomization), if deemed appropriate by the investigator and if consistent with local standards of care.

The definition of high-risk is based on composite criteria including size of the largest tumor, number of tumors, and presence of either microvascular invasion (defined as the presence of microscopic tumor emboli within the central vein, the portal vein, or large capsular vessels), minor macrovascular invasion of the portal vein (Vp1 or Vp2), or poorly differentiated microscopic appearance (histologic Grade 3 or 4). The criteria for high risk of HCC recurrence used in this study are presented by type of curative treatment in Table 2.

Table 2 High Risk Criteria for This Study

Curative Treatment	Criteria for High Risk of HCC Recurrence
Resection	• Up to three tumors, with largest tumor > 5 cm regardless of vascular invasion (microvascular invasion or minor macrovascular portal vein invasion of the portal vein-Vp1/Vp2), or poor tumor differentiation (Grade 3 or 4)[b]
	• Four or more tumors, with largest tumor ≤5 cm regardless of vascular invasion (microvascular invasion or minor macrovascular invasion of the portal vein-Vp1/Vp2), or poor tumor differentiation (Grade 3 or 4)[b]
	• Up to three tumors, with largest tumor ≤5 cm with vascular invasion (microvascular invasion or minor macrovascular invasion of the portal vein-Vp1/Vp2), and/or poor tumor differentiation (Grade 3 or 4)[b]
Ablation[a]	• Single tumor >2 cm but ≤5 cm
	• Multiple tumors (up to four tumors), all ≤5 cm

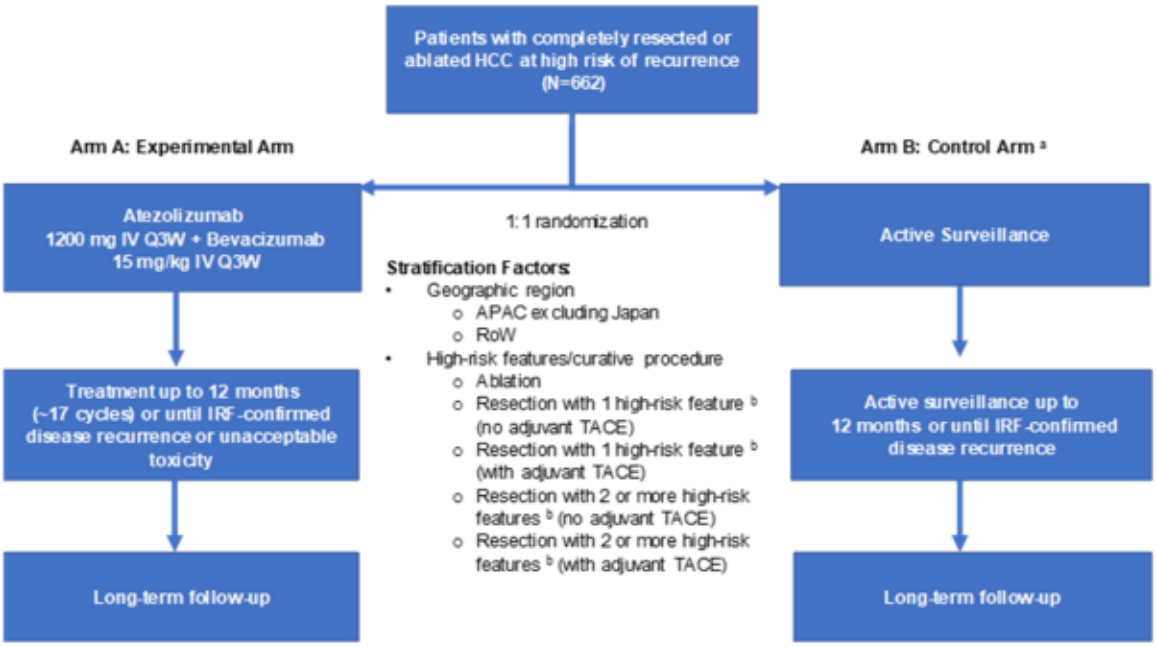
HCC=hepatocellular carcinoma; MWA=microwave ablation; RFA=radiofrequency ablation.

[a] Ablation must be RFA or MWA.

[b] In cases where a patient has evidence of mixed tumor differentiation, the worst differentiation status rather than the predominant differentiation status should be used to characterize high risk criteria.

An overview of the study design is shown in Figure 1.

Figure 1 Study Schema



APAC=Asia Pacific; HCC=hepatocellular carcinoma; IRF= Independent Review Facility; IV=intravenous; Q3W=every 3 weeks; ROW=rest of world; TACE=transarterial chemoembolization.

[a] Patients who are enrolled in Arm B who have documented recurrence during the active surveillance period or during follow-up and meet the criteria outlined in Section 4.1.1.1 of the protocol will have the option to cross over to treatment with atezolizumab 1200 mg Q3W plus bevacizumab 15 mg/kg Q3W.

[b] Risk features for resection include tumor size >5 cm, tumor number >3, vascular invasion (microvascular invasion or macrovascular invasion-Vp1/Vp2-of the portal vein), and poor tumor differentiation (defined as Grade 3 or 4).

This study will enroll approximately 662 patients from up to 175 sites in the United States, Europe, and Asia Pacific (including China) in a global enrollment phase. After completion of the global enrollment phase, additional patients may be enrolled in China in an extended China enrollment phase to ensure a minimum of approximately 250 patients in a China subpopulation. The global population will include all patients enrolled during the global enrollment phase (including patients enrolled at sites in China during that phase), and the China subpopulation will include all patients enrolled at sites in China (i.e., during both the global enrollment phase and the extended China enrollment phase).

Patients who withdraw from the study will not be replaced. If necessary, the Sponsor may choose to limit the number of patients with specific curative treatment procedures (resection or ablation).

Patients who fail their first screening for study eligibility may qualify for one re-screening opportunity (for a maximum of two screenings per patient) at the investigator's discretion. The investigator will record reasons for screen failure in the screening log.

Eligible patients will be randomized within 4–12 weeks of surgical resection or ablation (only radiofrequency ablation [RFA] or microwave ablation [MWA]) in a 1:1 ratio to one of the following two arms:

- Arm A (experimental arm): atezolizumab 1200 mg and bevacizumab 15 mg/kg, both administered by intravenous (IV) infusion on Day 1 of each 21-day cycle
- Arm B (control arm): active surveillance

Randomization will be stratified according to the following stratification factors:

- Geographic region
  - Asia Pacific excluding Japan
  - Rest of World
- High risk features/curative procedure
  - Ablation
  - Resection with 1 high risk feature (no adjuvant TACE)
  - Resection with 1 high risk feature (with adjuvant TACE)
  - Resection with 2 or more high risk features (no adjuvant TACE)
  - Resection with 2 or more high risk features (with adjuvant TACE)

Risk features for resection include tumor size >5 cm, tumor number >3, vascular invasion (microvascular invasion or macrovascular invasion-Vp1/Vp2-of the portal vein), and poor tumor differentiation (defined as Grade 3 or 4).

Treatment with atezolizumab+bevacizumab or active surveillance will continue for up to 12 months (or 17 cycles, whichever occurs first) or until Independent Review Facility (IRF)-confirmed disease recurrence or unacceptable toxicity (for patients receiving atezolizumab+bevacizumab), whichever occurs first. Patients may continue treatment beyond 12 months due to delayed or missed visits if deemed in the best interest of the patient by the investigator. After 12 months (or 17 cycles, whichever occurs first), information on recurrence, survival and subsequent anti-cancer therapies will be collected every 12 weeks until death (unless the patient withdraws consent or the Sponsor terminates the study).

Patients randomized to Arm A, who transiently withhold or permanently discontinue bevacizumab for adverse events, may continue on single-agent atezolizumab as long as they have not met the protocol-defined criteria for HCC recurrence. If atezolizumab is discontinued, bevacizumab should also be discontinued. If atezolizumab is transiently withheld for adverse events, bevacizumab should also be held.

Patients will undergo imaging assessments at every 12 weeks (± 7 days) during the study. Patients are required to undergo tumor biopsy sample collection at the time of radiographic confirmation of disease recurrence unless not clinically feasible as assessed and documented by the investigator to evaluate tumor tissue biomarkers related to mechanisms of acquired resistance and disease recurrence and clinical benefit of atezolizumab+bevacizumab.

Treatment with Atezolizumab+Bevacizumab following Recurrence for Patients in Arm B

Patients randomized to Arm B who experience documented IRF-assessed HCC recurrence per protocol-defined criteria either during the 12-month active surveillance period or during long-term follow-up and meet all eligibility criteria will be offered the option of crossing over to treatment with atezolizumab and bevacizumab. Patients can cross over directly after the first recurrence or after resection or ablation for the first recurrence (resulting in no evidence of disease [NED]). Patients undergoing resection are allowed to have 1 cycle of adjuvant TACE prior to crossover, if deemed appropriate by the investigator and if consistent with local standards of care. Day 1 of Cycle 1 of crossover treatment must be no later than 12 weeks after documentation of recurrence. Results of tests or examinations performed for crossover and per the relevant protocol-defined window may be used for screening assessments rather than repeating such tests.

Patients who transiently withhold or permanently discontinue bevacizumab for adverse events may continue on single-agent atezolizumab as long as they have not met the criteria for HCC recurrence (for patients with NED at crossover) or are experiencing clinical benefit in the opinion of the investigator (for patients with measurable disease at crossover) and after discussion with the Medical Monitor. If atezolizumab is discontinued, bevacizumab should also be discontinued. If atezolizumab is transiently withheld for adverse events, bevacizumab should also be held

Patients who crossover will undergo imaging assessments at scheduled intervals during the study.

1.2.1 Treatment Assignment and Blinding

This is a randomized, open-label study. After written informed consent has been obtained, all screening procedures and assessments have been completed, and eligibility has been established for a patient, the study site will obtain the patient's identification number and treatment assignment from the interactive voice or web-based response system (IxRS).

Patients will be randomly assigned to one of two study arms: Arm A (atezolizumab+bevacizumab) or Arm B (active surveillance). Randomization will occur in a 1:1 ratio through use of a permuted-block randomization method to ensure a

balanced assignment to each study arm. Randomization will be stratified according to the criteria outlined in Section 1.2.

1.2.2 Independent Review Facility

An IRF will be used to enable centralized, independent reviews of images and other clinical data (e.g., histopathology, tumor markers etc.) used for assessment of HCC recurrence. IRF reviews will be performed prior to the pre-specified efficacy analyses. IRF membership and procedures will be detailed in an IRF Charter.

1.2.3 Data Monitoring

An independent Data Monitoring Committee (iDMC) will evaluate safety and efficacy data during the study. Sponsor affiliates will be excluded from iDMC membership. The iDMC will follow a charter that outlines the iDMC roles and responsibilities.

Unblinded safety data will be reviewed on a periodic basis, approximately every 6 months from the time of enrollment of the first patient. Unblinded efficacy data will be reviewed as a part of the interim analysis of IRF-assessed recurrence-free survival (RFS), scheduled to occur when approximately 236 RFS events have occurred. All summaries and analyses for the iDMC review will be prepared by an independent Data Coordinating Center (iDCC).

After reviewing the data, the iDMC will provide a recommendation to the Sponsor as described in the iDMC Charter. Final decisions will rest with the Sponsor.

Any outcomes of these data reviews that affect study conduct will be communicated in a timely manner to the investigators for notification of their respective Institutional Review Boards or Ethics Committees (IRBs/ECs).

2. STATISTICAL HYPOTHESES

Hypotheses will be formally tested on primary endpoint (IRF-assessed RFS) and the key secondary endpoint (overall survival [OS]). Implementation of the statistical testing procedure will strongly control the overall type 1 error at 5% (two-sided).

The null hypothesis of no difference in IRF-assessed RFS between the two treatment arms in the intent-to-treat (ITT) population will be tested with the overall type I error controlled at a two-sided significance level of 0.05. The null and alternative hypotheses regarding IRF-assessed RFS can be phrased in terms of survival functions $S_A(t)$ and $S_B(t)$ for Arm A (atezolizumab+bevacizumab) and Arm B (active surveillance), respectively:

$$H_0: S_A(t) = S_B(t) \text{ vs. } H_1: S_A(t) \neq S_B(t)$$

The two-sided stratified log-rank test will be used as the primary analysis to compare RFS between the two arms. The stratification factors to be included in the stratified

analyses are described in Section 5.3.2. The unstratified log-rank test will be used to check the robustness of the results of the stratified log-rank test.

Overall survival (OS) will be a key secondary endpoint of this study. If the statistical significance is achieved in IRF-assessed RFS, then the null hypothesis of no difference in OS between two arms in the ITT population will be tested using a stratified log-rank test at a two-sided significance level of 0.05. The null and alternative hypotheses in terms of the survival functions $S_{OS\_A}(t)$ and $S_{OS\_B}(t)$ in Arm A and Arm B are phrased as below, respectively:

$$H_0: S_{OS\_A}(t) = S_{OS\_B}(t) \text{ versus } H_1: S_{OS\_A}(t) \neq S_{OS\_B}(t)$$

3. SAMPLE SIZE DETERMINATION

A total of approximately 662 patients will be randomized in the global enrollment phase of this study using a 1:1 randomization ratio to allocate patients to either the atezolizumab+bevacizumab arm (Arm A) or the active surveillance arm (Arm B). The primary efficacy endpoint for this study is IRF-assessed RFS, defined as the time from randomization to the first documented occurrence of intrahepatic or extrahepatic HCC, or death from any cause (whichever occurs first).

The sample size of this study is based on the number of RFS events required to demonstrate efficacy with regard to the primary efficacy endpoint of IRF-assessed RFS. To detect an improvement in RFS using a log-rank test at a two-sided significance level of 0.05, approximately 323 RFS events will be required to achieve 80% overall power assuming a target hazard ratio (HR) of 0.73 (median RFS improvement over active surveillance of 7.4 months). The minimum detectable difference (MDD) for RFS is a HR of 0.8 (median RFS improvement of 5 months).

The calculation of sample size and estimates of the RFS analysis timelines are based on the following assumptions:

- Patients will be randomized to Arm A and Arm B in a 1:1 ratio.
- RFS follows a one-piece exponential distribution.
- The median RFS in Arm B is 20 months.
- The O'Brien-Fleming boundaries approximated by the Lan-DeMets method (Gordon Lan and DeMets 1983) will be used as stopping boundaries for the interim and final analyses of RFS.
- The dropout rate is 15% for Arm A and 20% for Arm B over 12 months.

The recruitment of approximately 662 patients will take place over approximately 17 months.

To detect an improvement in OS using a log-rank test at a two-sided significance level of 0.05, approximately 319 deaths will be required to achieve 80% overall power assuming a target HR of 0.73 (median OS improvement over active surveillance of 22 months based on the assumption that the median OS for active surveillance is 60 months).

Sample Size for the China Subpopulation

The Sponsor is targeting a total enrollment of a minimum of approximately 250 patients from mainland China. The sample size of the China subpopulation was determined by characterizing the efficacy and safety profile of atezolizumab+bevacizumab. After approximately 662 patients have been randomized into the global portion of the study, in the event that fewer than 250 patients from mainland China are enrolled, additional patients in China may be subsequently randomized into the two study arms in a 1:1 ratio in an extended China enrollment phase to ensure a total of approximately 250 patients from mainland China for the China subpopulation.

4. ANALYSIS SETS

4.1 INTENT-TO-TREAT POPULATION

The ITT population is defined as all randomized patients, whether or not the patient has received the assigned study treatment.

Under the ITT principle, patients will be grouped according to the treatment assigned at randomization, regardless of whether they receive any assigned study drug, cause of going off-protocol treatment, crossover from Arm B to treatment with atezolizumab+bevacizumab, etc.

4.2 PD-L1-HIGH SUBGROUP

The PD-L1-high subgroup is defined as PD-L1 expression on tumor and immune cells that comprises at least 1% of the tumor area at the time of randomization. Submission of surgical tumor tissues at baseline will be mandated and the expression of PD-L1 will be examined by the central laboratory using a PD-L1 (SP263) immunohistochemistry assay and tumor area positive (TAP) scoring. OS, IRF- and investigator-assessed RFS outcome measures will also be performed for patients in the PD-L1-High subgroup within ITT population.

4.3 SAFETY POPULATION

The safety analysis population will consist of all patients randomized to Arm A who received at least one full or partial dose of study treatment (atezolizumab and/or bevacizumab) and all patients randomized to Arm B (active surveillance) who underwent at least one safety assessment. For patients crossing over from active surveillance to treatment, the safety analysis population will consist of all patients who underwent at least one safety assessment prior to crossover and all patients who received at least one full or partial dose of study treatment (atezolizumab and/or bevacizumab) post-crossover.

4.4 PHARMACOKINETIC-EVALUABLE POPULATION

The pharmacokinetic (PK) analysis population will consist of all patients whom had received any amount of study drug and had at least one measurable serum concentration result post-dose available at the clinical data cutoff for the clinical study report.

4.5 IMMUNOGENICITY ANALYSIS POPULATION

The immunogenicity analysis population will consist of all patients whom had received any amount of study drug and had at least one measurable anti-drug antibody (ADA) result post-dose available at the clinical data cutoff for the clinical study report.

5. STATISTICAL ANALYSES

5.1 GENERAL CONSIDERATION

All efficacy analyses will be performed in the ITT population, unless otherwise specified. Patients will be analyzed according to the treatment assigned at randomization by IxRS.

All IRF-assessed RFS analyses will use the recurrence assessments per the charter, unless otherwise specified.

All safety analyses will be performed in the safety-evaluable population, unless otherwise specified. Patients will be analyzed according to the treatment they actually received. Safety analyses will be conducted separately for the study periods prior to crossover and post-crossover.

5.2 PATIENT DISPOSITION

Enrollment and reasons for discontinuation from the study will be summarized by treatment arm for the ITT population. Study treatment administration and reasons for discontinuation from study treatment will be summarized for the safety-evaluable population.

5.3 PRIMARY ENDPOINT ANALYSIS

5.3.1 Definition of Primary Endpoint

The primary efficacy endpoint is IRF-assessed RFS, defined as the time from randomization to the first documented occurrence of intrahepatic or extrahepatic HCC, or death from any cause (whichever occurs first).

Patients who have not experienced disease recurrence or death at the time of analysis will be censored at the date of the last assessment for HCC occurrence. Patients with no post-baseline radiographic assessment will be censored at the date of randomization. For the purposes of analysis of RFS in the ITT population, in the rare event that the IRF identifies baseline disease, this patient will be assessed as having a recurrence event at the time of randomization.

Analysis of RFS will be performed on basis of the ITT population using the ITT principle: with patients grouped according to the study arm assigned at randomization, regardless of whether they receive any assigned study drug, cause of going off protocol treatment, crossover from Arm B to treatment with atezolizumab plus bevacizumab etc.

5.3.2 Main Analytical Approach for Primary Endpoint

Statistical hypothesis which will be tested for IRF-assessed RFS is given in Section 2. IRF-assessed RFS will be tested with the overall type I error controlled at a two-sided significance level of 0.05. The two-sided stratified log-rank test, stratified by the same stratification factors specified for randomization, will be used as the primary analysis to compare RFS between the two arms. The adjuvant TACE component will be removed from the randomization stratification factor High Risk Features/Curative Procedure in the stratified analyses such that this stratification factor becomes:

- Ablation
- Resection with 1 high risk feature
- Resection with 2 or more high risk features

Due to the potential risk of over-stratification (Akazawa et al. 1997), if at least one stratum (i.e., a combination of stratification factor levels across all stratification factors) has fewer than 5 RFS events across treatment arms, the stratification factor which contains the level with the smallest number of patients will be removed from the stratified analyses. The removal of the stratification factors will continue until there is no stratum with fewer than 5 RFS events. The unstratified log-rank test will be also used to check the robustness of the results of the stratified log-rank test.

The Kaplan-Meier method will be used to estimate the median RFS for each study arm. The Brookmeyer-Crowley method will be used to construct the 95% CI for the median RFS for each study arm. Stratified Cox proportional-hazards models will be used to estimate the HR and its 95% confidence interval (CI). The unstratified HR will also be estimated.

A group sequential design will be implemented for testing the RFS primary endpoint to account for the conduct of one interim analysis, which is to be conducted when approximately 236 IRF-assessed RFS events have occurred and approximately 26 months after the first patient is randomized. The overall type I error rate of 0.05 for testing of the RFS primary efficacy endpoint will be controlled through use of an α-spending function that utilizes the Lan-DeMets method approximating the O'Brien-Fleming boundaries (O'Brien and Fleming 1979).

The final analysis of RFS will be conducted when approximately 323 IRF-assessed RFS events have occurred, which is expected to take place approximately 39 months after the first patient is randomized.

Sensitivity analyses to assess the robustness of the analysis of RFS are described in Section 5.3.3.

To assess the homogeneity of the treatment effect with respect to the primary efficacy endpoint of RFS across important subgroups, Forest plots (including the estimated HRs) will be provided.

5.3.3 Sensitivity Analyses for Primary Endpoint

5.3.3.1 Impact of Missing Scheduled Tumor Assessments on Primary Endpoint IRF-Assessed RFS

The impact of missing scheduled tumor assessments on the primary endpoint of IRF-assessed RFS will be assessed depending on the number of patients who missed consecutive assessments scheduled immediately prior to the date of disease recurrence or death. If >5% of patients missed two or more consecutive recurrence assessments scheduled immediately prior to the first documented occurrence of intrahepatic or extrahepatic HCC as determined by an IRF-assessment or death in any treatment arm, the following two sensitivity analyses may be performed:

- Patients who missed two or more consecutive recurrence assessments scheduled immediately prior to the first documented occurrence of intrahepatic or extrahepatic HCC as determined by an IRF-assessment or death will be censored at the last tumor assessment prior to the missed visits.
- Patients who missed two or more consecutive recurrence assessments scheduled immediately prior to the first documented occurrence of intrahepatic or extrahepatic HCC as determined by an IRF-assessment or death will be counted as having recurred on the date of the first of these missing assessments.

5.3.3.2 Sensitivity Analyses on IRF-Assessed RFS for Patients in Any Treatment Arm Receiving Non-Protocol Therapy, Resection, or Ablation

The impact of non-protocol therapy or procedures on IRF-Assessed RFS may be assessed depending on the number of patients who receive such therapy before an IRF-assessed RFS event. If more than 5% of patients received non-protocol therapy or procedures before an IRF-assessed RFS event in any treatment arm, data for these patients will be censored at the last disease assessment date before they received such therapy in IRF-assessed RFS analyses that may be performed for the comparisons between the study arms.

5.3.3.3 Sensitivity Analyses on IRF-Assessed RFS for Missing Data

The missing data in this context are patients who are prematurely censored due to withdrawal of consent, lost to follow-up, or otherwise incomplete follow-up with the missing-at-random assumption. The impact of withdrawal of consent, lost to follow-up, or otherwise incomplete follow-up on IRF-assessed RFS may be assessed depending on the number of patients who have withdrawal of consent, lost to follow-up, or otherwise incomplete follow-up prior to an IRF-Assessed RFS event. If more than 5% of

patients have withdrawal of consent, lost to follow-up, or otherwise incomplete follow-up prior to an IRF-assessed RFS event in any treatment arm, data for these patients may be excluded in the IRF-assessed RFS that may be performed for the comparisons between the study arms.

5.3.3.4 Sensitivity Analyses on IRF-Assessed RFS for Patients with Baseline Disease

The impact of baseline disease on IRF–assessed RFS may be assessed depending on the number of patients who are identified with baseline disease according to the IRF. If more than 3% of patients among ITT have IRF baseline disease, data for these patients will be censored at the time of randomization for the comparisons between the study arms.

5.3.4 Supplementary Analyses for Primary Endpoint

5.3.4.1 Subgroup Analyses for Primary Endpoint

To assess the homogeneity of the treatment effect with respect to the primary efficacy endpoint of IRF-assessed RFS across important subgroups, forest plots (including the estimated HRs) will be provided, including, but not limited to, the following variables: age, sex, race, geographic region, high risk features/curative procedure defined in the study protocol, number of high risk features, Eastern Cooperative Oncology Group (ECOG) performance status, HCC etiology, Barcelona Clinic Liver Cancer (BCLC) staging at the time of study entry, TACE, and baseline PD-L1 expression in tumor tissue for patients with baseline tumor samples. Unstratified analysis results will be presented for subgroup analyses due to the potentially limited number of patients in each subgroup.

5.4 SECONDARY ENDPOINTS ANALYSES

The secondary efficacy endpoints are OS, investigator-assessed RFS, investigator-assessed and IRF-assessed time to recurrence (TTR), investigator- and IRF-assessed RFS rate at 24 months and 36 months, OS rate at 24 months and 36 months, investigator-assessed time to extrahepatic spread (EHS) or macrovascular invasion, and investigator-assessed and IRF-assessed RFS among patients in the PD-L1-high subgroup. OS will be a key secondary endpoint of this study.

5.4.1 Key Secondary Endpoint: Overall Survival

OS will be defined as the time from randomization to death from any cause. Patients who are alive at the time of the analysis will be censored at the last date the patient was known to be alive. Patients with no post-baseline information will be censored at the date of randomization. The methodology used for RFS will be applied to OS in the ITT population.

To detect an improvement in OS using a log-rank test at a two-sided significance level of 0.05, approximately 319 deaths will be required to achieve 80% overall power assuming a target HR of 0.73 (median OS improvement over active surveillance of 22 months).

The final OS analysis will be conducted when approximately 319 deaths have occurred and is expected to take place approximately 91 months after the first patient is randomized. Formal statistical treatment comparison for OS will only be performed after IRF-assessed RFS results have reached statistical significance. See Section 5.8 for additional details.

5.4.1.1 Sensitivity Analyses on OS for Patients in Control Arm Receiving Atezolizumab, Bevacizumab, Immunotherapy, Resection, or Ablation

The impact on OS by patients in control arm subsequently receiving atezolizumab plus bevacizumab in the crossover period, or another immunotherapy considered similar to atezolizumab or bevacizumab in its mechanism of action may be assessed, depending on the number of such patients. If more than 5% of patients in control arm have received atezolizumab plus bevacizumab or a similar immunotherapy, the following analyses may be performed to compare treatment arms at each of the interim analyses and final analysis, as applicable:

- OS in the control arm will be discounted according to a range of possible effects on OS after having received atezolizumab plus bevacizumab or another immunotherapy (e.g., 10%, 20%, 30%, etc.)
- The rank preserving structural failure time model may be applied to evaluate the impact of patient crossover on OS benefit

5.4.2 Supportive Secondary Endpoint(s)

5.4.2.1 Investigator-Assessed Recurrence Free Survival

The analysis methods for investigator-assessed RFS are analogous to those described for IRF-assessed RFS.

5.4.2.2 Time to Recurrence

TTR is defined as the time from randomization to first documented occurrence of intrahepatic or extrahepatic HCC. Patients who have not experienced disease recurrence at the time of the analysis will be censored at the date of the last assessment for HCC occurrence. Patients with no post-baseline information will be censored at the date of randomization. The analysis of TTR will be conducted separately based on investigator-assessed and IRF-assessed data and will follow the methods described for the primary endpoint.

5.4.2.3 IRF-Assessed and Investigator-Assessed RFS at 24-Month and 36-Month Landmark Analysis

IRF- and investigator-assessed RFS rates at 24 months and 36 months will be estimated for each study arm through use of the Kaplan-Meier method, with 95% CIs calculated through use of Greenwood's formula.

5.4.2.4 Overall Survival 24-Month and 36-Month Landmark Analysis

OS rates at 24 months and 36 months will be estimated for each study arm through use of the Kaplan-Meier method, with 95% CIs calculated through use of Greenwood's formula.

5.4.2.5 Time to Extrahepatic Spread or Macrovascular Invasion

Time to EHS or macrovascular invasion is defined as the time from randomization to the first appearance of EHS or macrovascular invasion. The analysis will be conducted for the study period prior to crossover. Events that occurred in the study period post-crossover will not be included. Patients who have not experienced either EHS or macrovascular invasion at the time of the analysis or before crossover will be censored at the date of the last assessment for HCC occurrence before the analysis data cutoff date and before crossover when applicable. Patients with no post-baseline information will be censored at the date of randomization. The analysis of time to EHS or macrovascular invasion will be conducted based on investigator-assessed data and will follow the methods described for the primary endpoint.

5.4.2.6 Recurrence-Free Survival among Patients with PD-L1-High Tumors

RFS among patients in the PD-L1-high subgroup is defined in an analogous manner to the primary endpoint and will be analyzed through use of same methods described for the primary endpoint. The analysis of RFS in the PD-L1-high subgroup will be conducted separately based on investigator-assessed RFS data and IRF-assessed data. PD-L1-high is defined by TAP ≥1% using PD-L1 immunohistochemistry (IHC) assay (SP263).

5.5 EXPLORATORY ENDPOINTS ANALYSIS

Exploratory Overall Survival Analyses

OS among patients in the PD-L1-high subgroup is defined in an analogous manner to OS in the ITT population and will be analyzed through use of the same methods described for OS in the ITT population. PD-L1-high is defined by TAP ≥1% using PD-L1 IHC assay (SP263).

Exploratory Patient-Reported Outcome Analyses

Visit mean summary and change from baseline will be performed for physical functioning, emotional functioning, role functioning, social functioning, and Global Health Status (GHS)/Quality of Life (QoL) scores as assessed by the IL42-European Organization for Research and Treatment of Cancer Quality of Life-Core 30 Questionnaire (Reduced) (IL42–EORTC QLQ-C30 [Reduced]). Summary statistics (number of patients, mean, mean change, standard deviation, median, minimum, maximum, 95% CI) of linearly transformed scores (per the EORTC scoring manual, Fayers et al. 2001) will be calculated at all assessment timepoints for each study arm. Additional descriptive visit mean summaries and changes from baseline for GHS/QoL

will be compared between the arms at 24 months and 36 months. Published minimally important differences (e.g., 10 points) will be considered to identify meaningful changes from baseline for each scale within each treatment group (Osoba et al. 1998). Alternative meaningful change thresholds will also be considered (Cocks et al. 2011; Cocks et al. 2012).

Completion rates and reasons for missing data will be summarized for the IL42-EORTC QLQ-C30 (Reduced) questionnaire at each assessment timepoint for both arms.

Exploratory Efficacy Analyses in Arm B patients who Cross Over to Treatment with Atezolizumab+Bevacizumab

The following exploratory efficacy endpoints will be analyzed in patients in Arm B who cross over to treatment with atezolizumab+bevacizumab after documented HCC recurrence:

- For patients with NED at the time of crossover:
  - RFS after first HCC recurrence, defined as the time from first exposure to any dose of crossover treatment to the second documented HCC recurrence or death from any cause (whichever occurs first), as determined by the investigator. HCC recurrence is defined as occurrence of intrahepatic or extrahepatic HCC.
- For patients with measurable disease at the time of crossover:
  - Progression-free survival (PFS) after first HCC recurrence, defined as the time from first exposure to any dose of crossover treatment to the first documented occurrence of disease progression beyond the initial unresectable disease recurrence as determined by the investigator according to Response Evaluation Criteria in Solid Tumors (RECIST) v1.1, or death from any cause (whichever occurs first)
  - Objective response rate (ORR), defined as the proportion of patients with a complete response (CR) or partial response (PR) as determined by the investigator according to RECIST v1.1

The 95% CI for ORR will be calculated using the Clopper-Pearson method (Clopper and Pearson 1934).

5.6 SAFETY ANALYSES

Unless specified otherwise, the safety analyses described below will be conducted for the safety-evaluable population (see Section 4.3), with all patients randomized to Arm A who received at least one full or partial dose of study treatment (atezolizumab and/or bevacizumab) and all patients randomized to Arm B who underwent at least one safety assessment. Safety analyses for crossover patients will be conducted separately for the study periods prior to crossover and post-crossover.

5.6.1 Extent of Exposure

Drug exposure will be summarized by descriptive statistics to include treatment duration, number of doses, and dose intensity.

5.6.2 Adverse Events

Verbatim adverse event terms will be mapped to Medical Dictionary for Regulatory Activities (MedDRA) thesaurus terms, and adverse event severity will be graded according to National Cancer Institute Common Terminology Criteria for Adverse Events (NCI CTCAE) v5.0.

The following events occurring during or after initiation of study treatment or surveillance on Day 1 of Cycle 1 will be summarized by study arm and NCI CTCAE grade:

- All adverse events
- All adverse events leading to death
- Serious adverse events
- Grade ≥ 3 adverse events
- Adverse event of special interest
- Adverse events leading to study drug discontinuation or interruption

Multiple occurrences of the same event will be counted once at the maximum severity. Laboratory data with values outside the normal ranges will be identified. In addition, selected laboratory data will be summarized by study arm and grade.

Descriptive statistics will be used to summarize changes in vital signs by study arm. Deaths and causes of death reported during the study treatment period and those reported during the follow-up period after treatment or surveillance completion/discontinuation will be summarized by arm.

Additional analyses may be performed as indicated.

5.6.3 Laboratory Data

Laboratory data will be classified according to NCI CTCAE v5.0. Laboratory data with values outside the normal ranges will be identified. In addition, selected laboratory data will be summarized by study arm and grade.

5.6.4 Vital Signs

Descriptive statistics will be used to summarize changes in vital signs by study arm.

5.6.5 Electrocardiogram (ECG)

Post-baseline abnormal electrocardiogram (ECG) (12-lead) records over time will be listed by study arm.

5.7 OTHER ANALYSES

5.7.1 Summaries of Conduct of Study

Study enrollment, study treatment administration, reasons for discontinuation from the study treatment, and reasons for study discontinuation will be summarized by study arm for all randomized patients (the ITT population). Major protocol deviations, including major deviations with regard to the inclusion and exclusion criteria, will be summarized by study arm for the ITT population.

5.7.2 Summaries of Treatment Group Comparability

Demographic characteristics such as age, sex, race/ethnicity, and baseline disease characteristics (e.g., ECOG Performance Status) will be summarized by study arm for the ITT population. Descriptive statistics (mean, median, standard deviation, and range) will be presented for continuous data, and frequencies and percentages will be presented for categorical data.

Baseline measurements are the last available data obtained prior to initiation of treatment or surveillance on Day 1 of Cycle 1.

5.7.3 Pharmacokinetic Analyses

PK analyses will be performed in the PK-evaluable population (see Section 4.4).

Serum concentrations of atezolizumab will be reported as individual values and summarized (geometric mean and geometric mean coefficient of variation) by cycle, when appropriate and as data allow. Individual and median serum atezolizumab concentrations will be plotted for PK-evaluable patients by day.

Atezolizumab concentration data may be pooled with data from other studies using an established population PK model to derive PK parameters such as clearance, volume of distribution, and area under the curve, as warranted by the data. Potential correlations of relevant PK parameters with dose, safety, efficacy, or biomarker outcomes may be explored.

5.7.4 Immunogenicity Analyses

The immunogenicity analyses will be performed in the immunogenicity analysis population (see Section 4.5).

The numbers and proportions of ADA-positive patients and ADA-negative patients at baseline (baseline prevalence) and after drug administration (post-baseline incidence) will be summarized for ADA-evaluable patients. When determining post-baseline incidence, patients are considered to be post-baseline ADA-positive if they are ADA-negative or have missing data at baseline but develop an ADA response following study drug exposure (treatment-induced ADA response), or if they are ADA-positive at baseline and the titer of one or more post-baseline samples is at least 0.60 titer unit

greater than the titer of the baseline sample (treatment-enhanced ADA response). Patients are considered to be post-baseline ADA-negative if they are ADA-negative or have missing data at baseline and all post-baseline samples are negative, or if they are ADA-positive at baseline but do not have any post-baseline samples with a titer that is at least 0.60 titer unit greater than the titer of the baseline sample (treatment unaffected).

The relationship between ADA status and safety, efficacy, PK, and biomarker endpoints may be analyzed and reported.

5.7.5 Biomarker Analyses

Although no formal statistical analysis of exploratory biomarkers will be performed, biomarker data may be analyzed in the context of this study and in aggregate with data from other studies.

5.7.6 Health Utility Analyses

Health utility data from the EuroQol 5-Dimension, 5-Level Questionnaire (EQ-5D-5L) will be evaluated in pharmacoeconomic models. The results from the health economic data analyses will be reported separately from the clinical study report.

5.7.7 Analyses of Subgroups of Interest

The PD-L1-high subgroup population is defined in Section 4.2. RFS among patients in the PD-L1-high subgroup is defined in an analogous manner to the primary endpoint and will be analyzed through use of same methods described for the primary endpoint. The analysis of RFS in the PD-L1-high subgroup will be conducted separately based on investigator-assessed RFS data and IRF-assessed data (see Section 5.4.2.6).

5.7.8 Analyses of China Subpopulation

A separate analysis will be performed for the China subpopulation, where data from all participants enrolled in mainland China, Taiwan and Hong Kong (during both the global enrollment phase and the extended China enrollment phase) will be combined and summarized. Data from the China extension cohort will not be included in the primary analysis of the main study.

All analyses described in this section will include all data from the China subpopulation collected up to the clinical cutoff date for the China subpopulation analysis as defined in Section 3. Data for the China subpopulation will be analyzed using the same statistical methods as described in Section 5.1-5.6 when data allow.

The primary efficacy objective of the China subpopulation analysis is to assess efficacy, as measured by the primary endpoint of IRF-assessed RFS of atezolizumab+bevacizumab compared with active surveillance in the Chinese patients. The China subpopulation is not powered to demonstrate statistical significance in terms of efficacy, and no formal hypothesis testing will be performed.

The China subpopulation analyses will be conducted when sufficient RFS events have occurred to demonstrate approximately 80% probability of maintaining 50% of RFS risk reduction compared with that estimated from the global population.

The analysis methods for China subpopulation will be the same as for the global population unless elsewhere noted. The results of the China subpopulation analyses will be summarized in a separate report from the clinical study report for the global population.

5.8 INTERIM ANALYSES

5.8.1 Planned Interim Analyses

One interim analysis of IRF-assessed RFS will be performed. The interim analysis will be performed when approximately 236 RFS events have occurred, which is expected to take place approximately 26 months after the first patient is randomized. For the interim analysis, the MDD for RFS is a HR of 0.734 (median RFS improvement over active surveillance of 7.2 months).

For the interim analysis, efficacy stopping boundaries will be determined through use of the Lan-DeMets method to approximate the O'Brien-Fleming boundaries. Analysis timing and stopping boundaries for the interim and final RFS analyses are summarized in Table 3.

The planned interim analysis of RFS will be conducted by an iDCC and reviewed by the iDMC. Interactions between the iDMC and Sponsor will be carried out as specified in the iDMC Charter.

Table 3 Analysis Timing and Stopping Boundaries for Recurrence-Free Survival Analyses

Analysis Timing[a]	Planned Information Fraction	Required No. of Events (Estimated)	Estimated Analysis Timing	Stopping Boundary (Two-Sided p-Value)[b]
RFS interim analysis	73%	236	26 months after first patient is randomized	MDD HR ≤0.734 (p-value ≤0.017)
RFS final analysis	100%	323	39 months after first patient is randomized	MDD HR ≤0.8 (p-value ≤0.045)

HR=hazard ratio; ITT=intent-to-treat; MDD=minimally detectable difference (based on an exponential distribution); RFS=recurrence-free survival.

[a] Analysis timing is estimated on the basis of protocol assumptions. Actual timing depends on the exact time that the required events have accrued.

[b] The actual stopping boundaries will be calculated at the time of the interim and final RFS analysis on the basis of the observed information fraction, that is, the actual number of RFS events observed at the time of analysis over the total planned target number of RFS events in the ITT population.

A group sequential design will be implemented for testing OS to account for the conduct of two interim analyses. The boundary for statistical significance at each interim analysis and the final analysis will be determined based on the Lan-DeMets implementation of the O'Brien-Fleming function (Lan and DeMets, 1983) to maintain the overall type 1 error rate at 0.05 level. The O'Brien-Fleming boundary for statistical significance is provided in Table 4. Formal statistical treatment comparison for OS will only be performed after IRF-assessed RSF results have reached statistical significance at either the interim or final RFS analyses.

If RFS is statistically significant at the interim analysis, then three analyses of the key secondary endpoint of OS are planned (see Table 4). The two interim analyses of OS will be conducted at the time of the planned interim and final analyses of the primary efficacy endpoint of IRF-assessed RFS.

The first interim analysis of OS will be performed at the time of the interim RFS analysis, which will be conducted when approximately 236 RFS events have occurred. On the basis of the projected median OS for each treatment arm, the projected number of deaths observed at the RFS interim analysis is approximately 107 deaths (16% of 662 patients), which corresponds to approximately 33.5% of the number of deaths required for the final analysis of OS. The observed HR of OS that is projected to result in a statistically significant difference between treatment arms is less than or equal to 0.489.

The second interim analysis of OS will be performed at the time of the final RFS analysis, which will be conducted when approximately 323 RFS events have occurred. On the basis of the projected median OS for each treatment arm, the projected number of deaths observed at the RFS final analysis is approximately 164 deaths (25% of 662 patients), which corresponds to approximately 51.4% of the number of deaths required for the final analysis of OS. The observed HR of OS that is projected to result in a statistically significant difference between treatment arms is less than or equal to 0.634.

The final analysis of OS will be performed when 319 deaths (48% of 662 patients in the ITT population, expected analysis timing approximately 91 months after the first patient is randomized) have occurred. The observed HR of OS that is projected to result in a statistically significant difference between treatment arms is less than or equal to 0.802.

Table 4 Projected Interim and Final OS Analysis Characteristics When Statistical Significance is Reached at the Interim Analysis for RFS

Analysis	No. of Events	% Information	Event to Patient Ratio	Projected Cutoff Date [a]	Projected MDD [b]	Projected Boundary (p-value) [c]
First OS interim (Performed at time of RFS interim analysis)	107	33.5%	16%	Month 26	0.489	p ≤0.0002
Second OS interim (Performed at time of RFS final analysis)	164	51.4%	25%	Month 39	0.634	p ≤0.0035
Final OS (Event driven)	319	100%	48%	Month 91	0.802	p ≤0.0488

MDD = minimally detectable difference; OS = overall survival; RFS = recurrence free survival

Note: Assumes 5% dropout rate over 12 months for OS analyses.

[a] Study month at which required number of events are projected to occur, where Study Month 1 is the month the first patient is enrolled. Analysis results will be available after data cleaning.

[b] The largest observed hazard ratio that is projected to be statistically significant.

[c] The projected boundary for statistical significance for the number of events shown (actual boundary to be calculated at time of analysis based on actual number of events).

If RFS is not statistically significant at the interim analysis, the study will continue to the RFS final analysis. If RFS is statistically significant at the final analysis, then one interim analysis of the secondary endpoint of OS is planned at the RFS final analysis (see Table 5).

The first interim analysis of OS will be performed at the time of the final RFS analysis, which will be conducted when approximately 323 RFS events (expected analysis timing approximately 39 months after the first patient is randomized) have occurred. On the basis of the projected median OS for each treatment arm, the projected number of deaths observed at the RFS final analysis is approximately 164 deaths (25% of 662 patients), which corresponds to approximately 51.4% of the number of deaths required for the final analysis of OS. The observed HR of OS that is projected to result in a statistically significant difference between treatment arms is less than or equal to 0.634.

The final analysis of OS will be performed when 319 deaths (48% of 662 patients in the ITT population, expected analysis timing approximately 91 months after the first patient is randomized) have occurred. The observed HR of OS that is projected to result in a statistically significant difference between treatment arms is less than or equal to 0.802.

Table 5 Projected Interim and Final OS Analysis Characteristics When Statistical Significance is Reached at the Final Analysis for RFS

Analysis	No. of Events	% Information	Event to Patient Ratio	Projected Cutoff Date [a]	Projected MDD [b]	Projected Boundary (p-value) [c]
First OS interim (performed at time of RFS final analysis)	164	51.4%	25%	Month 39	0.634	p ≤0.0035
Final OS (Event driven)	319	100%	48%	Month 91	0.802	p ≤0.0488

MDD = minimally detectable difference; OS = overall survival; RFS = recurrence free survival

Note: Assumes 5% dropout rate over 12 months for OS analyses.

[a] Study month at which required number of events are projected to occur, where Study Month 1 is the month the first patient is enrolled. Analysis results will be available after data cleaning.

[b] The largest observed hazard ratio that is projected to be statistically significant.

[c] The projected boundary for statistical significance for the number of events shown (actual boundary to be calculated at time of analysis based on actual number of events).

If RFS is not statistically significant in either interim or final analysis, the secondary endpoint of the OS will not be formally tested and will be descriptive only.

The boundaries for statistical significance at the planned IRF-assessed RFS and OS interim and final analyses will be determined based on the actual number of events observed.

6. SUPPORTING DOCUMENTATION

This section is not applicable, since there is no additional supporting document.

Appendix 1 Changes to Protocol-Planned Analyses

This section is not applicable since there are no changes to the protocol-planned analyses.

7. REFERENCES

Sponsor Reports

WO41535: A phase III, multicenter, randomized, open-label study of atezolizumab (anti-PD-L1 antibody) plus bevacizumab versus active surveillance as adjuvant therapy in patients with hepatocellular carcinoma at high risk of recurrence after surgical resection or ablation. Clinical study protocol, Version 4, dated 30 November 2021

WO41535: A phase III, multicenter, randomized, open-label study of atezolizumab (anti-PD-L1 antibody) plus bevacizumab versus active surveillance as adjuvant therapy in patients with hepatocellular carcinoma at high risk of recurrence after surgical resection or ablation. Independent Read Charter, Version 2, dated 21 December 2020.

WO41535: A phase III, multicenter, randomized, open-label study of atezolizumab (anti-PD-L1 antibody) plus bevacizumab versus active surveillance as adjuvant therapy in patients with hepatocellular carcinoma at high risk of recurrence after surgical resection or ablation. Charter for the independent data monitoring committee, Version 2, dated 5 June 2021.

Literature References

Akazawa K., Nakamura T, and Palesch Y. Power of logrank test and cox regression model in clinical trials with heterogeneous samples. Statist Med. 1997;16:583-597.

Clopper CJ, Pearson ES. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika. 1934 Dec 1;26(4):404-13.

Cocks, K., King, M. T., Velikova, G., Martyn St-James, M., Fayers, P. M., & Brown, J. M. Evidence-based guidelines for determination of sample size and interpretation of the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. J Clin Oncol. 2011;29(1), 89-96.

Cocks, K., King, M. T., Velikova, G., de Castro Jr, G., St-James, M. M., Fayers, P. M., & Brown, J. M. Evidence-based guidelines for interpreting change scores for the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. Eur J Cancer. 2012;48(11), 1713-1721.

Fayers, P. M., Aaronson, N. K., Bjordal, K. et al on behalf of the European Organisation for Research and Treatment of Cancer Quality of Life Study Group. (2001). The EORTC QLQ-C30 scoring manual (3rd ed.). Brussels: EORTC.

Gordon Lan KK, DeMets DL. Discrete sequential boundaries for clinical trials. Biometrika. 1983;Dec 1;70(3):659-63..

O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics. 1979;Sep 1:549-56..

Osoba D, Rodrigues G, Myles J et al. Interpreting the significance of changes in health-related quality-of-life scores. Journal of clinical oncology. 1998;Jan 1;16(1):139-44.

Signature Page for Statistical Analysis Plan - WO41535 - Published
System identifier: RIM-CLIN-453965

Approval Task	[redacted] Company Signatory 07-Oct-2022 22:47:17 GMT+0000