

**Bevacizumab (Avastin<sup>®</sup>, RO 487-6646)**

**ML 21380**

**Open-label, efficacy and safety study of bevacizumab (Avastin<sup>®</sup>) in combination with XELOX (Oxaliplatin plus Xeloda<sup>®</sup>) for the first-line treatment of patients with locally advanced or metastatic cancer of the colon or rectum – “OBELIX”**

**Phase IIIb**

**Document Version History**

Version Date	Version	Author	Signature	Change Description	Reason/Comment
06/09/2012	1	Elena Raimondi	<i>Elena Raimondi</i>	Initial release.	Not applicable.

### TABLE OF CONTENTS:

<b>LIST OF ABBREVIATIONS .....</b>	<b>2</b>
<b>1. INTRODUCTION .....</b>	<b>3</b>
<b>2. STUDY OBJECTIVES .....</b>	<b>3</b>
<b>3. INVESTIGATIONAL PLAN .....</b>	<b>3</b>
3.1. OVERALL STUDY DESIGN AND PLAN DESCRIPTION .....	3
3.2. SELECTION OF STUDY POPULATION .....	4
3.2.1. INCLUSION CRITERIA .....	4
3.2.2. EXCLUSION CRITERIA .....	4
3.3. TREATMENTS .....	5
3.3.1. METHOD OF ASSIGNING PATIENTS TO TREATMENT GROUP .....	6
3.4. EFFICACY AND SAFETY VARIABLES .....	6
3.4.1. EFFICACY AND SAFETY MEASUREMENTS ASSESSED AND FLOW CHART .....	6
3.4.2. PRIMARY EFFICACY VARIABLES .....	9
3.4.3. SECONDARY EFFICACY VARIABLES .....	9
3.4.4. SAFETY VARIABLES .....	11
3.5. DATA MANAGEMENT .....	12
3.5.1. DATA COLLECTION .....	12
3.5.2. DATABASE MANAGEMENT AND QUALITY CONTROL .....	12
3.6. STATISTICAL METHOD PLANNED IN THE PROTOCOL AND SAMPLE SIZE DETERMINATION .....	13
3.6.1. STATISTICAL AND ANALYTICAL PLAN .....	13
3.6.2. SAMPLE SIZE DETERMINATION .....	20
3.7. CHANGES IN THE CONDUCT OF THE STUDY OR PLANNED ANALYSES .....	20
3.8. INTERIM ANALYSIS .....	20
<b>4. STUDY PATIENTS .....</b>	<b>21</b>
4.1. DISPOSITION OF PATIENTS .....	21
4.2. PROTOCOL DEVIATION .....	21
<b>5. EFFICACY EVALUATION .....</b>	<b>23</b>
5.1. DATASET ANALYSED .....	23
5.2. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS .....	24
5.3. MEASUREMENT OF TREATMENT COMPLIANCE .....	30
5.4. EFFICACY RESULTS .....	31
5.4.1. ANALYSIS OF EFFICACY .....	31
5.4.2. STATISTICAL/ANALYTICAL ISSUES .....	46
5.4.3. EFFICACY CONCLUSION .....	47
<b>6. SAFETY EVALUATION .....</b>	<b>49</b>
6.1. EXTENT OF EXPOSURE .....	49
6.2. ADVERSE EVENT .....	50
6.3. DEATHS, OTHER SERIOUS ADVERSE EVENTS AND OTHER SIGNIFICANT ADVERSE EVENTS .....	55
6.4. CLINICAL LABORATORY EVALUATION .....	57
6.5. VITAL SIGNS, PHYSICAL FINDINGS AND OTHER OBSERVATIONS RELATED TO SAFETY .....	59
6.6. SAFETY CONCLUSION .....	59

### LIST OF ABBREVIATIONS

AE	Adverse Event
ANCOVA	ANalysis of COVariance
CR	Complete Response (RECIST definition)
CRF	Case Report Form
CT	Chemotherapy
ECG	ElectroCardioGram
ESF	Eligibility Screening Form
ITT	Intent-To-Treat
LOCF	Last Observation Carried Forward
NA	Not Assessable (RECIST definition)
NCI CTC	National Cancer Institute Common Terminology Criteria
NON-CR/NON-PD	Neither Complete Response/ Neither Progressive Disease (RECIST definition)
ORR	Overall Response Rate
PD	Progressive Disease (RECIST definition)
PFS	Progression Free Survival
PP	Per Protocol
PR	Partial Response (RECIST definition)
PT	Preferred Term
RECIST	Response Evaluation Criteria in Solid Tumours
RR	Relative Risk
SAE	Serious Adverse Event
SD	Standard Deviation
SD	Stable Disease (RECIST definition)
SOC	System Organ Class
TTF	Time to Treatment Failure
VAS	Visual Analogue Scale



## 1. INTRODUCTION

(See Chapter 1 on ML21380 amended protocol document version 3– 14 October 2010)

## 2. STUDY OBJECTIVES

According to protocol amendment n° 3 approved on 14<sup>th</sup> October 2010, this phase IIIb trial has been designed to test the efficacy and safety of Bevacizumab administered in combination with XELOX regimen until progressive disease. The study had also an explorative purpose to investigate if the tumoral K-Ras and B-Raf markers are predictor for Bevacizumab activity.

### Primary Objective:

To confirm the efficacy of Bevacizumab in combination with Oxaliplatin and Capecitabine (XELOX) based regimen, based on Progression Free Survival (PFS).

### Secondary objectives:

- To evaluate the safety profile of Bevacizumab in combination with Oxaliplatin and Capecitabine (XELOX) based regimen.
- To determine the Overall Response Rate (ORR), Time To Response (TTR), duration of response, overall survival rate, percentage of R0 resectability of metastatic lesions and quality of life of patients treated with Bevacizumab in combination with Oxaliplatin and Capecitabine (XELOX).
- To evaluate the correlation between the K-Ras mutation status and the B-Raf mutation status with the overall response rate, time to response and Progression Free Survival.
- To investigate the role of K-Ras and B-Raf mutations as predictor of Bevacizumab activity.

## 3. INVESTIGATIONAL PLAN

### 3.1. OVERALL STUDY DESIGN AND PLAN DESCRIPTION

Single arm, open label, multicenter phase IIIb study.

All subjects had to sign and date the most current IRB/IEC-approved written informed consent before any study specific assessments or procedures were performed.

A screening examination should have been performed between 1 and 14 days before the start of the study, except for imaging procedures (CT, MRI, X-Ray) which might have been performed within 30 days before starting treatment. An Eligibility Screening Form [ESF] (not collected on CRF) documenting the investigator's assessment of each screened subject with regard to the protocol's inclusion and exclusion criteria had to be completed by the investigator.

Once a patient has fulfilled the entry criteria, the informed consent has been signed and the eligibility has been verified, the ESF had to be sent as by the indication in the form. The complete registration form was then faxed back to the center after confirmation of the eligibility and the assignment of the patient number.

In this study, there was only a single arm and no randomization was foreseen.



### 3.2. SELECTION OF STUDY POPULATION

This was a multicenter clinical trial involving about 40 Italian centers and up to 200 patients were planned to be enrolled.

The target population was represented by patients with locally advanced or metastatic cancer of the colon or rectum starting first line combination therapy.

Under no circumstances enrolled subjects in this study were permitted for a second course of treatment.

#### 3.2.1. INCLUSION CRITERIA

- 1) Histologically or cytologically proven diagnosis of colorectal cancer.
- 2) Locally advanced or metastatic colorectal cancer not previously treated with chemotherapy for metastatic disease.
- 3) Age  $\geq 18$ .
- 4) ECOG Performance Status 0-1
- 5) Life expectancy of at least 12 weeks.
- 6) At least one measurable lesion according to RECIST criteria.
- 7) Neutrophils  $\geq 1.5 \times 10^9/L$  and Platelets  $\geq 100 \times 10^9/L$ .
- 8) Total bilirubin  $\leq 1.5$  time the upper-normal limits (UNL) of the Institutional normal values and ASAT (SGOT) and/or ALAT (SGPT)  $\leq 2.5 \times UNL$ , or  $\leq 5 \times UNL$  in case of liver metastases, alkaline phosphatase  $\leq 2.5 \times UNL$ ,  $\leq 5 \times UNL$  in case of liver metastases.
- 9) Creatinine clearance  $> 50$  mL/min or serum creatinine  $\leq 1.5 \times UNL$ .
- 10) Urine dipstick of proteinuria  $< 2+$ . Patients discovered to have  $\geq 2+$  proteinuria on dipstick urinalysis at baseline, should have undergone a 24-hour urine collection and had to demonstrate  $\leq 1$  g of protein/24 hr.
- 11) Written informed consent.
- 12) Patients had to be accessible for treatment and follow-up. Patients registered on this trial had to be treated and followed at the participating Center.

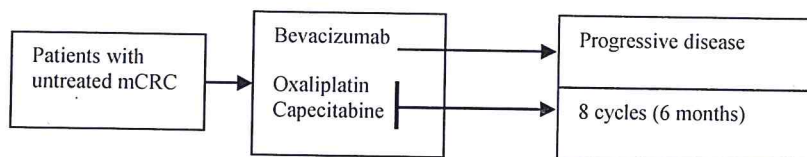
#### 3.2.2. EXCLUSION CRITERIA

- 1) Radiotherapy to any site within 4 weeks before the study.
- 2) Untreated brain metastases or spinal cord compression or primary brain tumours.
- 3) History or evidence upon physical examination of CNS disease unless adequately treated (e.g., seizure not controlled with standard medical therapy or history of stroke).
- 4) Serious, non-healing wound, ulcer, or bone fracture.
- 5) Evidence of bleeding diathesis or coagulopathy.
- 6) Uncontrolled hypertension.
- 7) Clinically significant (i.e. active) cardiovascular disease for example cerebrovascular accidents ( $\leq 6$  months), myocardial infarction ( $\leq 6$  months), unstable angina, New York Heart Association (NYHA) grade II or greater congestive heart failure, serious cardiac arrhythmia requiring medication.

- 8) Current or recent (within 10 days prior to study treatment start) ongoing treatment with anticoagulants for therapeutic purposes.
- 9) Chronic, daily treatment with high-dose aspirin (>325 mg/day).
- 10) Treatment with any investigational drug within 30 days prior to enrolment.
- 11) Patients with known allergy to Chinese hamster ovary cell proteins, or any of the components of the study medications.
- 12) Other co-existing malignancies or malignancies diagnosed within the last 5 years.
- 13) Major surgical procedure, open biopsy, or significant traumatic injury within 28 days prior to study treatment start, or anticipation of the need for major surgical procedure during the course of the study.
- 14) Lack of physical integrity of the upper gastrointestinal tract, malabsorption syndrome, or inability to take oral medication.
- 15) Pregnant or lactating women. Women of childbearing potential with either a positive or no pregnancy test at baseline. Postmenopausal women must have been amenorrheic for at least 12 months to be considered of non-childbearing potential. Sexually active males and females (of childbearing potential) unwilling to practice contraception during the study.
- 16) Symptomatic peripheral neuropathy  $\geq$  grade 1 according the NCI Common Toxicity Criteria.

### 3.3. TREATMENTS

Treatments were administered as follows:

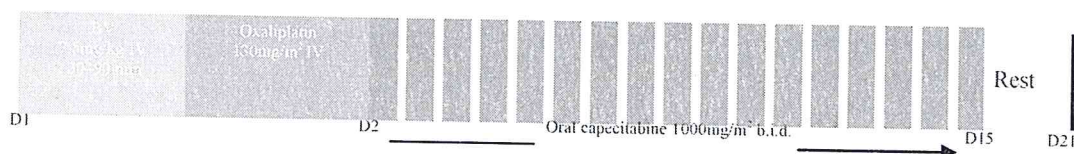


The treatment regimen followed a 21 days cycle:

- **Bevacizumab** had to be given by intravenous infusion at the dose of 7.5 mg/kg every 21 days (q3w).
- **Oxaliplatin** had to be administered by intravenous infusion at the dose of 130 mg/m<sup>2</sup> every 3 weeks (q3w).
- **Capecitabine** had to be taken orally at the dose of 1000 mg/m<sup>2</sup> twice daily (equivalent to a total daily dose of 2000 mg/m<sup>2</sup>) for 14 days continuously every 3 weeks (q3w).

The treatment regimen is summarized in the scheme below.





Treatment had to be continued on the basis of tumour assessment, according to the following criteria:

- Progressive disease: go off protocol treatment
- Stable disease, complete or partial response: Bevacizumab treatment had to be continued until progressive disease, unacceptable toxicity, patient or physician's decision. Oxaliplatin and Capecitabine were stopped after a maximum of 8 cycles.

Therapy after completion of the protocol was at the Investigator's discretion.

Patients with inoperable locally advanced or metastatic colorectal cancer, candidate for surgery following an initial response to treatment study, should had stopped bevacizumab therapy at least 6-8 weeks before surgery.

### 3.3.1. METHOD OF ASSIGNING PATIENTS TO TREATMENT GROUP

No randomization procedures for treatment allocation were required, since all the patients received the same treatment.

## 3.4. EFFICACY AND SAFETY VARIABLES

### 3.4.1. EFFICACY AND SAFETY MEASUREMENTS ASSESSED AND FLOW CHART

The duration of the study was planned from February 2008 to April 2011, assuming a 10 months median PFS, with a planned recruitment period of 12 months and a planned follow-up period of 1 years.

Patient had to be followed at the study Center according to the visit schedule and assessments outlined in the flow chart below.

### STUDY PROCEDURE

Parameters	Baseline	Treatment period		End of treatment	Long term follow-up
		Every cycle (3 weeks)	Every 3 cycles (9 weeks)		
Informed consent	X				
Medical history	X				
Physical examination	X	X			X



## Statistical Report

**Sponsor: Roche**

**Protocol: ML 21380**

Vital signs	X	X			X
ECOG PS	X	X			X
Concomitant diseases	X	X			
Concomitant treatments	X	X			
Toxicity evaluation		X		X <sup>5</sup>	
ECG	X			If clinically indicated	
Hematology <sup>1</sup>	X	X			
Blood chemistry <sup>2</sup>	X	X			
Chest X-ray	X				
INR/APTT <sup>6</sup>	X				
Dipstick proteinuria	X	X			
Chest/Abdomen CT or abdomen MRI <sup>4</sup>	X				If clinically indicated
Tumor evaluation	X		X <sup>3</sup>	X	X <sup>3</sup>
K-Ras mutation status <sup>7</sup>	X (OPTIONAL)				
B-Raf mutation status <sup>7</sup>	X (OPTIONAL)				
Quality of life	X		X	X	

1. White blood cell count with neutrophils/granulocytes, red blood cell count, hemoglobin, hematocrit and platelet count.
2. Bilirubin (total and direct), ASAT, ALAT, alkaline phosphatase, albumin, LDH, serum creatinine, glucose, electrolytes (sodium, potassium), calcium, pregnancy test (if clinically indicate).
3. Using the same technique performed at baseline.
4. Up to 30 days before the start of treatment.
5. Follow-up on any unresolved adverse event 28 days after last drugs intake.
6. For all patients at baseline. In addition, pre-cycle throughout the study treatment phase in those patients who were on anticoagulant therapy or at the occurrence of thromboembolic events or haemorrhage.
7. To be conducted as soon as this optional analysis was approved by local EC/CA and after the patient had signed the informed consent form specific for this part of the study. Tumour sample to be used was the one already available at the site, upon which the diagnosis of tumor was originally performed. Instructions on the handling of biopsy specimens and specifications of the genomic analysis were supplied in the Central Laboratory manual.

Tumor evaluations were made basing on RECIST criteria. Response to treatment had to be assessed every 9 weeks (three cycles) by using the same techniques performed at baseline. At the end of treatment tumour assessment had to be performed every 3 months, during follow up visits until PD, for all patients who went off study for other reason than PD. After PD, follow up visits continued until patient's death to collect survival data.

At baseline, each tumor lesion was categorized in measurable or nonmeasurable:

- Evaluable measurable lesions - Lesions that can be accurately measured in at least one dimension with longest diameter  $\geq 20$  mm using conventional techniques (CT scan or MRI) or  $\geq 10$  mm with spiral CT scan;
- Evaluable nonmeasurable lesions - The nonmeasurable lesions include the lesions with longest diameter  $<20$  mm with conventional techniques (CT scan or MRI) or  $<10$  mm with spiral CT scan and truly nonmeasurable lesions (bone lesions, leptomeningeal disease, ascites, pleural/pericardial effusion, inflammatory breast disease, lymphangitis cutis/pulmonis, abdominal masses that are not confirmed and followed by imaging techniques, and cystic lesion);
- Not evaluable lesions - Tumor lesions that are situated in a previously irradiated area might not be considered measurable.

For the evaluation, the lesions were classified at baseline as either target or nontarget:

- Target lesions: identified as all measurable lesions, up to a maximum of 5 lesions per organ and 10 lesions in total, representative of all involved organs. Target lesions should be selected on the basis of their size (lesions with the longest diameter) and their suitability for accurate repeated measurements (either by imaging techniques or clinically). A sum of the longest diameter for all target lesions was calculated and reported as the baseline sum of the longest diameter. It was used as reference by which to characterize the objective tumor response.
- Nontarget lesions: identified as all other lesions, that is lesions not fulfilling the criteria for target lesions at baseline. The presence or absence of each was noted every tumor re-evaluation assessment.

At each assessment, response was evaluated first separately for the target and nontarget lesions identified at baseline. The evaluation of target lesions was done using the following criteria:

- Complete Response (CR): disappearance of all target lesions;
- Partial Response (PR): at least a 30% decrease in the sum of the longest diameter of target lesions, taking as reference the baseline sum longest diameter;
- Stable disease (SD): neither sufficient shrinkage to qualify for partial response nor sufficient increase to qualify for progressive disease, taking as reference the smallest sum longest diameter since the treatment started;
- Progressive Disease (PD): at least a 20% increase in the sum of the longest diameter of target lesions, taking as reference the smallest sum longest diameter recorded since the treatment started. When the progression was observed before 12 weeks after entry into the study, the patient had to be considered as an "early progression";
- Not assessable (NA).

The evaluation of nontarget lesions was done using the following criteria:

- Complete Response (CR): the disappearance of all nontarget lesions and normalization of tumor marker level;



- Neither CR/ Neither PD (Non-CR/Non-PD): the persistence of one or more nontarget lesions and/or the maintenance of tumor marker level above the normal limits;
- Progressive Disease (PD): the appearance of one or more new lesions and/or unequivocal progression of existing non target lesions;
- Not assessable (NA).

These evaluations were used to calculate the overall lesion response considering both the target and non-target lesions together as well as the presence or absence of totally new lesions (see Table 3.4.1-1).

**Table 3.4.1-1: Overall lesion response.**

Target lesions	Nontarget lesions	New lesions	Overall lesion response
CR	CR	No	CR
CR	Non-CR/Non-PD	No	PR
PR	Non-PD	No	
SD	Non-PD	No	SD
PD	Any	Yes or No	PD
Any	PD	Yes or No	
Any	Any	Yes	
No PD	NA	No	NA
NA	No PD	No	

### 3.4.2. PRIMARY EFFICACY VARIABLES

The primary efficacy parameters was represented by the Progression Free Survival (PFS) defined as the time period from the start of the study treatment to the first disease progression or death for any cause, whichever occurred first. Patients without an event were censored at the time of last contact patient was to be known progression free or alive.

Disease progression was based on Investigator's assessment; evaluation of target lesions and nontarget lesions was in accordance with the RECIST criteria. For this reason, the evaluation of this parameter concerned the 1<sup>st</sup> line period of the study.

Patients with no tumour assessments after baseline but who were alive at the end of 1<sup>st</sup> line discontinuation study were censored at day 1.

For patients who undergone surgery after experiencing a sufficient shrinkage of the tumour, any relapse, new occurrence of colorectal cancer or death were considered as an event. Patients undergone surgery without any such event were censored at the date of the last tumour assessment that documented that neither a relapse nor a new colorectal cancer had occurred.

### 3.4.3. SECONDARY EFFICACY VARIABLES

The secondary efficacy parameters were represented by:



- Overall Response Rate (ORR): defined as the proportion of patients with a Best Overall Response of CR or PR during 1<sup>st</sup> line phase of the study. Best Overall Response was defined as the best response designation recorded from the start of treatment until disease progression or 1<sup>st</sup> line discontinuation study based on RECIST criteria.
- Time to Response: time to overall response (CR or PR) was calculated as the time between date of start of treatment until first documented response (CR or PR). This analysis included all patients and it concerned the 1<sup>st</sup> line period of the study. Patients who did not achieve a PR or CR were censored at the date of progression or death, or at last adequate tumour assessment date.
- Duration of response (DR):
  - only for patients whose Best Overall Response was CR or PR, the duration of overall response were measured from the time that measurement criteria was met for CR or PR (whichever occurred first) until the first date that progressive disease was objectively documented or death due to underlying cancer, whichever occurred first;
  - as to SD, only for patients whose Best Overall Response was CR, PR or SD the duration of stable disease was measured from the start of the treatment until the criteria for disease progression were met or death due to underlying cancer, whichever occurred first.

For both, in case of censoring the date of last adequate tumour assessment was used. For that, also the evaluation of this parameter concerned the 1<sup>st</sup> line period of the study.
- Time to Treatment Failure (TTF): defined as a composite endpoint measuring time from 1<sup>st</sup> day of treatment to discontinuation of treatment for any reason, including: death due to any cause, adverse event, insufficient therapeutic response (progression of disease), failure to return (lost to follow-up), refusing treatment (patients non-compliance), being unwilling to cooperate and withdrawing consent (patient withdrew consent). In case of censoring, the earlier of the date between the last tumour assessment and the date of the last intake of study medication was used.
- Overall survival (OS): defined as time from the date of 1<sup>st</sup> day of treatment until the date of death from any cause. If a patient was not known to have died, survival was censored at the last date the patient was known to be alive. For this parameter, all the study period were considered.
- R0 resectability rate: identified as the percentage of patients with substantial tumor shrinkage allowing the complete surgical resection with no macroscopic or microscopic tumor remains.
- K-Ras mutation status and the B-Raf mutation status.

- Quality of life: assessed by means of the EuroQoL EQ-5D questionnaire.

#### 3.4.4. SAFETY VARIABLES

Treatment Exposure. The administration of study treatment during the 1<sup>st</sup> line phase was considered.

Adverse Events. All adverse events occurred during the clinical study were reported on the Case Report Form. An adverse event is any adverse change from the patient's baseline (pre-treatment) condition, including intercurrent illness, which occurs during the course of a clinical study after treatment has started, whether considered related to treatment or not.

The intensity of clinical adverse events was graded according to the NCI Common Toxicity Criteria (CTC) version 3.0 grading system in the toxicity categories that have recommended grading. In addition to intensity, the duration, the relationship to the trial medication, the action taken regarding the drug administration, any treatment given for the adverse event, as well as the outcome of the adverse event had to be reported in the CRF.

AEs, especially those for which the relationship to study medication(s) is not "unrelated", should have been followed up until they have returned to baseline status or stabilized. All unrelated, mild or moderate events had to be followed up for 28 days after the last dose of the study drug.

All new adverse events occurred up till 28 days after the last dose of study treatment should have been recorded in the AE page of the CRF. Related non-serious new AEs occurring up to 6 months after the last dose of study drug should have been reported.

Related serious events should have been reported indefinitely.

Haematology and blood chemistry data. Each value was classified as normal (N), low (L) or high (H) according to the normal range of the site referring to. The values were also converted on the same unit of measurement, the Preferred one, by means of conversion factors (see Appendix I of "Statistical Analysis Plan" version 1 - 20/02/2012).

Vital signs. Systolic blood pressure (mmHg), diastolic blood pressure (mmHg), pulse rate (beats per minute), weight (Kg) and Body Surface Area (m<sup>2</sup>).

Performance status. It was valued by means of ECOG scale.

ECG. This exam was collected every 3 cycles and not only at baseline as erroneously showed in the flow chart on protocol.

Proteinuria. The values were converted on the same unit of measurement, the Preferred one, by means of conversion factors (see Appendix I of "Statistical Analysis Plan" version 1 - 20/02/2012).



### 3.5. DATA MANAGEMENT

#### 3.5.1. DATA COLLECTION

The study data collection was carried out through a paper CRF.

The original copy (white) was forwarded to the Data Entry personnel for data entry into the database.

Double Data Entry (DE) was carried out on an ongoing basis according to Sponsor's requirement: two data entry clerks entered data into the study database with interactive verification. The second DE clerk reviewed and resolved during the second run all the discrepancies between first and second entry.

Entry errors were documented by the electronic audit trail associated to the System. All indications and conventions used for homogeneous and correct data entry were collected in the Data Entry Instruction Document.

#### 3.5.2. DATABASE MANAGEMENT AND QUALITY CONTROL

The clinical database was implemented in Oracle Clinical (Oracle Pharmaceutical Applications Release 4i (4.5.0)).

In addition to the variables listed in the CRF, the following variables were included into the final SAS database: age at baseline, laboratory normal range and unit of measures, MedDRA codes for medical terms and WHO-DRL codes for therapies.

The data validation procedures were programmed inside Oracle Clinical, except for some checks that were implemented in SAS program: the system (Oracle Clinical/SAS) used to capture the data errors is specified in the Data Validation Document.

Specific missing data, data inconsistencies or discrepancies which were identified as clearly obvious were resolved by Data Manager himself without further Investigators' intervention. The permitted Obvious Corrections listings had been presented in the Obvious Corrections Document.

A laboratory normal range form as soon as possible with all information required for evaluation of laboratory parameters was filled.

In case of change even of only one piece of information present in the form, it was responsibility of the Clinical Project manager to provide a new form to the Data Manager with the update information.

In case of discrepancies, where it was possible, in the database the units reported in the Normal Ranges for laboratory parameters were entered instead of the ones of CRF.

Coding of medical conditions / trade names of drugs was performed using the International dictionaries MedDRA (version 11.1) and WHO-DRL (Q4\_2007).

Coding was handled by the DM by means of programs which automatically link the free text specified by the Investigators to the dictionary terms. If no automatic match was found, coding was made manually by the DM. In case of multi-axial ATC codes, the correct one was chosen on the basis of the drug indication reported by the Investigator.



Before the database lock, the DM sent to the Medical Advisor the listing of manual coded terms/therapies for approval and the listing of uncoded terms/therapies which were reviewed /corrected by the Medical Advisor and returned to DM dated and signed.

### 3.6. STATISTICAL METHOD PLANNED IN THE PROTOCOL AND SAMPLE SIZE DETERMINATION

#### 3.6.1. STATISTICAL AND ANALYTICAL PLAN

##### Population of analysis

The assignment of patients to populations was performed before closing the database, using a predefined protocol violation list. Three sets of Analysis Population were defined:

##### Safety population

All patients who signed the informed consent and who received at least one dose of all study medication.

##### Intent-To-Treat population (ITT)

All patients enrolled that meet the following criteria were included in the ITT:

- signed the informed consent;
- at least one dose of all study medication taken;
- at least one efficacy measurement at baseline available.

##### Per Protocol population (PP)

From the PP population were excluded who:

- did not receive at least 3 months of all study treatments (for reasons other than progressive disease, adverse event and death);
- did receive at least 3 months of all study treatments but received less than 50% of the anticipated treatment of at least one of study drug during the first 3 months;
- severely violated protocol inclusion or exclusion criteria.

Patient validation for the analyses was done before the DB lock by means of SAS program checks and patient data listings.

The results were summarized in the "Patient Validation Document" (version 2 – 14/03/2012) that was approved by the Sponsor study team.

Evaluability checks and rules for efficacy analysis exclusion are described in the following table:

Nr.	Type*	Evaluability checks	Excluded from...	Checked by DB**
1	I.C.	No histologically or cytologically proven diagnosis of colorectal cancer.	PP	x
2	I.C.	Not locally advanced or metastatic colorectal cancer previously treated with chemotherapy for metastatic disease.	PP	
3	I.C.	Age < 18 years.	PP	xx

## Statistical Report

**Sponsor: Roche**

**Protocol: ML 21380**

Nr.	Type*	Evaluability checks	Excluded from...	Checked by DB**
4	I.C.	ECOG Performance Status greater than 1.	PP	xx
5	I.C.	Life expectancy of less than 12 weeks.	PP	
6	I.C.	Not at least one measurable lesion according to RECIST criteria.	ITT/PP	xx
7	I.C.	Neutrophils < 1.5 x 10 <sup>9</sup> /L and Platelets < 100 x 10 <sup>9</sup> /L	PP	xx
8	I.C.	Total bilirubin > 1.5 time the upper-normal limits (UNL) of the Institutional normal values and ASAT (SGOT) and ALAT (SGPT) > 2.5 x UNL, or > 5 x UNL in case of liver metastases, alkaline phosphatase > 2.5 x UNL, > 5 x UNL in case of liver metastases.	PP	xx
9	I.C.	Creatinine clearance ≤ 50 mL/min or serum creatinine > 1.5 x UNL.	PP	x
10	I.C.	Urine dipstick of proteinuria ≥ 2+ at baseline without a 24-hour urine collection that demonstrated ≤ 1 g of protein/24 hr.	PP	xx
11	I.C.	Lack of written informed consent.	Safety/ITT/PP	xx
12	I.C.	Patients were not accessible for treatment and follow-up. Patients registered on this trial were not treated and followed at the participating Center.	PP	x
13	E.C.	Radiotherapy to any site within 4 weeks before the study.	PP	
14	E.C.	Untreated brain metastases or spinal cord compression or primary brain tumours.	PP	x
15	E.C.	History or evidence upon physical examination of CNS disease unless adequately treated (e.g., seizure not controlled with standard medical therapy or history of stroke).	PP	xx
16	E.C.	Active serious non-healing wound, ulcer or bone fracture.	PP	xx
17	E.C.	Evidence of active bleeding diathesis or coagulopathy.	PP	xx
18	E.C.	Uncontrolled active hypertension.	PP	xx
19	E.C.	Clinically significant (i.e. active) cardiovascular disease, for example cerebrovascular accidents (≤6 months), myocardial infarction (≤6 months), unstable angina, New York Heart Association (NYHA) grade II or greater congestive heart failure, serious cardiac arrhythmia requiring medication.	PP	x §
20	E.C.	Current or recent (within 10 days prior to study treatment start) ongoing treatment with anticoagulants for therapeutic purpose.	PP	x §
21	E.C.	Chronic, daily treatment with high-dose aspirin (>325 mg/day).	PP	x §
22	E.C.	Treatment with any investigational drug within 30 days prior to enrolment.	PP	
23	E.C.	Patients with known allergy to Chinese hamster ovary cell proteins or any of the components of the study medications.	PP	x §
24	E.C.	Other co-existing malignancies or malignancies diagnosed within the last 5 years.	PP	x §
25	E.C.	Major surgical procedure, open biopsy or significant traumatic injury within 28 days prior to study treatment start, or anticipation of the need for major surgical procedure during the course of the study.	PP	X
26	E.C.	Lack of physical integrity of the upper gastrointestinal tract, malabsorption syndrome or inability to take oral medication.	PP	XX
27	E.C.	Pregnant or lactating women. Women of childbearing potential with either a positive or no pregnancy test at baseline. Postmenopausal women must have been amenorrheic for at least 12 months to be considered of non-childbearing potential. Sexually active males and females (of childbearing potential) unwilling to practice contraception	PP	X



Nr.	Type*	Evaluability checks	Excluded from...	Checked by DB**
		during the study.		
28	E.C.	Symptomatic peripheral neuropathy $\geq$ grade 1 according the NCI Common Toxicity Criteria.	PP	XX
29	O.C.	Patients who didn't receive at least one dose of all study medications.	Safety/ITT/PP	XX
30	O.C.	Patients who didn't receive at least 3 months of treatment (for reasons other than progressive disease, adverse events and death).	PP	XX
31	O.C.	Patients who did receive at least 3 months of treatment but received less than 50% of the anticipated treatment of at least one study drug during the first 3 months.	PP	XX
32	O.C.	At least one measurable lesion not identified as foreseen by RECIST criteria.	PP	XX
33	O.C.	More than 5 target lesions per organ or more than 10 target lesions in total.	PP	XX
34	O.C.	Tumour evaluation assessed by using different techniques performed at baseline.	PP	XX
35	O.C.	Administration of not permitted concomitant therapies.	PP	X
36	O.C.	Informed consent signed after baseline visit.	PP	XX
37	O.C.	Study treatment began after more than 14 days from baseline visits.		XX
38	O.C.	No efficacy data about tumor assessment after baseline visits expect for patients who didn't receive at least 3 months of treatment due to progressive disease or adverse events or death.	PP	XX
39	O.C.	Patient didn't go off protocol treatment after progression of disease.		XX

\*I.C.=Inclusion Criteria; E.C.=Exclusion Criteria; O.C.=Other Criteria.

\*\* xx=check done from DB; x=check partially done from DB; §=check done from Sponsor.

## General methodology

The data were analysed by means of the SAS System for Windows version 9.2.

First of all a complete description of patients disposition was provided, specifying the number of enrolled patients, the number of patients at each visit, completed and discontinued patients with the reasons of the discontinuation of treatment. The number of enrolled patients was provided also by center.

A complete description of the number of protocol violators and the number of patients per violation was summarized in the "Patient Validation Document" (version 2 – 14/03/2012).

The numerosness of the analysis populations was described and the reason excluding the patient from any analysis population were provided.

A complete description of all data collected at baseline, including recorded and derived variables, was provided per treatment on the patients valid for Safety. They were analysed by means of usually descriptive statistics: mean, standard deviation, median, 25<sup>th</sup> and 75<sup>th</sup> percentiles, minimum and maximum values for continuous variables; absolute and relative frequencies for categorical ones.



As regards primary disease history, as patients could have more than one primary tumour in case of multiple status at first diagnosis new categories able to take into account the characteristics were created.

Previous and concomitant medications were described by ATC CODE description (Level 2) and by Preferred Term of WHODRL dictionary, version 2009. Similarly, medical history and concomitant diseases were described by means of Preferred Term of MedDRA dictionary, 11.1 version.

As far as descriptive statistics are concerned, the prevalence approach was adopted in case of missing data except for quality of life data as explained as follow in the corresponding section.

For all 'time to event' variables, other than the duration of overall response, the date of start of treatment was used as the start date. Instead, for the calculation of duration of overall response the date of RECIST assessment of 1<sup>st</sup> CR o PR as overall lesion response was used.

The end dates which were used to calculate 'time to event' variables are defined as follows:

- Date of death;
- Date of progression was identified as the first tumour assessment date at which the overall response was recorded as progressive disease. When there was no documentation of radiologic evidence of progression, but the patient discontinued treatment for 'Disease progression' due to documented clinical deterioration of disease, the date of treatment discontinuation was used as date of progression;
- Date of last adequate tumour assessment was identified as the date of the last tumour assessment with overall lesion response of CR, PR or SD which was made before an event or a censoring reason occurred. In this case the last tumor evaluation date at that assessment was used;
- Date of treatment discontinuation was identified as the date of the end of 1<sup>st</sup> line phase;
- Date of last contact was defined as the last date the patient was known to be alive. This corresponds to the latest date between the visit date, laboratory sample date, tumour assessment date, last treatment intake or date of discontinuation. If available, the last contact date from survival follow-up page was used too.

For the calculation of time to overall response, the date of RECIST assessment of 1<sup>st</sup> CR o PR as overall lesion response was used as event date.

All statistical tests were performed with a significance level  $\alpha=0.05$ .

Primary efficacy analysis

Primary efficacy analysis were based primarily on the ITT population, the analyses on the PP population were considered as supportive.

**Progression Free Survival**

Progression Free Survival was presented in Kaplan-Meier plots and summarized by median and 95% confidence interval.

As the hypothesis testing of this study was based on the confidence interval estimate, if the two-sided 95% confidence interval for Progression Free Survival median did not include the 7 months Progression Free Survival median historic controls, the evidence of improvement for Progression Free Survival was proved.

Secondary efficacy analyses

Secondary efficacy analyses were also based primarily on the ITT population, subsequently on the PP population. The secondary efficacy parameters considered are explained subsequently.

**Overall Response Rate (ORR).**

Descriptive statistics of the overall lesion response were provided by visit/tumour assessment.

The Best Overall Response, achieved within the time from start of treatment until disease progression or 1<sup>st</sup> line study discontinuation based on RECIST criteria, was described by means of the usual descriptive statistics for discrete variables.

The proportion of patients achieving CR or PR as best overall response was calculated and described. Confidence interval at 95% was also computed by means of the one sample exact binomial method.

A listing of Overall Tumor Response at each tumour assessment and Best Overall Response at the end of treatment or premature discontinuation was also provided by patient.

**Time to Response, Duration of overall response, Duration of stable disease, Time to Treatment Failure and Overall survival.**

All 'time to event' variables were presented in Kaplan-Meier plots and summarized by median and 95% confidence interval.

**R0 resectability rate.**

The surgery during the study period was described jointly by reason (i.e. curative, palliative, biopsy, reconstructive, other or unknown) and residual disease (i.e. R0 – no residual due to radical surgery, presence of residual disease, unknown or not applicable). Moreover, the proportion of radical surgery between patients with locally advanced or metastatic colorectal cancer who performed surgery was provided.



### **K-Ras and B-Raf mutation status.**

Absolute and relative frequencies of patients with K-Ras mutation and patients with B-Raf mutation were provided. Type of mutation of K-Ras and B-Raf were also tabulated.

To test the potential predictive value of K-Ras and B-Raf gene alterations in patients with colorectal cancer on the overall response rate a Chi-square test was performed and the Relative Risk with the 95% exact confidence limits were calculated.

Time to Response and Progression Free Survival were separately analysed, with an explorative purpose, by means of the Log-rank test, in order to test differences among patients with gene alterations (K-Ras and B-Raf) and patients without gene alterations. Median time to response and median Progression Free Survival with 95% confidence limits were calculated and Kaplan-Meier curves were provided.

### **Quality of life.**

Descriptive results were provided with prevalence approach and also by means of Last Observational Carried Forward (LOCF) approach applied to questionnaire's response. For this last approach, only patients who completed the questionnaire at least at baseline were considered.

Quality of life assessments were used to derive pre-specified QoL scores according to the QoL manual "EQ-5D-3L User Guide - Basic information on how to use the EQ-5D-3L instrument Version 4.0. EuroQol Group Executive Office on behalf of the EuroQol Group". See Paragraph 5.4.2 for more details.

These scores were summarized by descriptive summary tables at baseline and over time. To evaluate the overall health score changes from baseline to the end of treatment, an ANCOVA model must be applied to overall health score absolute changes and to overall health score indicated at the end of treatment as covariate.

The overall health score absolute changes were calculated for each patient as follows:

$$(\text{score at the end of treatment} - \text{score at baseline})$$

The normality assumption of these overall health score absolute changes was tested by means of graphical methods and Shapiro-Wilk test. In case of non-normality, a rank ANCOVA model must be adopted.

### **Safety analyses**

All the safety analyses were performed on the population of patients valid for safety. The safety parameters considered are explained subsequently.

### **Duration of the study period.**

It was calculated as the time in days elapsed from the date of the baseline visit until the date of death after treatment discontinuation or the date of last follow-up available and it was described with the usual descriptive statistics for continuous variables distinctly for 1<sup>st</sup> line period and follow-up time.

### **Treatment Exposure.**

Descriptive statistics of the number of cycles performed by patient for the three study treatments were provided. In addition, the number of doses modification/interruption and the specify reasons were described.

### **Adverse Events.**

Tables reporting a general summary of adverse events and tables displaying drug-related adverse events and serious adverse events were produced, specifying the number of total events and the absolute and relative frequency of subjects with events. As a patient may have more than one adverse event, the total number of adverse events could be greater than the total number of patients.

If the same patients reported more than one occurrence of the same adverse event also with the same intensity, no selection of the one with the worst intensity was adopted. So, for the analyses, all events occurred for each patient and reported on CRF, were considered to obtain a more exhaustive consideration about safety.

The incidence rates of all causalities and treatment emergent adverse events were tabulated by System Organ Class (level 2) and Preferred Terms of MedDRA dictionary, 11.1 version. Severity and relationship to study drug were presented per each event.

Distinct tables were provided for toxicities observed with Bevacizumab: in terms of PT, "Scleral haemorrhage", "Anal haemorrhage", "Enterocolitis haemorrhagic", "Haemorrhoidal haemorrhage", "Intestinal perforation", "Mouth haemorrhage", "Rectal haemorrhage", "Proteinuria", "Renal vein thrombosis", "Genital haemorrhage", "Pulmonary embolism", "Aortic thrombosis", "Axillary vein thrombosis", "Deep vein thrombosis", "Hypertension", "Hypertensive crisis", "Thrombosis" and "Vena cava thrombosis" were considered for this analysis, selected on base of indication reported on protocol even if AE was not signaled on CRF as related to Bevacizumab.

### **Haematology and blood chemistry data.**

Summary tables were produced for the converted values on all haematology and blood chemistry parameters, showing the number of observations at each time point, the mean, median, standard deviation, inter-quartile range and the lowest and highest values.

In addition, for each parameter a transition frequency table (4x4 contingency table where the forth category is for missing data) was adopted to describe patients with changes in relation to normal ranges, at the last available visit in which each laboratory data was collected for any patient versus baseline.

Incidence rates of abnormal lab values were tabulated. These incidence rates were calculated as the number of events divided by the number of patients at risk, where:

- the number of events was the number of patients reporting the specified abnormality at last available visit and without any abnormalities at baseline;
- the number of patients at risk was the number of patients who did not report any abnormality at baseline.



Moreover, for the computation of rates in case of missing data at baseline, patients were considered at risk; while in case of missing data at both baseline and at the end of treatment the patients were excluded.

Individual patient listings of abnormal values of hematology and blood chemistry data were also provided.

### **Other safety data.**

Summary tables were produced for all vital signs showing the number of observations at each time point, mean, standard deviation, median, 25<sup>th</sup> and 75<sup>th</sup> percentiles, minimum and maximum values.

The ECG, performance status, and proteinuria data were described by visit.

### **3.6.2. SAMPLE SIZE DETERMINATION**

The sample size chosen for this exploratory study was such that it provided enough precision to allow for a valid assessment of study result with regard to historic controls.

Assuming that median Progression Free Survival (PFS) of historic controls was 7.3 months (XELOX vs. FOLFOX-4: efficacy results from XELOX-I), an improvement by 2 months would lead to a 64% Progression Free Survival rate at 7 months for the Bevacizumab + chemotherapy treatment in the study. With 190 evaluable patients, the two-sided 95% confidence interval for Progression Free Survival would not include the 7 months Progression Free Survival rate for historic controls of 50%, which would provide evidence for improved Progression Free Survival. Therefore 200 patients had to be recruited into this study, which allows for a 5% dropout rate to ensure the evaluability of 190 patients.

### **3.7. CHANGES IN THE CONDUCT OF THE STUDY OR PLANNED ANALYSES**

Considering specific codon K-Ras mutations and B-Raf, frequencies of patients with gene alterations were very low. For that, all the secondary efficacy analyses were done comparing a unique group identified as patients with at least a gene alterations (K-Ras or B-Raf) vs. patients without any gene alterations.

The EQ-5D-3L answers and EQ-5D-3L raw index values by visit were also graphically represented by means of an appropriate bar-chart for categorical data and of line chart of the mean distribution for continuous data.

The overall health score changes from baseline to the end of treatment, were tested by means of paired T-test for data with normal distribution or by means of Signed-rank test otherwise.

### **3.8. INTERIM ANALYSIS**

The ethical performance and the safety data were periodically supervised by a specific scientific committee. From October 2008, the committee was updated through a formal report about every three months in which AE, SAE and toxicities arising during the study were summarized.

An interim analysis was also performed on February 2011 on all the 205 patients enrolled up to 28<sup>th</sup> January 2011 (without any patient validation) with an explorative purpose and data were presented during the Investigator's Meeting planned on 1<sup>st</sup> March 2011. Characteristics at baseline, Progression free survival, overall response rate, duration of response, overall survival and adverse events were analyzed. As concerning the primary efficacy analysis, the 50% of patients at risk had a first PD or died due to any cause at 10.26 months since treatment start, with the two-sided 95% confidence interval for Progression Free Survival median corresponding to [8.79-11.21]. Consequently, the evidence of improvement for PFS could be provide since the 7 months Progression Free Survival median historic controls was not included in the two-sided 95% confidence interval for Progression Free Survival.

## 4. STUDY PATIENTS

### 4.1. DISPOSITION OF PATIENTS

Two hundred and five patients were enrolled in the study in 34 Centers. All of them discontinued from the 1<sup>st</sup> line of the study and the main reasons for discontinuation were the progression of disease (45.37%) and the occurrence of adverse event(s) (25.37%). All the reasons for discontinuation are reported in the following Table 4.1-1 and further details were provided in Tables 1.1.1 and 1.1.3 in the "Tables and Figures" document.

**Table 4.1-1: Reasons for premature discontinuation.**

		(N=205)	
		N	%
Reason for end of treatment	Adverse event(s)	52	25.37
	Progression of disease	93	45.37
	Protocol violation	5	2.44
	Patient withdrew consent	13	6.34
	Patient non-compliance	5	2.44
	Need for surgery	17	8.29
	Medical decision	14	6.83
	Death	6	2.93

Source Table 1.1.1 of "Tables and Figures" document.

The disposition of patients by follow-up visits is reported in Tables 1.1.2 and 1.1.4 in the "Tables and Figures" document.

### 4.2. PROTOCOL DEVIATION

According to the evaluability checks defined in the "Patient Validation Document" (version 2 – 14/03/2012), 54.15% of enrolled patients had at least one violation. The number of patients found per each violation is summarized in the following table:



## Statistical Report

**Sponsor: Roche**

**Protocol: ML 21380**

*Table 4.2-1: Summary of protocol violators (Enrolled population).*

Type	Nr. of violation	Violation	(N=205)	
			N	%
I.C.	6	Not at least one measurable lesion according to RECIST criteria.	1	0.49
	7	Neutrophils < 1.5 x 10 <sup>9</sup> /L and Platelets < 100 x 10 <sup>9</sup> /L.	2	0.98
	8	Total bilirubin > 1.5 time the UNL and ASAT (SGOT) and ALAT (SGPT) > 2.5 x UNL, or > 5 x UNL in case of liver metastases, alkaline phosphatase > 2.5 x UNL, > 5 x UNL in case of liver metastases.	26	12.68
	9	Creatinine clearance = 50 mL/min or serum creatinine > 1.5 x UNL.	1	0.49
	10	Urine dipstick of proteinuria = 2+ at baseline without a 24-hour urine collection that demonstrated = 1 g of protein/24 hr.	14	6.83
	12	Patients were not accessible for treatment and follow-up. Patients registered on this trial were not treated and followed at the participating Center.	5	2.44
E.C.	15	History or evidence upon physical examination of CNS disease unless adequately treated (e.g., seizure not controlled with standard medical therapy or history of stroke).	1	0.49
	16	Active serious non-healing wound, ulcer or bone fracture.	2	0.98
	17	Evidence of active bleeding diathesis or coagulopathy.	1	0.49
	18	Uncontrolled active hypertension.	7	3.41
	25	Major surgical procedure, open biopsy or significant traumatic injury within 28 days prior to study treatment start, or anticipation of the need for major surgical procedure during the course of the study.	10	4.88
	27	Pregnant or lactating women, women of childbearing potential with either a positive or no pregnancy test at baseline.	7	3.41
O.C.	29	Patients who didn't receive at least one dose of all study medications.	8	3.90
	30	Patients who didn't receive at least 3 months of treatment (for reasons other than progressive disease, adverse events and death).	22	10.73
	31	Patients who did receive at least 3 months of treatment but received less than 50% of the anticipated treatment of at least one study drug during the first 3 months.	23	11.22
	32	At least one measurable lesion not identified as foreseen by RECIST criteria	19	9.27
	33	More than 5 target lesions per organ or more than 10 target lesions in total.	7	3.41
	34	Tumour evaluation assessed by using different technique performed at baseline.	18	8.78
	36	Informed consent signed after baseline visit.	1	0.49
	37	Study treatment began after more than 14 days from baseline visits.	12	5.85
	38	No efficacy data about tumor assessment after baseline visits expect for patients who didn't receive at least 3 months of treatment due to progressive disease or adverse events or death.	11	5.37
	39	Patient didn't go off protocol treatment after progression of disease.	2	0.98

Source Table 1.2.1 of "Tables and Figures" document.

I.C.=Inclusion Criteria; E.C.=Exclusion Criteria; O.C. =Other Criteria.

Detailed listing of protocol violators is available in Table 1.2.2 in the "Tables and Figures" document.

## 5. EFFICACY EVALUATION

### 5.1. DATASET ANALYSED

The analysis populations are shown in the following Table 5.1-1:

*Table 5.1-1: Analysis populations.*

	N	%
Enrolled patients	205	100.00
Safety population	197	96.10
ITT population	197	96.10
PP population	98	47.80
Not valid	8	3.90

*Source Table 1.3.1 of "Tables and Figures" document.*

Patients '122708\_174', '122727\_12', '122729\_153', '122730\_71', '122730\_81', '122733\_44', '122738\_96' and '122747\_74' were considered as 'not valid' for the analyses because they didn't receive at least one dose of all study medications.

For this reason, these eight patients were also excluded from Safety and ITT population. Moreover, among these patients, patient '122730\_81' had not any measurable lesions according to RECIST criteria. See Table 1.3.3 in the "Tables and Figures" document for details about protocol violations.

One hundred and seven patients (including patients mentioned before) were excluded from PP population.

Reasons that excluded the patients from the PP population are reported in the following Table 5.1-2 (further details are collected in Tables 1.3.2 and 1.3.3 in the "Tables and Figures" document):



**Table 5.1-2: Reasons for exclusion of patients from PP population.**

Type	Nr. of violation	Violation	(N=205)	
			N	%
I.C.	6	Not at least one measurable lesion according to RECIST criteria.	1	0.49
	7	Neutrophils < 1.5 x 10 <sup>9</sup> /L and Platelets < 100 x 10 <sup>9</sup> /L.	2	0.98
	8	Total bilirubin > 1.5 time the UNL and ASAT (SGOT) and ALAT (SGPT) > 2.5 x UNL, or > 5 x UNL in case of liver metastases, alkaline phosphatase > 2.5 x UNL, > 5 x UNL in case of liver metastases.	26	12.68
	9	Creatinine clearance = 50 mL/min or serum creatinine > 1.5 x UNL.	1	0.49
	10	Urine dipstick of proteinuria = 2+ at baseline without a 24-hour urine collection that demonstrated = 1 g of protein/24 hr.	14	6.83
	12	Patients were not accessible for treatment and follow-up. Patients registered on this trial were not treated and followed at the participating Center.	5	2.44
E.C.	15	History or evidence upon physical examination of CNS disease unless adequately treated (e.g., seizure not controlled with standard medical therapy or history of stroke).	1	0.49
	16	Active serious non-healing wound, ulcer or bone fracture.	2	0.98
	17	Evidence of active bleeding diathesis or coagulopathy.	1	0.49
	18	Uncontrolled active hypertension.	7	3.41
	25	Major surgical procedure, open biopsy or significant traumatic injury within 28 days prior to study treatment start, or anticipation of the need for major surgical procedure during the course of the study.	10	4.88
	27	Pregnant or lactating women, women of childbearing potential with either a positive or no pregnancy test at baseline.	7	3.41
O.C.	29	Patients who didn't receive at least one dose of all study medications.	8	3.90
	30	Patients who didn't receive at least 3 months of treatment (for reasons other than progressive disease, adverse events and death).	22	10.73
	31	Patients who did receive at least 3 months of treatment but received less than 50% of the anticipated treatment of at least one study drug during the first 3 months.	23	11.22
	32	At least one measurable lesion not identified as foreseen by RECIST criteria	19	9.27
	33	More than 5 target lesions per organ or more than 10 target lesions in total.	7	3.41
	34	Tumour evaluation assessed by using different technique performed at baseline.	18	8.78
	36	Informed consent signed after baseline visit.	1	0.49
	38	No efficacy data about tumor assessment after baseline visits expect for patients who didn't receive at least 3 months of treatment due to progressive disease or adverse events or death.	11	5.37

Source Table 1.3.2 of "Tables and Figures" document.  
I.C.=Inclusion Criteria; E.C.=Exclusion Criteria; O.C. =Other Criteria.

## 5.2. DEMOGRAPHIC AND OTHER BASELINE CHARACTERISTICS

A description of demographic characteristics for Safety population is provided in the following Table 5.2-1.

**Table 5.2-1: Demographic data at baseline (Safety population).**

		(N=197)
Age (years)	N	197
	Mean	62.25
	SD	9.94
	25th percentile	56.00
	Median	63.00
	75th percentile	70.00
	Min	34
	Max	80

		(N=197)	
		N	%
Sex	Male	111	56.35
	Female	86	43.65

Source Table 1.4.1 of "Tables and Figures" document.

The majority of patients were male (56.35%) and the mean age at baseline was 62.25±9.94 years with a range between 34-80 years.

Summary of sitting systolic and diastolic blood pressure, sitting pulse rate and other vital signs are shown in Table 1.4.2 in the "Tables and Figures" document.

As foreseen from one inclusion criterion of study protocol, all patients had an ECOG Performance Status at baseline with Grade 0 or 1. In particular, the most part of patients reported a good ECOG Performance Status at baseline: 80.71% were fully active, able to carry out all pre-disease performance without restriction (Grade 0), while the rest of patients were restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, light housework, office work (Grade 1). For more details, see Table 1.4.3 in the "Tables and Figures" document.

Details about laboratory data at baseline are provided in Tables 1.4.7 and 1.4.8 in the "Tables and Figures" document; while in Tables 1.4.9 and 1.4.10 there are the information about Urinalysis and Serum pregnancy test.

Twenty-four patients (12.18%) had an abnormal interpretation of ECG, 78.17% a normal evaluation, while 19 patients (9.64%) did not perform the exam. At baseline, 17 patients (8.63%) had an abnormal interpretation of chest X-ray due to tumour involvement, 10 patients (5.08%) an abnormal interpretation due to other, 37.56% a normal evaluation, but nearly the half of patients did not perform this exam (48.73%). Further details are shown in Tables 1.4.11 and 1.4.12 in the "Tables and Figures" and "Patient Data Listing" documents.

With relation to baseline, 100 patients (50.76%) had at least one past medical condition and 67.51% (133 patients) had at least one active medical condition.



The most frequent active medical condition were: hypertension (41.12%), diabetes mellitus (10.15%) and smoking (9.14%). Further details about medical history are provided in Table 1.4.4a in the "Tables and Figures".

Many patients reported at least one past (21.32%) or active (38.58%) other condition different from those specified on the Medical history CRF page. The complete description of other specified medical condition by SOC and PT is presented in Table 1.4.4b in the "Tables and Figures".

Six patients (3.05%) took at least one prior therapy, 118 patients (59.90%) took at least one prior and concomitant therapy and almost all patients (97.97%) took at least one concomitant therapy. The complete description of taken medication by ATC CODE description - Level 2 and Preferred Term is presented in Table 1.4.16 in the "Tables and Figures" document.

Information about colorectal cancer primary disease history is displayed in the following Table 5.2-2.

**Table 5.2-2: Colorectal cancer primary disease history: age at diagnosis, status and location (Safety population).**

		(N=197)
Age at colorectal cancer diagnosis (years)	N	197
	Mean	61.20
	SD	9.93
	25th percentile	55.00
	Median	62.00
	75th percentile	69.00
	Min	33
	Max	80

		(N=197)	
		N	%
Status at first diagnosis	Locally advanced	13	6.60
	Metastatic	69	35.03
	Metastatic and locally advanced	41	20.81
	Primary operable	61	30.96
	Primary operable + Locally advanced	1	0.51
	Primary operable + Metastatic	9	4.57
	Primary operable + Metastatic and locally advanced	3	1.52

		(N=197)	
		N	%
Location	Left colon	37	18.78
	Left colon + Sigmoid colon	3	1.52
	Left colon + Trasverse colon	1	0.51
	Rectum	46	23.35
	Rectum + Left colon	1	0.51
	Rectum + Sigmoid colon	7	3.55
	Right colon	41	20.81
	Right colon + Left colon + Trasverse colon	1	0.51
	Right colon + Trasverse colon	1	0.51
	Sigmoid colon	52	26.40
	Trasverse colon	5	2.54
	Unknown	2	1.02

Source Table 1.4.5a of "Tables and Figures" document.

Each patient could have more than one primary tumour.

Each patient is counted only once in each row.

Mean age at colorectal cancer diagnosis was  $61.20 \pm 9.93$  years, with a minimum of 33 and a maximum of 80 years among patients included in the Safety population.

As regards first diagnosis, the main status were metastatic (35.03%), primary operable (30.96%) and metastatic and locally advanced (20.81%).

Principally, the location was at sigmoid colon (26.40%), at rectum (23.35%) and at colon – right (20.81%) or left (18.78%).

One hundred and fifty-two patients (77.16%) had a surgical resection of primary tumour, in average at  $61.18 \pm 10.15$  years (range between 33-80 years). Principally, the surgery of primary tumour was left (25.66%) and right (23.03%) hemicolectomy, low anterior resection (16.45%) and sigmoid colectomy (15.13%). All the type of surgery performed are presented in Table 1.4.5b in the "Tables and Figures" document.

Before the study enrolment, 21 patients (10.66%) received radiotherapy, 148 patients (75.13%) did neither previous adjuvant nor previous neo-adjuvant therapies. Table below showed the type of chemotherapy done:



**Table 5.2-3: Colorectal cancer primary disease history: chemotherapy (Safety population).**

			(N=197)	
			N	%
Type of Chemotherapy	None		148	75.13
	5-FU		3	1.52
	5-FU + Capecitabine		1	0.51
	5-FU + Irinotecan		1	0.51
	5-FU + Leucovorin		14	7.11
	5-FU + Leucovorin + Irinotecan		2	1.02
	5-FU + Leucovorin + Oxaliplatin		12	6.09
	5-FU + Other		1	0.51
	5-FU + Oxaliplatin		4	2.03
	5-FU + Oxaliplatin + Other		1	0.51
	Capecitabine		10	5.08

Source Table 1.4.5c of "Tables and Figures" document.

Further details about therapy concerning colorectal cancer primary disease history are reported in Table 1.4.5c in the "Tables and Figures" document.

The number of metastases for each patient is presented in the following Table 5.2-4:

**Table 5.2-4: Metastatic disease history (Safety population).**

		(N=197)	
		N	%
Nr. of metastases for each patient	1	105	53.30
	2	72	36.55
	3	19	9.64
	4	1	0.51

Source Table 1.4.6 of "Tables and Figures" document.

Patients had at least one metastasis mainly at liver (76.14%), lung (35.53%), lymph nodes (13.71%) and peritoneum (7.11%). Twenty-one patients (10.66%) had at least one surgical resection of metastasis. More details on site of metastases are reported in Table 1.4.6 in the "Tables and figures" document.

A description of number of target/non-target lesions and sum of longest diameter of all target lesions at baseline is shown in the following Table 5.2-5:

**Table 5.2-5: RECIST tumour assessment at baseline (Safety population).**

Target lesions		(N=197)	
		N	%
Number of lesions	1	64	32.49
	2	40	20.30
	3	39	19.80
	4	26	13.20
	5	14	7.11
	6	5	2.54
	7	4	2.03
	8	2	1.02
	9	1	0.51
	10	2	1.02

		(N=197)
Sum of longest diameters of all target lesions (mm)	N	197
	Mean	92.85
	SD	77.45
	25th percentile	37.00
	Median	70.00
	75th percentile	123.00
	Min	10
	Max	373

Non-Target lesions		(N=197)	
		N	%
Number of lesions	0	55	27.92
	1	72	36.55
	2	40	20.30
	3	17	8.63
	4	5	2.54
	5	4	2.03
	6	3	1.52
	8	1	0.51

Source Tables 1.4.14a-1.4.14b of "Tables and Figures" document.

All Safety patients had at least one target lesion: 32.49% had one target lesion, 20.30% two target lesions, 19.80% three target lesions, 13.20% four target lesions and 14.23% five or more target lesions. The maximum number of target lesions by patient was 10.



The locations of target lesions were various as reported in details in Table 1.4.14a in the "Tables and Figures" document. In particular, the major of target lesions (326 out of 538 target lesions, i.e. 60.59%) were localized at liver measured by CT scan; 14.87% and 8.36% of all target lesions were at lung and lymph node respectively, measured again by means of CT scan.

At baseline sum of longest diameter of all target lesions was on average  $92.85 \pm 77.45$  mm, with an interquartile range between 37.00 and 123.00 mm, a minimum of 10 and a maximum of 373 mm.

As regards non-target lesions, 55 patients (27.92%) had no-one and the 56.85% had one or two lesions. The maximum number of non-target lesions for patient was 8. The major of this type of lesions were localized again at lung, at liver and at lymph node (84, 76 and 53 out of 269 non-target lesions respectively) measured by means of CT scan.

More details about RECIST tumour assessment of target and non-target lesions are reported in Tables 1.4.14a and 1.4.14b in the "Tables and Figures" document.

One hundred and sixty-eight patients (85.28%) fulfilled the EQ-5D-3L questionnaire at baseline mainly by himself and 3 patients compiled some section of the questionnaire. Twenty-six patients (13.20%) did not entirely compiled the questionnaire principally due to logistic problems or mistakes.

As regards single dimension of health profile, 87 patients (44.16%) have reported problems concerning anxiety/depression, 32.99% in relation to pain/discomfort, 23.35% with performing usual activities, 9.14% in mobility and 7.61% with self-care.

On the EQ VAS, on average patients indicated their overall health state as  $72.65 \pm 17.49$ , with an interquartile range between 60.00 and 82.50, a minimum of 10 and a maximum of 100 (corresponding to the best state).

Considering each levels of health profile specified by the patient, the mean EQ-5D-3L raw index was  $78.85 \pm 15.43$ , with an interquartile range between 70.49 and 97.66, a minimum of 32 and a maximum of 98.

In Table 1.4.15 in the "Tables and Figures" document there are all the information about EQ-5D-3L questionnaire results at baseline.

### 5.3. MEASUREMENT OF TREATMENT COMPLIANCE

Compliance for the three study treatments were not computed due to several cases of dose reductions not quantified and check not possible on study treatment doses changes. See Paragraph 6.1 for summaries of number of cycles performed by patient for the three study treatments.

## 5.4. EFFICACY RESULTS

### 5.4.1. ANALYSIS OF EFFICACY

#### Primary efficacy parameter

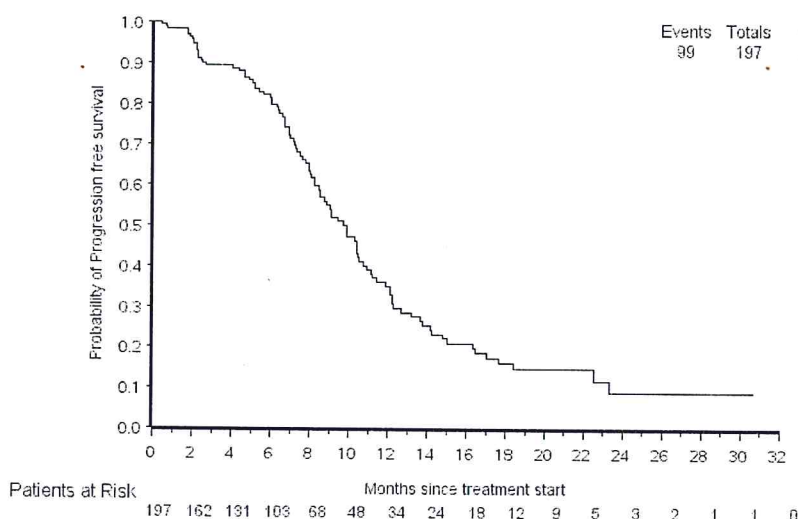
Relatively to the ITT population, 99 patients (50.25%) experimented during 1<sup>st</sup> line period the failure event defined as first progression disease or death due to any cause; the 49.75% patients could be considered censored, that is alive without a progression disease or patients without tumour assessment after baseline whose reason for treatment discontinuation of 1<sup>st</sup> line was not progression of disease or death. In particular, 6 patients died without having a documented progression of disease and 27 patients not failed without tumour assessment after baseline were censored at day 1 as they were alive at the end of premature discontinuation.

Description of PFS and its graphical representation created applying Kaplan-Meier methodology are provided in the following Table 5.4.1-1.

**Table 5.4.1-1: Progression Free Survival - Kaplan-Meier analysis (ITT population).**

PROGRESSION FREE SURVIVAL		(N=197)				
		N	%	Estimate	Confidence Limits	
					Lower	Upper
Numerousness of events	Failed	99	50.25	-	-	-
	Censored	98	49.75	-	-	-
Description of times (months)	25th percentile	-	-	6.69	5.93	7.44
	Median	-	-	9.70	8.43	10.49
	75th percentile	-	-	14.16	12.13	17.64





Source Tables 2.2.1-1 and Figure 2.2.1-1 of "Tables and Figures" document.

After 9.70 months since treatment start, 50% of patients at risk had a first PD or died due to any cause. In addition, the two-sided 95% confidence interval for Progression Free Survival median [8.43-10.49] did not include the 7 months Progression Free Survival median historic controls.

In the PP population, 52 patients (53.06%) experimented the event failure. Among these, 3 patients died without having a documented progression of disease. Eleven patients not failed without tumour assessment after baseline were censored at day 1 as they were alive at the end of premature discontinuation.

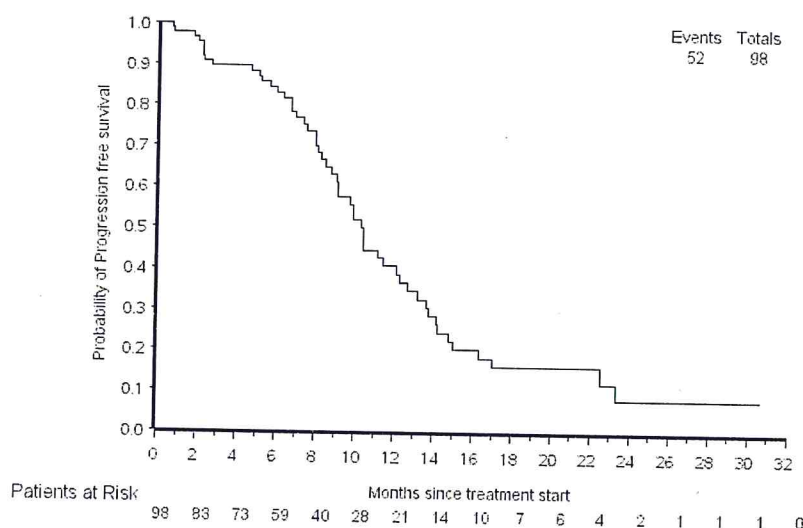
After 10.26 months since treatment start, the 50% of patients at risk had a first PD or have died due to any cause. Also in the PP population, the two-sided 95% confidence interval for PFS median [9.02-12.30] did not include the 7 months PFS median historic controls.

Compared to the ITT population, in this population there were no major difference as it can be observed in the following graph:

**Table 5.4.1-2: Progression Free Survival - Kaplan-Meier analysis (PP population).**

PROGRESSION FREE SURVIVAL		(N=98)				
		N	%	Estimate	Confidence Limits	
					Lower	Upper
Numerousness of events	Failed	52	53.06	-	-	-
	Censored	46	46.94	-	-	-

PROGRESSION FREE SURVIVAL		(N=98)				
		N	%	Estimate	Confidence Limits	
					Lower	Upper
Description of times (months)	25th percentile	-	-	7.44	5.57	8.43
	Median	-	-	10.26	9.02	12.30
	75th percentile	-	-	14.23	12.30	22.52



Source Tables 2.2.1-2 and Figure 2.2.1-2 of "Tables and Figures" document

Secondary efficacy parameter

**Overall Response Rate (ORR).** As regards ITT population, summary of overall lesion response based on RECIST criteria by scheduled visit is reported in Tables 2.3.1-1 in the "Tables and Figures" document. Descriptive statistics of the Best Overall Response achieved during 1<sup>st</sup> line phase (considering also unscheduled visits) based on RECIST criteria are presented in the following Table 5.4.1-3.

**Table 5.4.1-3: Summary of Best Overall Response (ITT population).**

	(N=197)	
	N	%
Best Overall Response		
CR	10	5.08
PR	87	44.16



## Statistical Report

**Sponsor: Roche**

**Protocol: ML 21380**

	(N=197)	
	N	%
SD	55	27.92
PD	13	6.60

Source Tables 2.3.2-1 of "Tables and Figures" document.

In the ITT population, 32 patients (16.24%) had no any tumour assessment after baseline. Principally, patients had a PR (44.16%) and 6.60% had a PD as best overall response. The following Table 5.4.1-4 reports the proportion of patients who achieved the objective response, i.e. a best overall response of CR or PR (overall response rate) calculated over all patients with at least one tumour assessment after baseline.

**Table 5.4.1-4: Overall Response Rate (ITT population with at least one tumor assessment after baseline).**

Overall Response Rate	(N=165)			
	N	%	Proportion	Exact 95% Two-sided Confidence Interval
Response achieved	97	58.79	0.59	[0.51-0.66]
Response not achieved	68	41.21		

Source Tables 2.3.3-1 of "Tables and Figures" document.

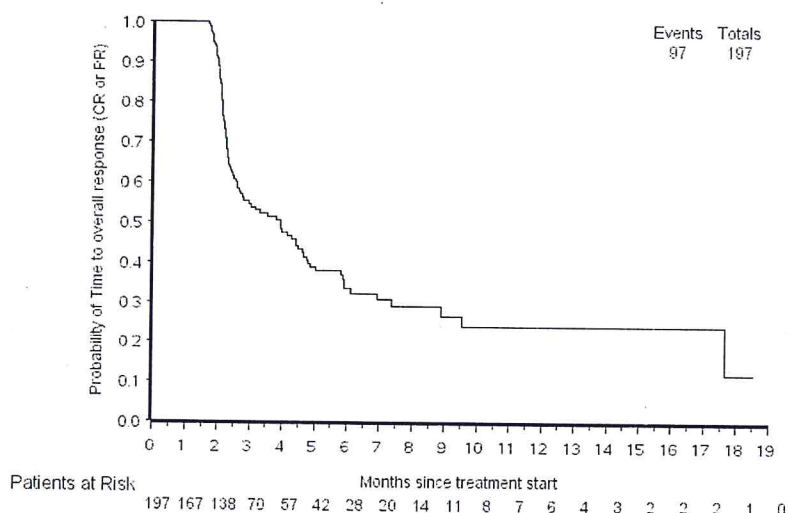
Among the evaluable 165 patients, 97 (58.79%) had the objective response achieved. The listing of overall tumor response and best overall response based on RECIST criteria at the end of treatment or premature discontinuation by each patient with at least one tumour assessment after baseline is shown in Table 2.3.4 in the "Tables and Figures" document.

Relatively to the PP population, the results of analyses were similar: a best overall response of CR or PR was observed in 64.29% of patients with at least one tumor assessment (54 out of 84 patients); principally the main best overall response was yet again PR. The complete description of overall lesion response analysis in the PP population is reported in Tables 2.3.1-2, 2.3.2-2 and 2.3.3-2 in the "Tables and Figures" document.

**Time to Response.** Concerning ITT population, any CR or PR as overall lesion response (based on RECIST criteria) during 1<sup>st</sup> line was not observed for 50.76% of patients. In particular, 27 patients not failed without tumour assessment after baseline whose reason for treatment discontinuation of 1<sup>st</sup> line was not progression of disease or death were censored at day 1. Results on time to response and its graphical representation created applying Kaplan-Meier methodology are provided in the following Table 5.4.1-5.

**Table 5.4.1-5: Time to overall response - Kaplan-Meier analysis (ITT population).**

TIME TO OVERALL RESPONSE (CR or PR)		(N=197)				
		N	%	Estimate	Confidence Limits	
					Lower	Upper
Numerousness of events	Failed	97	49.24	-	-	-
	Censored	100	50.76	-	-	-
Description of times (months)	25th percentile	-	-	2.13	2.07	2.23
	Median	-	-	3.93	2.56	4.66
	75th percentile	-	-	9.54	6.10	-



Source Tables 2.3.5-1 and Figure 2.3.5-1 of "Tables and Figures" document.

In the PP population, similar results were observed even if 55.10% of patients had a CR or PR as overall lesion response (based on RECIST criteria) during 1<sup>st</sup> line. Results on time to response for PP populations are displayed in Table 2.3.5-2 and Figure 2.3.5-2 in "Tables and Figures" document.

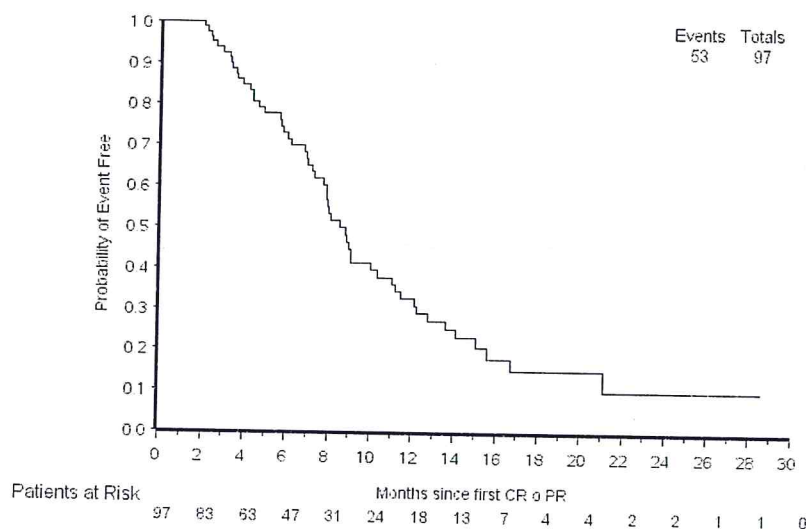
**Duration of overall response - Overall response (CR or PR).** Relatively to the ITT, duration of overall response (CR or PR) was calculated only for the 97 patients whose Best Overall Response was CR or PR based on RECIST criteria during 1<sup>st</sup> line of the study. Among these, 53 patients (54.64%) experimented during 1<sup>st</sup> line period a progression of disease (overall lesion response); 45.36% patients could be considered censored, that is alive without a progression disease objectively documented. No-one of these patients died during 1<sup>st</sup> line phase.



Description of the duration of overall response and its graphical representation created by means of Kaplan-Meier methodology are provided in the following Table 5.4.1-6.

**Table 5.4.1-6: Duration of Response: Overall response - Kaplan-Meier analysis (ITT population).**

OVERALL RESPONSE		(N=97)				
		N	%	Estimate	Confidence Limits	
					Lower	Upper
Numerousness of events	Failed	53	54.64	-	-	-
	Censored	44	45.36	-	-	-
Description of times (months)	25th percentile	-	-	5.64	3.84	7.18
	Median	-	-	8.52	7.28	10.33
	75th percentile	-	-	14.10	11.05	21.11



Source Tables 2.3.6-1a and Figure 2.3.6-1a of "Tables and Figures" document

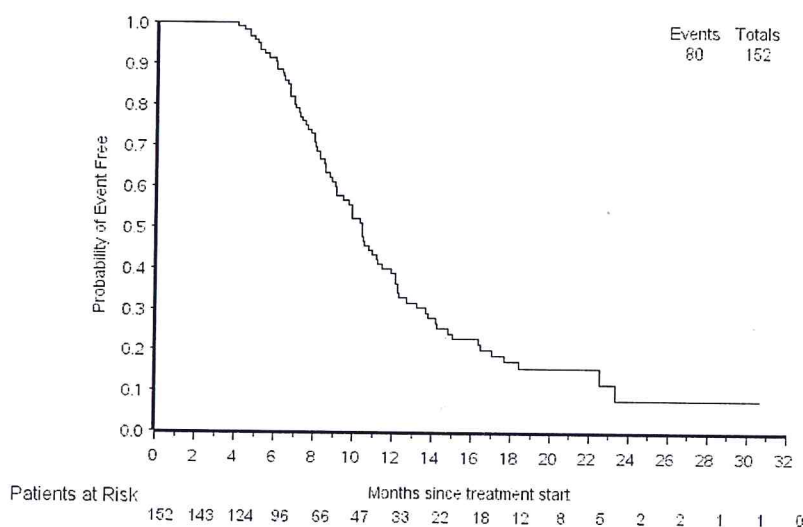
After 8.52 months since the date of RECIST assessment of 1<sup>st</sup> CR or PR as overall lesion response, 50% of patients at risk experienced a PD.

In the PP population, the duration of overall response was calculable only for 54 patients. Among these, 50.00% of patients had a PD and the median duration of overall response was slightly longer, i.e. 10.33 months. The complete description of the duration of overall response and Kaplan-Meier curve for the PP population are provided in the Table 2.3.6-2a and Figure 2.3.6-2a in the "Tables and Figures" document.

**Duration of Response - Stable disease (CR or PR or SD).** As regards the ITT population, duration of stable disease (CR or PR or SD) was calculated for the 152 patients whose Best Overall Response was CR or PR or SD based on RECIST criteria during 1<sup>st</sup> line of the study. Among these, 80 patients (52.63%) had a progression of disease (overall lesion response) during 1<sup>st</sup> line period; 47.37% patients could be considered censored, that is alive without a progression disease objectively documented. Only one of these patients (122748\_73) with SD as best overall response died during 1st line. However, death was reported as an unknown event so it was not considered as death due to underlying cancer and the patient was censored at the date of last RECIST assessment. Description of the duration of stable disease and its graphical representation created by means of Kaplan-Meier methodology are provided in the following Table 5.4.1-7.

**Table 5.4.1-7: Duration of Response: Stable response - Kaplan-Meier analysis (ITT population).**

STABLE DISEASE		(N=152)				
		N	%	Estimate	Confidence Limits	
					Lower	Upper
Numerousness of events	Failed	80	52.63	-	-	-
	Censored	72	47.37	-	-	-
Description of times (months)	25th percentile	-	-	7.44	6.69	8.20
	Median	-	-	10.39	9.02	11.44
	75th percentile	-	-	14.82	12.20	18.43



Source Tables 2.3.6-1b and Figure 2.3.6-1b of "Tables and Figures" document



The 50% of patients at risk had a PD after 10.39 months since treatment start.

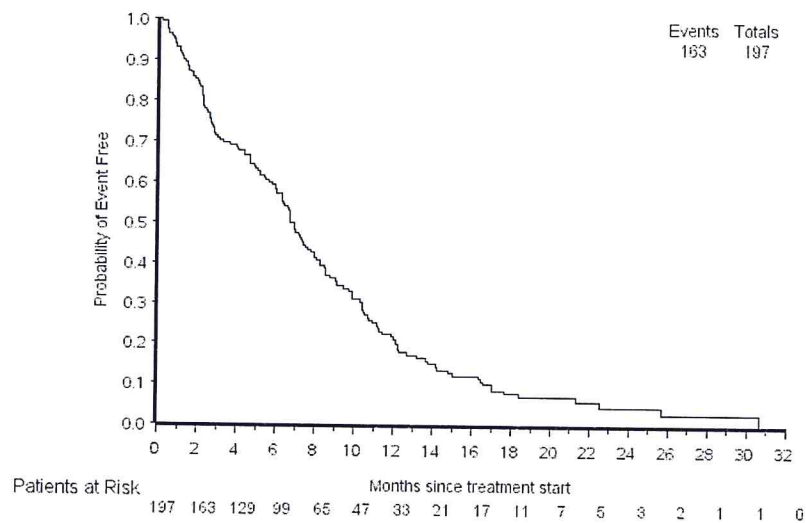
In the PP population, this analysis was performed on 78 patients and results were similar. The complete description of the duration of stable disease and Kaplan-Meier curve for the PP population are provided in the Table 2.3.6-2b and Figure 2.3.6-2b in the "Tables and Figures" document.

**Time to Treatment Failure (TTF).** As regards the ITT population, 163 patients (82.74%) experimented during the 1<sup>st</sup> line period the failure event defined as occurrence of discontinuation of treatment for any reason, including: death due to any cause, adverse event, insufficient therapeutic response (progression of disease), failure to return (lost to follow-up), refusing treatment (patients non-compliance), being unwilling to cooperate and withdrawing consent (patient withdrew consent). The Table 5.4.1-8 summarizes the result on TTF.

The median TTF was 6.69 months since treatment start.

**Table 5.4.1-8:** Time to treatment failure - Kaplan-Meier analysis (ITT population).

TIME TO TREATMENT FAILURE		(N=197)				
		N	%	Estimate	Confidence Limits	
					Lower	Upper
Numerousness of events	Failed	163	82.74	-	-	-
	Censored	34	17.26	-	-	-
Description of times (months)	25th percentile	-	-	2.59	2.16	3.90
	Median	-	-	6.69	5.97	7.74
	75th percentile	-	-	11.15	9.87	12.20



Source Tables 2.3.7-1 and Figure 2.3.7-1 of "Tables and Figures" document

In the PP population, 83.67% patients experimented the event failure and the median time to treatment failure was 7.90 months. Descriptions of TTF are reported in the Table 2.3.7-2 and Figure 2.3.7-2 in the "Tables and Figures" document.

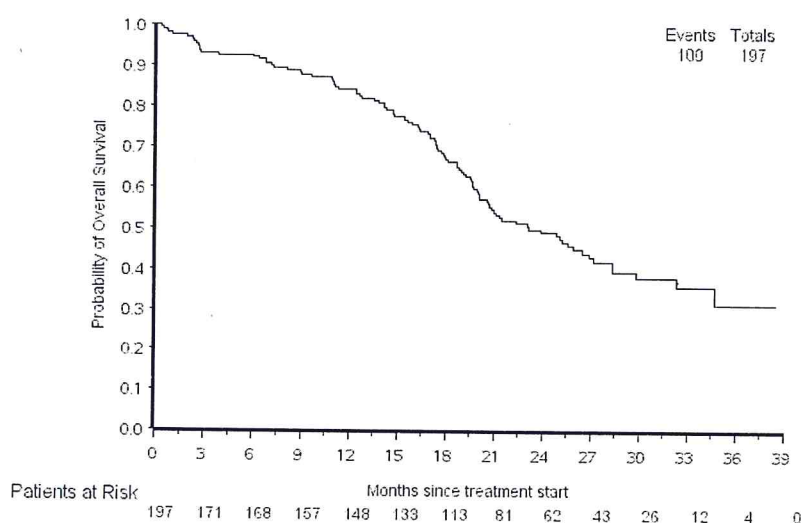
**Overall survival (OS).** In the ITT population, the 50.76% of patients experimented during 1<sup>st</sup> line period and follow-up the failure event defined as occurrence of death due to any cause, while 97 patients (49.24%) could be considered censored, that is alive at last available follow-up visit.

Description of OS and its graphical representation created by means of Kaplan-Meier methodology are provided in the following Table 5.4.1-9.

**Table 5.4.1-9: Overall Survival - Kaplan-Meier analysis (ITT population).**

OVERALL SURVIVAL		(N=197)				
		N	%	Estimate	Confidence Limits	
					Lower	Upper
Numerousness of events	Failed	100	50.76	-	-	-
	Censored	97	49.24	-	-	-
Description of times (months)	25th percentile	-	-	16.30	13.48	17.70
	Median	-	-	23.15	20.07	27.15
	75th percentile	-	-	-	-	-





Source Tables 2.3.8-1 and Figure 2.3.8-1 of "Tables and Figures" document

During 1<sup>st</sup> line period and follow-up, there were no sufficient events to estimate 75<sup>th</sup> percentile and its confidence limits. At about 23 months after start of treatment 50% of patients at risk in the ITT population had died.

Description of OS and its graphical representation performed with Kaplan-Meier methodology in PP population are provided in Table 2.3.8-2 and Figure 2.3.8-2 in the "Tables and Figures" document.

**R0 resectability rate.** During the study, 47 patients had one surgery and 5 two surgical procedure, for a total of 52 ITT patients (26.40%).

The reason and the residual disease for patients who performed any surgery are described jointly in the following Table 5.4.1-10.

**Table 5.4.1-10:** Summary of number of patients by reason for surgery and residual diseases (ITT patients who performed surgery during the study).

		(N=52)	
Reason for surgery	Residual disease	N	%
Curative	No (Radical surgery)	29	55.77
	Yes	7	13.46
	Unknown	2	3.85
	Not Applicable	4	7.69
Palliative	No (Radical surgery)	2	3.85
	Yes	4	7.69
	Unknown	2	3.85
	Not Applicable	1	1.92
Biopsy	Yes	1	1.92
	Not Applicable	1	1.92
Unknown	Unknown	1	1.92
	Not Applicable	2	3.85
Other	Yes	1	1.92
	Not Applicable	2	3.85

Source Tables 2.3.9-1 of "Tables and Figures" document.

Patients are counted only once in each row.

Each patient could have had more than one surgery also with the same reason and residual disease during the study period.

Overall, 31 patients (59.62%) out of 52 who performed any surgery had obtained at least one radical procedure.

The description of surgery analysis in the PP population is reported in Tables 2.3.9-2 in the "Tables and Figures" document.

**K-Ras and B-Raf mutation status.** Tumor sample was evaluated by the Central Laboratory only for 34 enrolled patients. All these patients were eligible for ITT population.

Among these, overall 16 patients (47.06%) had at least one gene alteration (K-Ras or B-Raf mutation). In details, 12 patients (35.29%) had a K-Ras mutation in codon 12 or 13 (see Table 2.3.10-1 in the "Table and Figures" document for details about type), 1 patient (2.94%) had a K-Ras mutation in codon 61 of type Q61H, 1 patient (2.94%) had a K-Ras mutation in codon 146 of type A146T and 3 patients (8.82%) had a V600E B-Raf mutation. In particular, one patient ('122701\_42') had both a G12D K-Ras mutation and a V600E B-Raf mutation.

Compared to wild type patients, patients with K-Ras or B-Raf mutation had a lower overall response rate: 66.67% (10 out of 15) vs. 88.89% (16 out of 18). Indeed, the estimate RR (0.75) suggested that the probability of achieving a CR or PR as best overall response during 1<sup>st</sup> line is lower for patients with gene alterations than patients without gene alterations (more exactly 0.75 times minor). However RR was not statistically significant

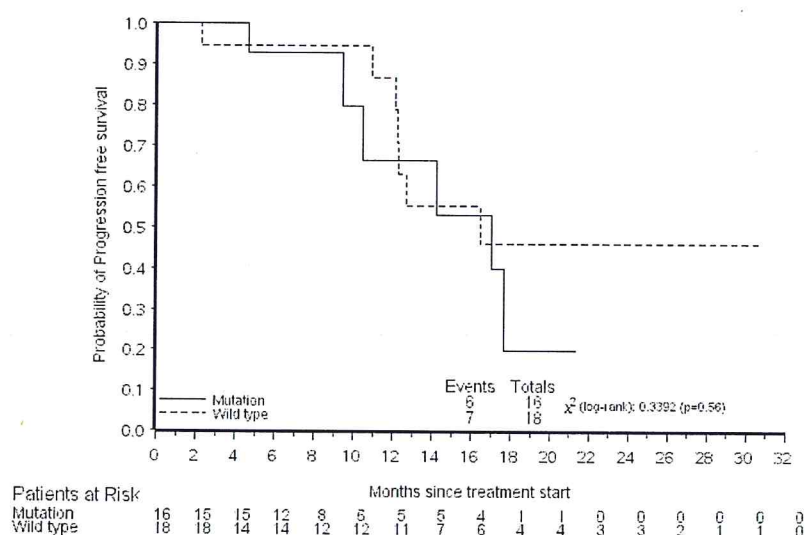


(95% CI: [0.51-1.11], p-value=0.2028). One patient ('122729\_126') with a G12C K-Ras mutation did not perform any tumour assessment after baseline; for that, it was not considered for this analysis.

No statistically significant difference between patients with K-Ras or B-Raf mutation vs. wild type patients in terms of Progression Free Survival was detected (Log-rank test p-value=0.56). The median PFS in patient with gene alteration was 17.02 months and in patients without gene alteration 16.43 months and the incidence of failure event was similar. Moreover during 1<sup>st</sup> line period there were sufficient events to estimate 75<sup>th</sup> percentile only for patients with gene alterations.

Kaplan-Meier curves for PFS are represented in the following Figure 5.4.1-1.

**Figure 5.4.1-1:** Kaplan Meier curve of Progression Free Survival by K-Ras and B-Raf mutation status (ITT patients with tumor sample evaluated by the Central Laboratory)

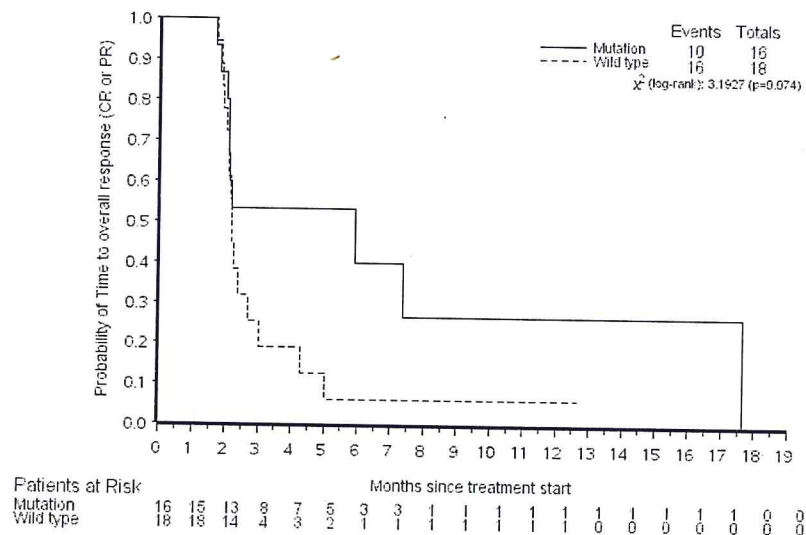


Source Figure 2.3.12-1 of "Tables and Figures" document

A difference was noted in the median time to response: 5.90 months for patients with K-Ras or B-Raf mutation and 2.16 months for patients wild type; although the difference was not statistically significant (Log-rank test p-value=0.074). In addition, the incidence of event (i.e. CR or PR as overall lesion response during 1<sup>st</sup> line) was higher in wild type patients (16 out of 18, i.e. 88.89%) than in patients with gene alterations (10 out of 16, i.e. 62.50%).

Kaplan-Meier curves for time to response are represented in the following Figure 5.4.1-2.

**Figure 5.4.1-2: Kaplan Meier curve of time to response by K-Ras and B-Raf mutation status (ITT patients with tumor sample evaluated by the Central Laboratory)**



Source Figure 2.3.13-1 of "Tables and Figures" document

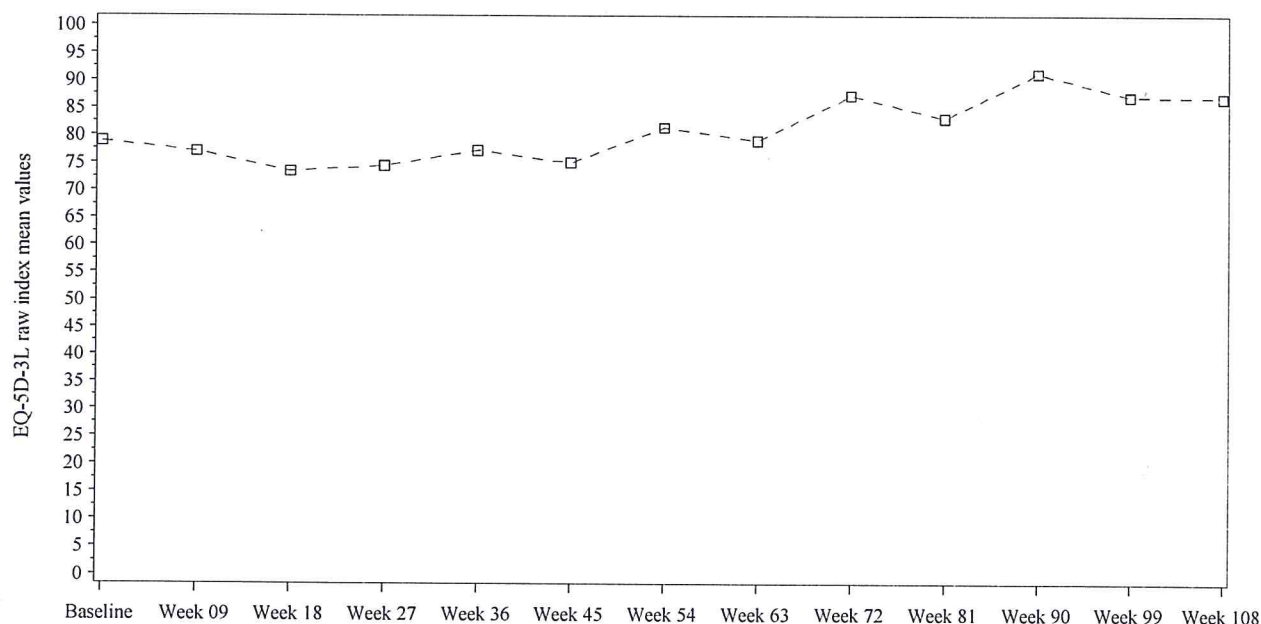
Further details about results of analysis of ORR, PFS and time to response by K-Ras and B-Raf mutation status in ITT population are reported in Tables and Figures 2.3.11-1, 2.3.12-1 and 2.3.13-1 in the "Tables and Figures" document.

Only for 17 patients in the PP population the tumour sample was evaluated by the Central Laboratory and among these, 7 patients (41.18%) had at least one gene alteration (K-Ras or B-Raf). See Table 2.3.10-2 for summary of K-Ras and B-Raf mutation status, Tables and Figures 2.3.11-2, 2.3.12-2 and 2.3.13-2 in the "Tables and Figures" document for analysis of ORR, PFS and time to response by K-Ras and B-Raf mutation status in PP population.

**Quality of life.** Summary statistics and graphical representation of EQ-5D-3L answers by visit according to prevalence approach in ITT population are reported in Tables and Figure 2.3.14-1a. A graphical representation of the EQ-5D-3L raw index mean values by visit is displayed in the following Figure 5.4.1-3:



**Figure 5.4.1-3: EQ-5D-3L raw index: summary by visits - Prevalence approach (ITT population).**



Source Figure 2.3.15-1a of "Tables and Figures" document

A slightly tendency to increasing is observed in the graph above. However, as it can be observed in Table 2.3.15-1a in "Table and Figures" document, after visits week 54 the number of observed patients was minor than ten.

Changes vs. baseline of EQ-5D-3L raw index were calculable only for patients with both baseline and last visit score not missing, i.e. 114 patients. A summary of EQ-5D-3L raw index at baseline, at last visit with questionnaire compiled by each patient and of their changes vs. baseline according to prevalence approach are reported in Table 5.4.1-11.

**Table 5.4.1-11: EQ-5D-3L raw index: changes end of treatment versus baseline - Prevalence approach (ITT population).**

EQ-5D-3L raw index	(N=197)		
	Baseline	Last visit with questionnaire compiled	Absolute change vs baseline
N	114	114	114
Mean	80.24	74.94	-5.30
SD	14.32	19.08	19.13
25th percentile	70.49	62.40	-18.37
Median	78.64	77.32	0.00
75th percentile	97.66	97.66	1.85
Min	34	29	-51
Max	98	98	43

Source Table 2.3.16-1a of "Tables and Figures" document

In average, EQ-5D-3L raw index was  $80.24 \pm 14.32$  at baseline and  $74.94 \pm 19.08$  at last visit with questionnaire compiled by each patient. The absolute changes vs. baseline was negative in average:  $-5.30 \pm 19.13$ , with a range between -51 and 43. The decreasing evaluation of overall health during the study was stressed by the p-values of Signed-rank test performed on changes values (p-value= 0.0076).

The EQ-5D-3L raw index according to LOCF approach at baseline and last visit with questionnaire compiled by each patient were also explored. In particular, LOCF approach could be applied only for patients who fulfilled the questionnaire at baseline, i.e. 168. Among these, absolute changes vs. baseline were calculable only for 142 patients. The Table 5.4.1-12 shows descriptive statistics about changes vs. baseline according to LOCF approach.

**Table 5.4.1-12:** EQ-5D-3L raw index: changes end of treatment versus baseline - LOCF approach (ITT patients with EQ-5D questionnaire fulfilled at baseline).

EQ-5D-3L raw index*	(N=168)		
	Baseline	Last visit with questionnaire compiled	Absolute change vs baseline
N	142	142	142
Mean	80.23	76.27	-3.96
SD	14.40	18.58	17.13
25th percentile	70.49	65.71	-10.46
Median	78.64	78.30	0.00
75th percentile	97.66	97.66	0.00
Min	34	29	-51
Max	98	98	43

Source Table 2.3.16-1b of "Tables and Figures" document

No main difference compared to prevalence approach was observed. Indeed, the decreasing evaluation of overall health during the study (in average  $-3.96 \pm 17.13$ ) was stressed again by the p-values of Signed-rank test performed on changes values (p-value= 0.0131).

There is only a difference as regard 75<sup>th</sup> percentile of absolute change since many patients fulfilled the questionnaire only at baseline.

Further information about quality of life analyses according to LOCF approach is provided in Tables and Figures 2.3.14-1b and 2.3.15-1b in the "Tables and Figures" document.

Analyses concerning questionnaire EQ-5D-3L in PP population according to prevalence and LOCF approaches are provided in Tables 2.3.14-2a, 2.3.15-2a, 2.3.16-2a and 2.3.14-2b, 2.3.15-2b, 2.3.16-2b respectively in the "Tables and Figures" document. Overall, similar results were obtained.



### 5.4.2. STATISTICAL/ANALYTICAL ISSUES

On CRF, date of treatment discontinuation of 1<sup>st</sup> line was not collected. For analyses, this date was identified as follows depending on the reason for discontinuation occurred in this study:

- 'Death' or 'Progression of disease': date of death or date of PD respectively collected on section 'End of treatment or premature discontinuation' on CRF;
- 'Protocol violation', 'Patient withdrew consent', 'Patient non-compliance', 'Need for surgery', 'Medical decision': the maximum between date of last 1<sup>st</sup> line visit, date of last treatment taken, date of last sample collected for laboratory exam and date of last tumour assessment;
- 'Adverse event(s)': identifying last date of 1<sup>st</sup> line as before (the maximum between date of last 1<sup>st</sup> line visit, date of last treatment taken, date of last sample collected for laboratory exam and date of last tumour assessment), there were two cases:
  - if last date of 1<sup>st</sup> line was later or equal than start date of AE with action taken 'study treatment permanently discontinued', then last date of 1<sup>st</sup> line was considered;
  - otherwise, start date of AE was used.

If patient had more than one AE with action taken 'study treatment permanently discontinued', the last AE occurred was evaluated.

For overall survival analysis, in order to identifying the occurrence of death for any patient after 1<sup>st</sup> line discontinuation, the presence of date of death collected in section 'Follow-up' of CRF or in section 'Adverse events' of CRF was checked. Instead, if a patient was not died, time was censored at the last date defined as the maximum between following dates: date of 1<sup>st</sup> line (defined as mentioned before) and date of last available follow-up visit in which patient was known to be alive.

As regards progression of disease, the first date between the date of PD as overall lesion tumour response (i.e., the date of RECIST tumour assessment during 1<sup>st</sup> line) and the date of PD as reason for premature discontinuation was considered as date of first progression of disease. Instead, for duration of overall response, only tumour evaluation based on RECIST criteria was used to detect the occurrence of progression of disease.

The potential predictive value of K-Ras and B-Raf gene alterations in patients with colorectal cancer on the overall response rate was tested by means of Fisher exact test instead of Chi-square test since one or more expected frequencies were of five or less.

The 95% confidence limits was calculated for Relative Risk.

In order to evaluate quality of life, the EQ-5D-3L descriptive system is constituted by 5 dimensions: mobility, self-care, usual activity, pain/discomfort and anxiety. Each dimension has 3 levels of perceived problems: no problem (level 1), some problems (level 2) and extreme problems (level 3).

A unique respondent's health state is defined by combining one level from each of the 5 dimensions.

EQ-5D health states may be converted into a single summary index by applying a formula that attaches weights (called values sets) to each level in each dimension, in order to represent health status as a continuous variable.

In details, for this study EQ-5D health states were converted into EQ-5D-3L raw index value by applying the scoring algorithm based on the European EQ-net VAS set. The raw index was chosen instead of rescaled index since the questionnaire was used in order to obtain a quality of life assessment.

The EQ VAS presents on EQ-5D-3L questionnaire, records the global respondent's self-rated health on a vertical visual analogue scale where the endpoints are labelled 'Best imaginable health state' (i.e. 100) and 'Worst imaginable health state' (i.e. 0). These data can be used as a quantitative measure of overall health outcome as judged by the individual respondents.

The EQ-5D-3L raw index can be interpreted as EQ VAS (ranges are similar), but for its computation it's a more specific description of perceived patient's health state by considering each dimension. For that, analyses were focused on EQ-5D-3L raw index.

The EQ-5D-3L raw index was calculable only if patient had replied to all EQ-5D descriptive system's five dimensions. Since many patients not fulfilled the questionnaire after baseline visit, LOCF approach was adopted in order to reduce the impact of missing values on quality of life analysis.

### 5.4.3. EFFICACY CONCLUSION

As described on Paragraph 3.6.2, the sample size chosen for this exploratory study was based on the confidence interval estimate for median Progression Free Survival (primary efficacy analysis). In ITT population, the 50% of patients at risk had a first PD or died due to any cause at 9.70 months since treatment start, with the two-sided 95% confidence interval for PFS median corresponding to [8.43-10.49]. The evidence of improvement for Progression Free Survival can be provide since the 7 months PFS median historic controls was not included in the two-sided 95% confidence interval for PFS median.

These results were supported on PP population: the Progression Free Survival median was 10.26 months and the two-sided 95% confidence interval was equal to [9.02-12.30].

Among the 165 ITT patients with at least one tumour assessment after baseline, 97 (58.79%) had a CR or PR as best overall response based on RECIST criteria during 1<sup>st</sup> line period (*Overall response rate endpoint*). The exact 95% two-sided CI was [0.51-0.66]. Similar results were observed in the PP population.

Relatively to the ITT population, after about 4 months from the 1<sup>st</sup> day of study treatment, 50% of patients at risk had a CR or a PR as overall lesion response based on RECIST criteria during 1<sup>st</sup> line period (*Time to response endpoint*).

In the ITT population, after about 6 months and half from treatment start, 50% of patients at risk have discontinued study treatment for death due to any cause or adverse event or insufficient therapeutic response (progression of disease) or failure to return



(lost to follow-up) or refusing treatment (patients non-compliance) or being unwilling to cooperate and withdrawing consent (patient withdrew consent) (*TTF endpoint*).

Among the 152 ITT patients whose Best Overall Response was CR or PR or SD based on RECIST criteria during 1<sup>st</sup> line period, 80 patients (52.63%) had a progression of disease during 1<sup>st</sup> line period and the median duration of stable disease was about 10 months since treatment start (*Duration of Response - Stable disease*).

Instead, among the 97 ITT patients whose Best Overall Response was CR or PR based on RECIST criteria during 1<sup>st</sup> line period, 53 patients (54.64%) experienced during 1<sup>st</sup> line period a progression of disease and the median duration of overall response was 8 months and half after the date of RECIST assessment of first CR or PR as overall lesion response (*Duration of overall response*).

Considering all the study period, 50% of patient at risk had died due to any cause at about 23 months after 1<sup>st</sup> day of study treatment (*OS endpoint*).

Concerning all these survival analyses, no main differences were observed in the PP population. Only considering OS endpoint, the median overall survival was slightly later, i.e. 26.46 months after treatment start.

During the study, among the 52 ITT patients who performed at least one surgery, overall 31 patients had obtained at least one radical procedure.

Tumour sample was evaluated by the Central Laboratory only for 34 enrolled patients, all eligible for ITT population. Among these, 16 patients had at least one K-Ras or B-Raf mutation.

The exploratory analysis indicates a proportion of CR and PR as Best overall response during 1<sup>st</sup> line period lower in patients with gene alterations (10 out of 15, i.e. 66.67%) than in wild type patient (16 out of 18, i.e. 88.89%). The minor probability of achieving a CR or PR as best overall response for patients with K-Ras or B-Raf mutation was confirmed by RR estimate, i.e. 0.75. Nonetheless, RR was not statistically significant (95% CI: [0.51-1.11], p-value= 0.2028).

A no statistically significant difference between patients with K-Ras or B-Raf mutation vs. wild type patients in terms of PFS during 1<sup>st</sup> line was detected (p-value=0.56). The median PFS were similar: 17.02 months for patient with gene alteration and 16.43 months for patients without gene alteration. The incidence of PD or death due to any cause was homogeneous between the two groups.

More wild-type patients had a CR or PR as overall lesion response during 1<sup>st</sup> line than patients with gene alterations. Moreover, the median time to achieve the event is shorter: 2.16 months for patients wild type vs. 5.90 months for patients with K-Ras or B-Raf mutation. This difference can be observed in the Kaplan-Meier plots, but it was not statistically significant (p-value=0.074).

Nevertheless, for the very lower number of tumour sample evaluated, all these secondary efficacy results must be read with a descriptive meaning.

During 1<sup>st</sup> line period, compared to baseline, at last visit with questionnaire compiled by each patient a decreasing on the judgment on overall health (measured by means of EQ-5D-3L raw index values) was observed: the mean absolute change was

-5.30±19.13 according to Prevalence approach (p-value= 0.0076) and -3.96±17.13 according to LOCF approach (p-value= 0.0131). However, the evaluation on quality of life perceived by patients was good.

## 6. SAFETY EVALUATION

### 6.1. EXTENT OF EXPOSURE

The following Table 6.1-1 reports the duration of the study period distinctly for 1<sup>st</sup> line period and follow-up time: duration of 1<sup>st</sup> line period was defined as the time in days elapsed from the date of the baseline visit until the date of the discontinuation identified as last day of 1<sup>st</sup> line; follow-up time was defined as the time in days elapsed from the date of discontinuation of 1<sup>st</sup> line treatment until the date of death after discontinuation or the date of last available follow-up.

**Table 6.1-1:** Duration of the study period (Safety population).

		(N=197)
First line period (days)	N	197
	Mean	223.65
	SD	175.93
	25th percentile	79.00
	Median	191.00
	75th percentile	309.00
	Min	10
	Max	936
Follow-up time (days)	N	178
	Mean	387.74
	SD	238.79
	25th percentile	195.00
	Median	386.50
	75th percentile	547.00
	Min	1
	Max	1085

Source Tables 2.1.1 of "Tables and Figures" document

The mean duration of 1<sup>st</sup> line period was 223.65±175.93 days (interquartile range: 79-309 days) and in average patients were observed during follow-up for 387.74±238.79 days (interquartile range: 195-547 days). The follow-up time was not calculable for 19 patients of safety population: 6 died during 1<sup>st</sup> line, 12 discontinued 1<sup>st</sup> line for withdrew consent and no follow-up visit was collected for patient '122719\_125' who discontinued 1<sup>st</sup> line for AE.



The summary statistics of the number of cycles performed by patient for the three study treatments and the number of patients with at least a cycle with a dose modification/interrupted by reason during 1<sup>st</sup> line treatment phase are shown in the following Table 6.1-2:

**Table 6.1-2: Treatment exposure (Safety population)**

Number of cycles performed*	(N=197)		
	Bevacizumab	Capecitabine	Oxaliplatin
N	197	197	197
Mean	9.16	5.94	5.85
SD	7.27	2.69	2.62
25th percentile	3.00	3.00	3.00
Median	8.00	7.00	7.00
75th percentile	12.00	8.00	8.00
Min	1	1	1
Max	39	13	13

	(N=197)					
	Bevacizumab		Capecitabine		Oxaliplatin	
	N	%	N	%	N	%
Number of patients without any dose modification/interruption	44	22.34	68	34.52	58	29.44
Number of patients with at least a cycle with a dose modification/interruption	153	77.66	129	65.48	139	70.56
Number of patients with at least a cycle with a dose modification/interruption due to AE	98	49.75	98	49.75	90	45.69
Number of patients with at least a cycle with a dose modification/interruption due to refusal	3	1.52	-	-	-	-
Number of patients with at least a cycle with a dose modification/interruption due to non-compliance	6	3.05	7	3.55	1	0.51

Source Tables 2.1.2 of "Tables and Figures" document

\*Cycles with a dose modification/interruption signaled with a dose different from zero specified were considered as cycles performed.

During the study period, 8 patients (4.06%) received radiotherapy: 5 patients performed one cycle and 3 patients two cycles. Site of radiotherapy was hepatic (3 patients), lung (1 patient) and another site different from hepatic, local/regional and lung (4 patients). Further details about radiotherapy are present in Table 1.4.17 in the "Tables and Figures" document and in "Patient Data Listings" document.

## 6.2. ADVERSE EVENT

A summary of all adverse events occurred during the study period and follow-up (any new AE occurred up to 28 days after the last dose of study treatment, any related non-serious

new AE occurred up to 6 months after the last dose of study drug and any related SAE had to be collected on CRF) is displayed in the following Table 6.2-1.

**Table 6.2-1: Summary of patients with adverse events (Safety population).**

	N	%
Patients evaluable for AEs	197	-
Number of AEs	1877	-
Number of distinct AEs per patient*	1253	-
Patients with AEs	186	94.42
Patients with SAEs	56	28.43
Patients with severe AEs **	104	52.79
Patients with a drug related AE	179	90.86
Patients with AEs related to CT	172	87.31
Patients with AEs related to Bevacizumab	61	30.96
Patients with AEs related to the combination of CT and Bevacizumab	16	8.12
Patients with AEs unknown related	45	22.84
Patients with at least one study treatment discontinued due to AEs	51	25.89
Patients with treatment adjusted/temporary interrupted due to AEs	106	53.81
Patients who died due to AEs	12	6.09

Source Tables 3.1.2 of "Tables and Figures" document

Except for the number of AEs, patients are counted only once in each row.

Each patient could experience more than one AE.

\*Distinct number of AEs is calculated by patient and Low Level Term.

\*\*Toxicity grade 3-4 AEs were considered severe.

The percentage of patients who experienced at least one adverse event was 94.42%; 11 patients had no-one AE.

In terms of SOC, the majority of occurred adverse events (with relative frequency higher than 30%) were classified as "Gastrointestinal disorders" (74.62%), "General disorders and administration site conditions" (58.88%), "Nervous system disorders" (56.85%) and "Vascular disorders" (31.98%). The complete description of number of patients with AE by SOC is reported in Table 3.1.3 in the "Table and Figures" document.

More specifically, adverse events happened with relative frequency higher than 5% are reported in Table 6.2-2 by SOC and PT. All adverse events signaled on CRF during the study period are shown in Table 3.1.4 in the "Tables and Figures" document.



**Table 6.2-2:** Number of patients with adverse events by SOC and PT, whether or not study drug related (Safety population).

		(N=197)	
SYSTEM ORGAN CLASS	PREFERRED TERM	N	%
Blood and lymphatic system disorders	Anaemia	12	6.09
	Leukopenia	13	6.60
	Neutropenia	32	16.24
	Thrombocytopenia	26	13.20
Gastrointestinal disorders	Abdominal pain	29	14.72
	Abdominal pain upper	12	6.09
	Constipation	28	14.21
	Diarrhoea	75	38.07
	Nausea	84	42.64
	Stomatitis	10	5.08
	Vomiting	48	24.37
General disorders and administration site conditions	Asthenia	60	30.46
	Fatigue	29	14.72
	Mucosal inflammation	21	10.66
	Pyrexia	41	20.81
Metabolism and nutrition disorders	Anorexia	18	9.14
Musculoskeletal and connective tissue disorders	Musculoskeletal pain	10	5.08
	Pain in extremity	17	8.63
Nervous system disorders	Headache	10	5.08
	Neuropathy peripheral	38	19.29
	Neurotoxicity	12	6.09
	Paraesthesia	59	29.95
Renal and urinary disorders	Proteinuria	17	8.63
Respiratory, thoracic and mediastinal disorders	Dyspnoea	13	6.60
	Epistaxis	14	7.11
Skin and subcutaneous tissue disorders	Palmar-plantar erythrodysesthesia syndrome	18	9.14
Vascular disorders	Hypertension	45	22.84

Source Tables 3.1.4 of "Tables and Figures" document

Overall, 87.31% of patients had at least one AE suspect related to chemotherapy. The more frequent AEs were classified as "Diarrhoea" (35.53%), "Nausea" (39.59%), "Vomiting" (21.83%), "Asthenia" (26.40%) and "Paraesthesia" (28.93%) mainly with a toxicity grade equal to 1.

As regards AEs of grade 4 suspect related to CT, 2 patients had "Diarrhoea", 1 patients "Gastrointestinal disorders", 1 patients "Vomiting" and 2 patients "Asthenia". The complete description of AEs related to chemotherapy is shown in Tables 3.1.5a and 3.1.6a in the "Tables and Figures" document.

Sixty-one patients (30.96%) experienced at least one AE suspect related to Bevacizumab. In details, all AEs suspect related to Bevacizumab occurred in more than one patient are reported in the Table 6.2-3. Concerning AEs of grade 4: 1 patient had "Intestinal perforation", 2 patients "Pulmonary embolism", 1 patient had "Hypertension" and 1 patient had "Hypertensive crisis". The complete description of number of patients with AE suspect related to Bevacizumab is shown in Table 3.1.5b and 3.1.6b in the "Table and Figures" document.

**Table 6.2-3:** Number of patients with adverse events by SOC and PT, suspect related to Bevacizumab (Safety population).

		(N=197)	
SYSTEM ORGAN CLASS	PREFERRED TERM	N	%
Gastrointestinal disorders	Gingival bleeding	2	1.02
	Intestinal perforation	2	1.02
	Rectal haemorrhage	2	1.02
General disorders and administration site conditions	Asthenia	2	1.02
Renal and urinary disorders	Proteinuria	13	6.60
Respiratory, thoracic and mediastinal disorders	Epistaxis	7	3.55
	Pulmonary embolism	5	2.54
Vascular disorders	Deep vein thrombosis	2	1.02
	Hypertension	31	15.74
	Hypertensive crisis	4	2.03

Source Tables 3.1.5b of "Tables and Figures" document

Patients are counted only once in each row.

Each patient could experience more than one adverse event.

Sixteen patients (8.12%) experienced at least one AE suspect related to the combination of CT and Bevacizumab. In details, AEs suspect related to the combination of CT and Bevacizumab occurred in more than one patient are presented in the Table 6.2-4. Moreover, one patient had "Cardio-respiratory arrest" with toxicity grade 4 suspect related to the combination of CT and Bevacizumab. The complete description of number of patients with AE suspect related to the combination of CT and Bevacizumab is shown in Table 3.1.5c and 3.1.6c in the "Table and Figures" document.



**Table 6.2-4:** Number of patients with adverse events by SOC and PT, suspect related to the combination of CT and Bevacizumab (Safety population).

		(N=197)	
SYSTEM ORGAN CLASS	PREFERRED TERM	N	%
Gastrointestinal disorders	Nausea	3	1.52
	Vomiting	2	1.02
General disorders and administration site conditions	Asthenia	2	1.02
	Fatigue	2	1.02
Vascular disorders	Hypertension	2	1.02

Source Tables 3.1.5c of "Tables and Figures" document

Patients are counted only once in each row.

Each patient could experience more than one adverse event.

A summary of SAEs collected on CRF is shown in the following Table 6.2-5.

**Table 6.2-5:** Number of patients with serious adverse events (Safety population).

	N	%
Patients evaluable for AEs	197	-
Number of SAEs	84	-
Number of distinct SAEs per patient*	82	-
Patients with SAEs	56	28.43
Patients with severe SAEs**	50	25.38
Patients with a drug related SAE	32	16.24
Patients with SAEs related to CT	18	9.14
Patients with SAEs related to Bevacizumab	14	7.11
Patients with SAEs related to the combination of CT and Bevacizumab	4	2.03
Patients with SAEs unknown related	3	1.52
Patients with with at least one study treatment discontinued due to SAEs	29	14.72
Patients with treatment adjusted/temporarily interrupted due to SAEs	17	8.63
Patients who died due to SAEs	12	6.09

Source Tables 3.2.1 of "Tables and Figures" document

Except for the number of SAE, patients are counted only once in each row.

Each patient could experience more than one SAE

\*Distinct number of SAEs is calculated by patient and Low Level Term.

\*\*Toxicity grade 3-4 SAEs were considered severe.

Fifty-six patients (28.43%) had at least one serious adverse event and the total number of SAEs collected during the study was 84. Twenty-nine patients (14.72%) permanently interrupted at least one study treatments due to SAE. The listing of patients with SAEs is displayed in Table 3.2.2 in the "Table and Figures" document.

The number of patients with SAEs by SOC and PT occurred in more than one patient are reported in the Table 6.2-6.

**Table 6.2-6:** Number of patients with serious adverse events by SOC and PT, whether or not study drug related (Safety population).

		(N=197)	
SYSTEM ORGAN CLASS	PREFERRED TERM	N	%
Gastrointestinal disorders	Diarrhoea	4	2.03
	Intestinal obstruction	2	1.02
	Intestinal perforation	2	1.02
	Subileus	4	2.03
General disorders and administration site conditions	Asthenia	2	1.02
Immune system disorders	Hypersensitivity	3	1.52
Infections and infestations	Pneumonia	3	1.52
Metabolism and nutrition disorders	Cachexia	2	1.02
Respiratory, thoracic and mediastinal disorders	Pulmonary embolism	7	3.55
Vascular disorders	Deep vein thrombosis	2	1.02
	Hypertensive crisis	3	1.52

Source Tables 3.2.3 of "Tables and Figures" document.

Patients are counted only once in each row.

Each patient could experience more than one adverse event.

Eighteen patients (9.14%) had at least one SAE related to CT, 14 patients (7.11%) had at least one SAE related to Bevacizumab and 4 patients (2.03%) had at least one SAE related to the combination of CT and Bevacizumab. For details, see the Tables 3.2.4a, 3.2.4b and 3.2.4c in the "Table and Figures" document.

For all descriptions of adverse events see Tables from 3.1.1 to 3.2.6 in the "Tables and Figures" document. The listing of the patients who discontinued at least one study treatment permanently due to adverse event(s) is reported in Table 3.2.5 of "Tables and Figures" document.

### 6.3. DEATHS, OTHER SERIOUS ADVERSE EVENTS AND OTHER SIGNIFICANT ADVERSE EVENTS

In the Safety population a total of 100 patients (50.76%) died during the conduct of the study. Six patients (3.05%) discontinued study prematurely for death. In addition, 94 patients (47.72%) died during follow-up principally due to colorectal cancer (86 patients). Listing of the 12 patients who died due to a SAE is reported in the following Table 6.3-1.



## Statistical Report

Sponsor: Roche

Protocol: ML 21380

**Table 6.3-1: Listing of patients who died due to a SAE (Safety population)**

Patient	System Organ Class	Preferred Term	AE	NCI Grade	Start date	End date	Outcome	Relationship to therapy	Action taken
122703_35	Infections and infestations	Pneumonia	PNEUMONIA	4	23SEP2008	13OCT2008	Death	Not related	2 4
122708_79	Cardiac disorders	Cardiovascular disorder	HEART CIRCULATORY BLOCK	4	04FEB2009	04FEB2009	Death	Not related	1
122709_165	Gastrointestinal disorders	Subileus	INTESTINAL SUBOCCLUSION	3	16JUN2009	20JUN2009	Death	Not related	4
122709_190	Renal and urinary disorders	Renal failure	RENAL FAILURE	4	12SEP2009	13SEP2009	Death	Not related	5
122713_30	Cardiac disorders	Cardio-respiratory arrest	CARDIO-RESPIRATORY ARREST	4	02AUG2008	06AUG2008	Death	Related to the combination of CT and Bevacizumab	2
122717_131	General disorders and administration site conditions	Fatigue	FATIGUE	3	25MAY2009	27JUN2009	Death	Related to the combination of CT and Bevacizumab	2 5
122721_43	Gastrointestinal disorders	Intestinal perforation	INTESTINAL PERFORATION	4	18NOV2008	19NOV2008	Death	Related to Bevacizumab	2
122730_139	Cardiac disorders	Myocardial infarction	MYOCARDIAL INFARCTION	4	19MAY2009	19MAY2009	Death	Not related	1
122745_87	Hepatobiliary disorders	Hepatic failure	HEPATIC FAILURE	3	24DEC2008	30DEC2008	Death	Not related	2 4
122745_88	Metabolism and nutrition disorders	Cachexia	CACHEXIA	3	18FEB2009	18FEB2009	Death	Not related	2
122748_73	General disorders and administration site conditions	Death	DEATH	4	22JUN2009	22JUN2009	Death	Not related	1
122748_97	Cardiac disorders	Cardiac arrest	CARDIAC COLLAPSE	4	09JAN2009	09JAN2009	Death	Not related	1

Source Tables 3.2.2 of "Tables and Figures" document

Action taken: 1=No action taken; 2=Study treatment permanently discontinued due to this adverse event; 3=Study treatment adjusted/temporarily interrupted; 4=Concomitant medication taken; 5=Other.

Terms have been codified with MedDRA dictionary, version 11.1.

The complete description of toxicities observed with Bevacizumab, whether or not signaled on CRF as related to Bevacizumab (see Paragraph 3.6.1 for details about selection) by SOC, PT and severity is shown in Table 3.2.6 in the "Tables and figures" document. More in details, the number of patients with these toxicities by SOC, PT and severity occurred in more than one patient are reported in the following Table 6.3-2.

**Table 6.3-2:** Number of patients with toxicities seen with Bevacizumab by SOC, PT and severity, whether or not study drug related (Safety population)

			(N=197)	
SYSTEM ORGAN CLASS	PREFERRED TERM	SEVERITY	N	%
Gastrointestinal disorders	Rectal haemorrhage	1	7	3.55
Renal and urinary disorders	Proteinuria	1	7	3.55
		2	9	4.57
		3	5	2.54
Respiratory, thoracic and mediastinal disorders	Pulmonary embolism	3	3	1.52
		4	4	2.03
Vascular disorders	Deep vein thrombosis	1	2	1.02
		2	3	1.52
		3	3	1.52
	Hypertension	1	22	11.17
		2	27	13.71
		3	4	2.03
	Hypertensive crisis	3	2	1.02
	Thrombosis	4	2	1.02

Source Tables 3.2.6 of "Tables and Figures" document

Patients are counted only once in each row.

Each patient could experience more than one adverse event.

Eighty patients (40.61%) had at least a toxicity observed with Bevacizumab. In terms of PT, the most frequent toxicities were "Proteinuria" and "Hypertension".

Concerning toxicity with grade 4: one patients had "Intestinal perforation", 4 patients "Pulmonary embolism", 1 patient "Deep vein thrombosis", 1 patient "Hypertension", 1 "Hypertensive crisis" and 2 patients "Thrombosis".

## 6.4. CLINICAL LABORATORY EVALUATION

Descriptive statistics of Hematology and Blood Chemistry at each visit are reported in Tables 3.3.1a and 3.3.1b respectively in the "Tables and Figures" document.

The transition frequency tables of Hematology and Blood Chemistry abnormalities based on normal ranges, displaying the disposition of patients from baseline category to the last available visit in which each laboratory data were collected for each patient are presented in Tables 3.3.2a and 3.3.2b respectively in the "Tables and Figures" document.



## Statistical Report

**Sponsor: Roche**

**Protocol: ML 21380**

The incidence rates of low and high abnormal lab values based on normal ranges for Hematology and Blood chemistry are shown in the following Tables 6.4-1 and 6.4-2 respectively (see Paragraph 3.6.1 for details about computation).

**Table 6.4-1:** Incidence rates at the end of treatment of Hematology abnormalities based on normal ranges (Safety population).

		(N=197)			
		INCIDENCE OF LOW ABNORMALITIES		INCIDENCE OF HIGH ABNORMALITIES	
Parameter	Unit of measure	Events At risk	%	Events At risk	%
APTT*	ratio	5/45	11.11	0/45	0.00
	sec	8/129	6.20	3/129	2.33
HAEMOGLOBIN	g/dL	11/97	11.34	5/97	5.15
HEMATOCRIT	%	12/88	13.64	5/88	5.68
INR	ratio	6/155	3.87	9/155	5.81
NEUTROPHILS	10E9/L	14/168	8.33	5/168	2.98
PLATELETS	10E9/L	30/164	18.29	2/164	1.22
RBC	10E12/L	41/144	28.47	4/144	2.78
WBC	10E9/L	22/160	13.75	2/160	1.25

Source Tables 3.3.3a of "Tables and Figures" document

\*APTT was analyzed with different units of measurement.

Patient '122701\_40' had missing APTT baseline value and was considered on 'sec' computation since all other values were collected with this unit of measure.

**Table 6.4-2:** Incidence rates at the end of treatment of Blood Chemistry abnormalities based on normal ranges (Safety population).

		(N=197)			
		INCIDENCE OF LOW ABNORMALITIES		INCIDENCE OF HIGH ABNORMALITIES	
Parameter	Unit of measure	Events At risk	%	Events At risk	%
ALAT (SGPT)	U/L	4/165	2.42	24/165	14.55
ALBUMIN*	%	2/9	22.22	1/9	11.11
	g/L	17/145	11.72	2/145	1.38
ALKALINE PHOSPHATASE	U/L	2/130	1.54	15/130	11.54
ASAT (SGOT)	U/L	0/153	0.00	28/153	18.30
CALCIUM	mg/dL	5/182	2.75	5/182	2.75
CREATININE	mg/dL	12/179	6.70	3/179	1.68
DIRECT BILIRUBIN	mg/dL	1/160	0.63	29/160	18.13
GLUCOSE	mg/dL	1/135	0.74	21/135	15.56
LDH	U/L	1/119	0.84	34/119	28.57
POTASSIUM	mEq/L	13/180	7.22	2/180	1.11

		(N=197)			
		INCIDENCE OF LOW ABNORMALITIES		INCIDENCE OF HIGH ABNORMALITIES	
Parameter	Unit of measure	Events At risk	%	Events At risk	%
SODIUM	mEq/L	8/180	4.44	3/180	1.67
TOTAL BILIRUBIN	mg/dL	0/180	0.00	23/180	12.78

Source Tables 3.3.3b of "Tables and Figures" document

\*Albumin was analyzed with different units of measurement.

Patient '122701\_108' had missing Albumin baseline value and was considered on 'g/L' computation since all other values were collected with this unit of measure.

Further details about Hematology and Blood Chemistry abnormalities based on normal ranges are listed in Tables 3.3.4a and 3.3.4b respectively in the "Tables and Figures" document.

### 6.5. VITAL SIGNS, PHYSICAL FINDINGS AND OTHER OBSERVATIONS RELATED TO SAFETY

Summary statistics of vital signs (excluded height collected only at baseline) at each visit during 1<sup>st</sup> line period are reported in Table 3.4.1 in the "Tables and Figures" document. In addition, descriptive statistics about sitting pulse rate, sitting systolic and diastolic blood pressure during follow-up period are provided.

No important differences were observed in time profiles of body weight, body surface area and sitting pulse rate. A tendency to increasing of sitting systolic and diastolic blood pressure was perceived during 1<sup>st</sup> line period. However, the number of patients observed during the last visits was lowered and therefore all descriptions were not much reliable.

Description of ECOG Performance Status by visit during 1<sup>st</sup> line period and follow-up period is provided in Table 3.4.2 in the "Tables and Figures" document. In all visits, the major of patients had a grade 0 of ECOG performance status. At week 3 and at week 9 one patient had a grade 3 and one patient a grade 4.

Description of ECG every 3 cycles during 1<sup>st</sup> line period is shown in Table 3.4.3 in the "Tables and Figures" document. After baseline, ECG was principally not done, so it has not been possible to observe patients' condition during the study. Among patients who done the exam, ECG interpretation was mainly normal. See "Patient Data Listings" document for specifications about abnormal interpretation.

Descriptive statistics of proteinuria (dipstick) at each visit during 1<sup>st</sup> line period are described in Table 3.4.4 in the "Tables and Figures" document.

### 6.6. SAFETY CONCLUSION

In Safety population, only 11 patients had no one adverse event: the 94.42% of patients had experienced at least one adverse event during study period and follow-up. Fifty-one patients (25.89%) permanently discontinued at least one study treatment due to AE.



## Statistical Report

**Sponsor: Roche**

**Protocol: ML 21380**

Many patients had at least one AE of toxicity grade 3 or 4 (52.79%) and 56 patients (28.43%) had a SAE. In particular 14 patients had a SAE suspected related to Bevacizumab and 4 to the combination of CT and Bevacizumab.

Six patients discontinued study prematurely for death and 94 patients died during follow-up principally due to colorectal cancer. In particular, 12 patients died due to a SAE.