FINAL REPORT

Study title: MIMEB - Molecular Imaging and Molecular Markers in NSCLC treated with Erlotinib and Bevacizumab

A Clinical Pilot Study to Evaluate the Accuracy of FDG-/FLT-PET and DCE-MRI for Early Prediction of Non-Progression in Patients with Advanced Non Squamous Cell Non-Small Cell Lung Cancer (NSCLC) treated with Erlotinib and Bevacizumab and to Associate Imaging Findings with Molecular Markers (phase II)

Investigational products: Erlotinib, Bevacizumab, FDG, FLT, Gadolinium-DPTA

Indication studied: Advanced non-small cell lung cancer (NSCLC)

Short title: MIMEB

EudraCT: 2009-012607-26

ClinicalTrials.gov: NCT01047059

Protocol-ID: MIMEB, ML21801 (Roche), Uni-Koeln-1111 (University of Cologne)

First patient enrolled: 18.01.2010

Last patient last visit: 23.11.2013

Sponsor: University of Cologne, Albertus-Magnus-Platz, D-50923 Cologne; represented by

PI/LKP: Prof. Dr. Jürgen Wolf, Department I of Internal Medicine, Center for Integrated Oncology Köln Bonn, University Hospital of Cologne, 50924 Cologne, Germany. Tel.: +49 221 478 89050, Fax: +49 221 478 89051

Final report author: Dr. Matthias Scheffler, Department I of Internal Medicine, Center for Integrated Oncology Köln Bonn, University Hospital of Cologne, 50924 Cologne, Germany. Tel.: +49 221 478 98766, Fax: +49 221 3978

Version/Date: V1-03f, 05.04.2016

Signature page

The signing authors hereby confirm that the study was performed in compliance with Good Clinical Practice (GCP), with the Declaration of Helsinki and with local laws:

Sponsor/PI

Prof. Dr. Jürgen Wolf	Date

Author

Dr. Matthias Scheffler	Date

Statistician

Prof. Dr. Martin Hellmich	Date

1 Synopsis

Title:	MIMEB - Molecular Imaging and Molecular Markers in NSCLC treated with Erlotinib and Bevacizumab. A Clinical Pilot Study to Evaluate the Accuracy of FDG-/FLT-PET and DCE-MRI for Early Prediction of Non-Progression in Patients with Advanced Non Squamous Cell Non Small Cell Lung Cancer (NSCLC) treated with Erlotinib and Bevacizumab and to Associate Imaging Findings with Molecular Markers
Study drugs:	Name of Finished Product: Avastin® and Tarceva® Name of Active Ingredients: Erlotinib and Bevacizumab
Diagnostic agents	Magnevist® (Gadolinium-DPTA) FDG-EZAG Bonn® (Fluoro-Deoxy-Glukose)
Indication:	First-line treatment of non-squamous cell NSCLC St. IIIb (wet)/IV
Time schedule:	18.01.2010 - 23.11.2013
Kind of trial / Number of centers:	Pilot study (phase II)/ exploratory diagnostic pilot trial / single arm / monocentric
Trial centre:	University Hospital of Cologne / Department I of Internal Medicine/ Kerpenerstr. 62 / 50937 Cologne, Germany
Objectives:	Primary objective: • To evaluate the accuracy of imaging findings in FDG-/FLT-PET and DCE-MRI after one week of treatment for early prediction of RECIST-based non-progression (CR+PR+SD) after 6 weeks of therapy in patients with NSCLC stage IIIb/IV treated first line with erlotinib and bevacizumab • To evaluate the accuracy of imaging findings in FDG-/FLT-PET and DCE-MRI after one week of treatment for early prediction of PFS in patients with NSCLC stage IIIb/IV treated first line with erlotinib and bevacizumab Secondary objectives: • To identify imaging characteristics after one week of treatment predicting RECIST-defined progressive disease (PD), stable disease (SD) and response after 6 weeks of treatment with erlotinib and bevacizumab • To compare the potential of FDG-/FLT-PET and DCE-MRI for early prediction of non-progression • To compare the potential of FDG-/FLT-PET and DCE-MRI regarding patient prognosis • To compare imaging characteristics after one and after 6

	weeks of treatment with regard to their predictive potential for therapy outcome • To compare EGFR- and KRAS-mutational status and imaging characteristics with regard to their potential for early prediction of non-progression • To describe the correlation between pharmacokinetics of bevacizumab and erlotinib with imaging results and clinical outcome • To determine the efficacy of the combination therapy descriptively (response rate [RR], progression free survival [PFS], time on treatment [TOT], disease control rates [DCR], overall survival [OS]) • To describe exploratively correlations between results of high-throughput mutational profiling in tumour tissue, expression profiling in tumour tissue and peripheral blood, imaging results and clinical characteristics • To evaluate safety and efficacy of combination therapy with erlotinib and bevacizumab in patients with clinically stable brain metastases • To evaluate reliability of DCE-MRI in a subset of patients
Methodology:	Single-arm, open label
Number of patients:	40 patients, both genders
Amendments	none
Trial interruption	none

Inclusion criteria:	• Patients with histologically or cytologically proven non-squamous NSCLC stage IIIB with pleural effusion or stage IV • ≥ 18 years of age • Performance status ECOG 0-2 • Estimated life expectancy of at least 12 weeks • Subjects with at least one measurable (CT or MRI) lesion according to RECIST • Adequate bone marrow, liver and renal function as assessed by the following laboratory requirements to be conducted within 7 days prior to screening: - Hemoglobin ≥ 9.0 g/dL - Absolute neutrophil count (ANC) ≥ 1,500 /mm^3 - Platelet count ≥ 100 000/µL - Total bilirubin ≤ 2 x ULN - ALT, AST and alkaline phosphatase (AP) ≤ 2,5 x ULN - PT-INR/PTT < 1.5 x ULN - Creatinine clearance (CrCl) ≥ 60 ml/min calculated by either Cockcroft-Gault or by 24 hours urine collection • Written informed consent (after adequate explanation of the trial) to participate in the trial and to adhere to trial procedures, as well as consenting to data protection procedures • No clinical or radiological sign of interstitial lung disease, no interstitial lung disease in the past • Patients must be able to take oral medication • In case of female patients with childbearing potential: - negative serum or urine HCG in women with childbearing potential - effective method of contraception (Pearl-Index not greater than 1%) - at least 12 months after last menstruation

Exclusion criteria:	• Patient has received prior chemotherapeutic regimens for advanced disease. Prior chemotherapy given as neoadjuvant or adjuvant therapy for early stage disease, completed at least 12 months prior to diagnosis of advanced stage disease, will not be considered as exclusion criterion. • Patient has received prior EGFR-targeted therapy • Squamous-cell carcinoma (SCC) histology, SCLC histology or mixed histology • Evidence of tumor invading or abutting major blood vessels • Patient has signs or symptoms of acute infection requiring systemic therapy (acute or within the last 14 days) • Uncontrolled diabetes mellitus with HbA1c > 7,5% or elevated blood glucose levels levels of > 200 mg/dL • History of uncontrolled heart disease (congestive heart failure > NYHA class 2; active Coronary Arterial Disease (CAD), (MI more than 6 months prior to study entry is allowed); cardiac arrythmias requiring anti-arrythmic therapy (except, when controlled by beta blockers or digoxin) and/or uncontrolled hypertension (> 150/100 mmHg) • Impairment of gastrointestinal function or gastrointestinal disease that may significantly alter the absorption of erlotinib and (e.g. ulcerative disease, uncontrolled nausea, vomiting, diarrhea, malabsorption syndrome or small bowel resection, total parenteral nutrition with lipids) • History of HIV infection or previously sero-positive for the virus • History of Hepatitis B or/and C or previously sero-positive for the Hepatitis B or/and C virus • Patients with seizure disorder requiring CYP3A4-inducing antiepileptics • History of organ allograft • Patients with evidence or history of bleeding diathesis • History of thrombotic disorders within the last 6 months prior to enrolment • Fine needle biopsy or open biopsy within one week prior inclusion • Clinically symptomatic leptomeningeal or brain metastases (patients with clinically stable brain metastases may be enrolled)

Exclusion criteria (c'td)	• Impaired wound healing, non-healing wounds, ulcers, fractures or any condition that provokes uncontrolled bleeding • Preexisting neuropathia ≥ grade 2 • History of grade ≥2 hemoptysis (bright red blood of at least 2.5 ml) • Patients undergoing renal dialysis • Past or current history of cancer other than the entry diagnosis EXCEPT cervical carcinoma in situ, treated basal cell carcinoma, superficial bladder tumors [Ta, Tis & T1] or any cancer curatively treated > 3 years prior to study entry. • Any person being in an institution on assignment of the respective authority • Urine protein qualitative value of > 30 in urinalysis or > +1 in proteinuria testing by dipstick • Any medical, mental or psychological condition which in the opinion of the investigator would not permit the patient to complete the study or understand the patient information • Concomitant or intented anticoagulation therapy • Planned surgical or dental invasive intervention (e.g. tooth extraction, planned surgeries) during the course of the study • Any serious medical condition with organ impairment • Hypersensitivity to bevacizumab or erlotinib or any of their ingredients • Major surgery or significant traumatic injury within the last 4 weeks before inclusion • Parallel participation in another clinical trial or participation in another clinical trial within the last 30 days or 7 half-life's, whatever is of longer duration, prior study start • Pregnancy, breast feeding • Claustrophobia • Known allergic reaction to Gadolinium • Heart pacemaker • Ferromagnetic and electronic implants in special locations (e. g. cerebral) • Cochlea implants • known allergic reaction to non-ionic iodinated computed tomography contrast agents • known hyperthyroidism

Duration of therapy:	Study specific therapy was administered for six weeks.
Statistical methods:	NOTE: Hypothesis generation only: All statistical analyses are of exploratory nature. - ROC analysis of baseline values and changes in SUVs of FDG-/FLT-PET. ROC analysis of baseline values and changes in K^{trans}/k_{ep}/K_i/IAUC for DCE-MRI. Explorative analysis of BF, BV, PS and MTT values and changes during treatment in DCE-MRI. PFS/OS analysis with possibly predefined cut-off values in FDG-/FLT-PET for metabolic response. Correlation of baseline values and analysis of possibly predefined cut-off values in correlation with PFS/OS/RR. Statistical association of molecular status and clinical response (Fisher's exact test). Descriptive assessment of survival (Kaplan-Meier (curve) estimation). Pharmacokinetic modelling using NONMEM. Descriptive safety analysis.
GCP-Conformity:	This trial was performed according to ICH-GCP, including the archivation of essential documents

2 Table of contents

3 List of abbreviations and definition of terms

Abbreviation	Meaning
AE	Adverse event
ALT	Alanine transaminase
AMG	Arzneimittelgesetz (German regulations concerning the production and sale of medicines)
An	Anamnesis case report form
γ-GT	γ-glutamyl transferase
AP	Alkaline phosphatase
AR	Adverse reaction
ASCO	American Society for Clinical Oncology
ASR	Annual safety report
AST	Aspartate transaminase
ATL-1	Aspirin-triggered lipoxin A(4) analogue
AUC	Area under (ROC) curve
BAL	Bronchoalveolar lavage
BF	Blood flow
BfArM	Bundesinstitut für Arzneimittel und Medizinprodukte (German medicines and medical products authority)
BM	Bone marrow
BMBF	Bundesministerium für Bildung und Forschung (Governmental department of education and research)
BFS	Bundesamt für Strahlenschutz (German federal authority for the safety of radiation)
BV	Blood volume
CHD	Coronary heart disease
CNS	Central nervous system
CIO	Center for Integrated Oncology Cologne-Bonn
CR	Complete response
CRF	Case report form
CRP	C-reactive protein
CRr	CR with residual abnormalities (= CU)

CS	Clinical stage
CT	Computed tomography
CTC	Common toxicity criteria
DCE	Dynamic contrast enhanced
DCO	Diffusion capacity for oxygen
DICOM	Digital imaging and communications in medicine
DMC	Data monitoring committee
EDTA	Ethylene-diamine-tetra-acetic acid
EC	Erythrocyte concentrate
ECG	Electrocardiogram
ECOG	Eastern Cooperative Oncology Group
ED	Effective dosage
EGF	Epidermal growth factor
ELISA	Enzyme-linked Immunosorbent Assay
EMEA	European medicines agency
EORTC	European Organization for Research and Treatment of Cancer
ESR	Erythrocytes sedimentation rate
F	Follow up case report
FDA	Food & Drug Administration
FA	Final (data) analysis
FDG	Fluoro-deoxy-glucose
FLT	Fluoro-L-thymidine
FFTF	Freedom from treatment failure
FPFV	First patient, first visit
GCP	Good clinical practice
Gd-DPTA	Gadolinium-Diethyltriaminepentaacetic acid
gGT	Gamma-glutamyl-transferase
GOT	Glutamate oxalacetate transaminase
GPT	Glutamate pyruvate transaminase
Hb	Hemoglobin
Hk	Hematocrit
HR-CT	High resolution computed tomography

IA	Interim (data) analysis
IAUC	Initial area under the contrast agent concentration - time curve
ICDO	International Classification on Diseases for Oncology
ICH	International Conference on Harmonization
ICRU	International Commission on Radiation Units and Measurements
IL-8	Interleukin-8
INR	International normalized ratio
ISF	Investigator site file
ITT	Intention to treat
ICAM-1	Inter-cellular adhesion molecule 1
K^{trans}	Bidirectional volume transfer coefficient
k_{ep}	Rate constant
K_i	Unidirectional influx constant
KRAS	Kirsten rat sarcoma viral oncogene homolog
LCGC	Lung cancer group Cologne
LDH	Lactate dehydrogenase
LN	Lymph node(s)
LPLV	Last patient, last visit
LTCG	Laboratory of translational cancer genomics
LV	Left ventricle
MDS	Myelodysplastic syndrome
MFI	Multidimensional Fatigue Inventory
MI	Myocardial infarction
MPI	Max-Planck-Institute
MRI	Magnetic resonance imaging
MTD	Maximal tolerated dose
mTOR	Mammalian target of rapamycin
MTT	Mean transit time
NC	No change
NCI-CTC	National Cancer Institute Common Toxicity Criteria
NMR	Nuclear magnetic resonance
NONMEM	Non-linear mixed effects modelling

NSCLC	Non-small cell lung cancer
NYHA	New York Heart Association
OS	Overall survival
$_{P}CO2$	Carbon dioxide partial pressure
PCR	Polymerase chain reaction
PD	Progressive disease
PDGF	Platelet-derived growth factor
PDGFR	Platelet-derived growth factor receptor
PEI	Paul-Ehrlich-Institut
PET	Positron emission tomography
PFS	Progression-free survival
PIGF	Phosphatidylinositol glycan anchor biosynthesis, class F
PK	pharmacokinetics
PO2	Oxygen partial pressure
PR	Partial response
PS	Permeability surface area product
PT	Prothrombin time
PTT	Partial thromboplastin time
QM	Quality management
QoL	Quality of life
Raf	V-raf murine sarcoma viral oncogene
RBC	Red blood cells
RE	Restaging case report form
RECIST	Response evaluation criteria in solid tumors
ROC	Receiver operating characteristics
REG	Registration case report form
RF	Risk factor
RFS	Relapse free survival
SD	Stable disease
SAE	Serious adverse event
SCC	Squamous cell carcinoma
SCLC	Small-cell lung cancer

SNP	Single-nucleotide polymorphism
SOPs	Standard operating procedures
SUSAR	Suspected unexpected serious adverse reaction
SUV	Standard uptake value
sSUV	Sum of SUV
sVCAM	Soluble vascular cell adhesion molecule
T1	In bladder cancer: Tumor infiltrating submucosa
Ta	In bladder cancer: Non-invasive papillary carcinoma of the urothel
TEE	Transesophageal echocardiography
Th	Therapy case report form
Tis	Carcinoma in situ
TKI	Tyrosine kinase inhibitor
TMF	Trial master file
TSH	Thyroid stimulating hormone
TSP-1	Thrombospondin-1
ULN	Upper limit of normal
VEGF	Vascular endothelial growth factor
WBC	White blood cells
WHO	World Health Organisation
ZKS Köln	Zentrum für Klinische Studien Köln (Clinical Trials Center Cologne)

4 Ethics

4.1 Independent Ethics Committee (IEC) and Institutional Review Board (IRB)

The study including all amendments and safety reports were reviewed and approved by the Ethics Committee of the University of Cologne. The termination of this trial was supervised by the Institutional Review Board (Studienkommission der Klinik I für Innere Medizin, University Hospital of Cologne).

4.2 Ethical conduct of the study

The conduct of this study was performed based on the Declaration of Helsinki (1996 version), the German Drug Law (Arzneimittelgesetz §§40-42 in the current versions) and the ICH-GCP guidelines. In concordance with German law, all enrolled patients received a trial-specific insurance.

4.3 Patient information and consent

The patients were informed about the trial, including the following aspects: title and aim of the trial, nature of the treatment, side effects, risks of the imaging procedures including side effects of the contrast agents and the radiation exposure, reason for recruitment, passing on of data and material samples, insurance, the ethics committee vote and the patient's freedom to decide. The patients received an informed consent form.

Informed consent was provided before any trial-specific procedure had taken place, i. e. usually at allocation.

5 Investigators and study administrative structures

Principal Investigator:	Prof. Dr. Jürgen Wolf Department I for Internal Medicine Center of Integrated Oncology Köln Bonn University Hospital of Cologne 50924 Cologne
Coordinating investigator:	Dr. Matthias Scheffler Department I for Internal Medicine Center of Integrated Oncology Köln Bonn University Hospital of Cologne 50924 Cologne
Trial Coordination:	Lung Cancer Group Cologne (LCGC) Department I of Internal Medicine Center for Integrated Oncolgy University Hospital of Cologne Kerpener Str. 62 D-50937 Cologne

Monitoring:	ZKS Köln (Clinical Trials Center Cologne) University of Cologne Gleueler Str. 269 D-50935 Cologne
Data management:	ZKS Köln (Clinical Trials Center Cologne) University of Cologne Gleueler Str. 269 D-50935 Cologne
SAE-management:	ZKS Köln (Clinical Trials Center Cologne) University of Cologne Gleueler Str. 269 D-50935 Cologne
Scientific advice:	Prof. Dr. Oliver Cornely ZKS Köln (Clinical Trials Center Cologne) University of Cologne Gleueler Str. 269 D-50935 Cologne
Quality management:	ZKS Köln (Clinical Trials Center Cologne) University of Cologne Gleueler Str. 269 D-50935 Cologne
Laboratories:	Laboratory of Translational Cancer Genomics (LTCG) Max-Planck-Institute for Neurological Research Gleueler Str. 50 D-50931 Cologne
	Institute of Clinical Chemistry University Hospital of Cologne Kerpener Str. 62 D-50937 Cologne
	Institute of Pharmacology Clinical Pharmacology University Hospital of Cologne Gleueler Str. 24 D-50931 Cologne
Technical Facilities:	Max-Planck-Institute for Neurological Research Department of Radiochemistry Gleueler Str. 50 D-50931 Cologne

Max-Planck-Institute for Neurological Research
Multimodality Imaging
Gleueler Str. 50
D-50931 Cologne

Department of Nuclear Medicine
University Hospital of Cologne
Kerpener Str. 62
D-50937 Cologne

Department of Radiology
University Hospital of Cologne
Kerpener Str. 62
D-50937 Cologne

Radiology Practice Köln-Sülz
Berrenrather Str. 296
D-50937 Cologne

Institute of Pathology
University Hospital of Cologne
Kerpener Str. 62
D-50937 Cologne

Department I of Internal Medicine
University Hospital of Cologne
Kerpener Str. 62
D-50937 Cologne

Department of Pathology and Neuropathology
University of Bonn Medical Center
Siegmund-Freud-Str. 25
D-53127 Bonn

6 Introduction

Erlotinib and bevacizumab represent new substances in cancer treatment which act beyond conventional chemotherapeutical cytotoxicity. Nevertheless, identifying patients who might benefit from this treatment is a major concern, as not all patients represent the optimal targets for this therapy (i. e., EGFR mutations for erlotinib efficacy, an angiogenic phenotype for bevacizumab efficacy, and a potential "cross-talk" between both substances by inducing tumor hypoxia). Within this trial, we set out to identify patients benefitting clinically from the combination therapy by analysing early metabolic or prolif-

erative changes as well as changes in the tumor vasculature independently from the tumor genotype. The rationale of this trial is the notion that metabolic and proliferative changes measured by PET are earlier detectable than morphologic changes measured by CT scans and therefore might spare potential ineffective treatment for patients. Further, there is a hypothesis that for tumor vessels, there is a small and early window of normalization under specific therapy which might enhance the delivery of erlotinib to the tumors induced by bevacizumab. We explanatorily analyzed if this normalization is non-invasively detectable by DCE-MRI.

A total number of participants of 40 was preplanned in order to detect the accuracy of changes in the early imaging parameters to predict clinical outcome in terms of response rates and progression-free survival. The medication and financial support was provided by Roche Pharmaceuticals. The protocol followed closely the one of the terminated ERLO-PET trial (ClinicalTrials.gov: NCT00568841) and was performed in accordance with the aforementioned guidelines. Support was also provided by the German Ministry for Research and Education (BMBF, grant 01KN0706).

7 Study objectives

Primary objective:

- To evaluate the accuracy of imaging findings in FDG-/FLT-PET and DCE-MRI after one week of treatment for early prediction of RECIST-based non-progression (CR+PR+SD) after 6 weeks of therapy in patients with NSCLC stage IIIb/IV treated first line with erlotinib and bevacizumab
- To evaluate the accuracy of imaging findings in FDG-/FLT-PET and DCE-MRI after one week of treatment for early prediction of PFS in patients with NSCLC stage IIIb/IV treated first line with erlotinib and bevacizumab

Endpoints:

The predictive value/accuracy of FDG-/FLT-PET and DCE-MRI regarding non-progression in NSCLC patients treated with erlotinib and bevacizumab after 6 weeks will be evaluated by ROC analysis of percentual changes of (semi-)quantitative parameters (see parameters listed above; area under the ROC curve, AUC); H_0: $AUC_{FDG/FLT\ PET/DCE\text{-}MRI} \leq 0.5$, H_A: $AUC_{FDG/FLT\ PET/DCE\text{-}MRI} > 0.5$ (one-sided level 2.5% , not adjusted for multiple testing).

To evaluate the accuracy of imaging findings in FDG-/FLT-PET and DCE-MRI after one week of treatment for early prediction of PFS in patients with NSCLC treated first line with erlotinib and bevacizumab. Cut-offs maximising the Youden index for each method will be used for Kaplan-Meier (curve) estimations.

Additionally, predefined percentual reductions of (semi-)quantitative imaging parameters will be analyzed:

- reductions of at least 10%
- reductions of at least 20%
- reductions of at least 30%
- reductions of at least 40%
- reductions of at least 50%

Further reduction values can be analyzed depending on ROC analyses.

Secondary objectives:

- To identify imaging characteristics after one week of treatment predicting RECIST-defined progressive disease (PD), stable disease (SD) and response after 6 weeks of treatment with erlotinib and bevacizumab
- To compare the potential of FDG-/FLT-PET and DCE-MRI for early prediction of non-progression
- To compare the potential of FDG-/FLT-PET and DCE-MRI regarding patient prognosis
- To compare imaging characteristics after one and after 6 weeks of treatment with regard to their predictive potential for therapy outcome
- To compare EGFR- and KRAS-mutational status and imaging characteristics with regard to their potential for early prediction of non-progression
- To describe the correlation between pharmacokinetics of bevacizumab and erlotinib with imaging results and clinical outcome
- To determine the efficacy of the combination therapy descriptively (response rate [RR], progression free survival [PFS], time on treatment [TOT], disease control rates [DCR], overall survival [OS])
- To describe exploratively correlations between results of high-throughput mutational profiling in tumor tissue, expression profiling in tumor tissue and peripheral blood, imaging results and clinical characteristics
- To evaluate safety and efficacy of combination therapy with erlotinib and bevacizumab in patients with clinically stable brain metastases
- To evaluate reliability of DCE-MRI in a subset of patients
- FLT and FDG standard uptake value, EGFR mutational status, KRAS mutational status, EGFR positivity in immunohistochemistry, quantification of biomarkers, standardized hemodynamic parameters in DCE-MRI (K^{trans}, V_e, K_{ep}, BV, BF, MTT)**.**
- One year PFS rate, one-year OS rate.

Endpoints:

The predictive/prognostic value/accuracy of FDG/FLT-PET/DCE-MRI regarding clinical failure, SD and clinical response in NSCLC patients treated with erlotinib and bevaci-

zumab will be compared by ROC analysis; cut-offs maximizing the Youden index will be taken to compare PFS/OS.

Association between EGFR mutation, KRAS mutation and EGFR positivity with clinical response will described (Fisher's exact test).

Secondary endpoints (PFS and OS) will be analyzed using descriptive statistics after at least one-year follow-up of all patients.

Overall survival time will be measured descriptively after follow-up.

Response rates will be measured in accordance with RECIST-criteria (see below).

Safety of erlotinib and bevacizumab in combination will be assessed by evaluating adverse events and serious adverse events, descriptively analyzing the clinical outcome.

The statistical association of EGFR respectively KRAS mutational status, EGFR positivity and clinical response in NSCLC patients treated with erlotinib and bevacizumab will be evaluated by Fisher's exact test (at one-sided level 2.5%; not adjusted for multiple testing).

Association between biomarkers and imaging characteristics will be exploratory described and analyzed.

Pharmacokinetics will be explanatorily described regarding average plasma levels, AUC concentrations and proposed peak plasma levels. The findings will be correlated with imaging results and with the parameters of the safety analysis, as well as with biomarker analysis results.

Optional CYP polymorphism description will be performed exploratory, correlating the results with the results mentioned above.

8 Investigational plan

8.1 Overall study design and plan - description

This study was performed as a therapeutic exploratory clinical pilot study. Because of our primary objective in this trial, no placebo-group was established. For the same reason randomization after enrollment was not done. A fixed combination of erlotinib and bevacizumab was given for six weeks and, in cases where no progression in restaging at week 7 is confirmed, thereafter as long as the patient benefitted from therapy. Crossover was not applicable.

After enrollment, baseline procedures were performed within 14 days before first administration of therapy (d-14 – d1, if a procedure takes place the same day as therapy start. Nevertheless, baseline procedures had to be performed before administration of therapy). Baseline procedures included FDG- and FLT-PET imaging, and DCE-MRI imaging (in the subset of the first patients to be enrolled, DCE-MRI baseline imaging were performed twice). The last available CT scan of the involved regions was considered baseline-CT and should preferably not have been older than 14 days. In cases where the available CT-

data did not meet quality demands in the opinion of the investigator regarding clinical practice, a new CT scan should have been performed between d-14 and d1.

Tumor material from which the diagnosis was made was sent to LTCG and to the Department of Pathology and Neuropathology, University of Bonn Medical Centre. This department later became the Institute of Pathology of the University Hospital of Cologne and analyzed the probes after establishment of a next-generation sequencing (NGS) pipeline again. Phenotyping of CYP1A2 and CYP3A4 was not performed. Brain metastases were confirmed either by a CT or MRI scan of the skull.

Between d7 and d14, week 2 imaging procedures were performed, including an FDG-PET, an FLT-PET and DCE-MRI-imaging.

During the six-week-course of therapy, pharmacokinetic evaluation and biomarker assessment were not performed due to lack of feasibility. Adverse events and concomitant medication will be documented (see "Adverse events").

After the six weeks of the treatment cycle, restaging will take place in week 7 with a CT-scan (according to standard medical/clinical care) and FDG-PET-, FLT-PET- and DCE-MRI-imaging. The MRI-scan were evaluated according to RECIST in comparison with baseline MRI scan retrospectively by experienced radiologists and reviewed by external experts. The assessment of PFS was performed the same way.

Further treatment depends on the results of week-7-CT scan in accordance with standard clinical practice.

The protocol of this trial was comparable with the one used for the ERLOPET-trial (see above). In this trial, erlotinib was given in monotherapy, and DCE-MRI was not performed. The results of this trial demonstrated the accuracy of early PET-imaging in non-progression of the disease later under therapy regardless from the mutational status.

Figure 1 shows the flow-sheet of the ERLOPET trial, whereas Figure 2 the one of the present MIMEB trial.

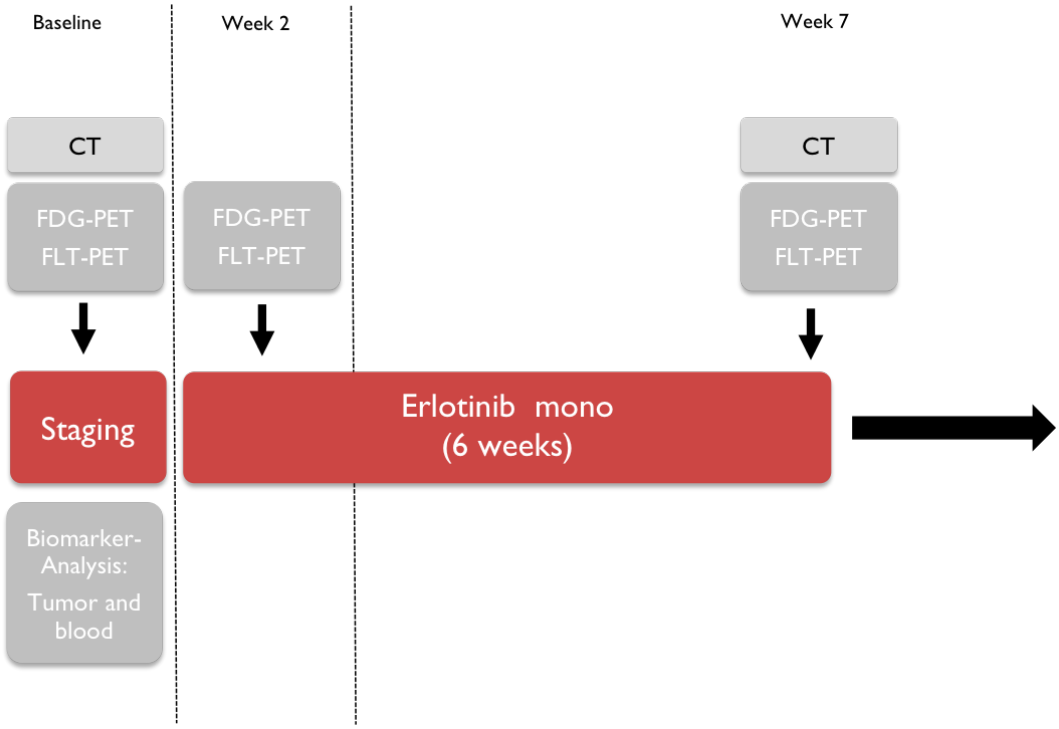


Figure 1: Flow-sheet of ERLOPET

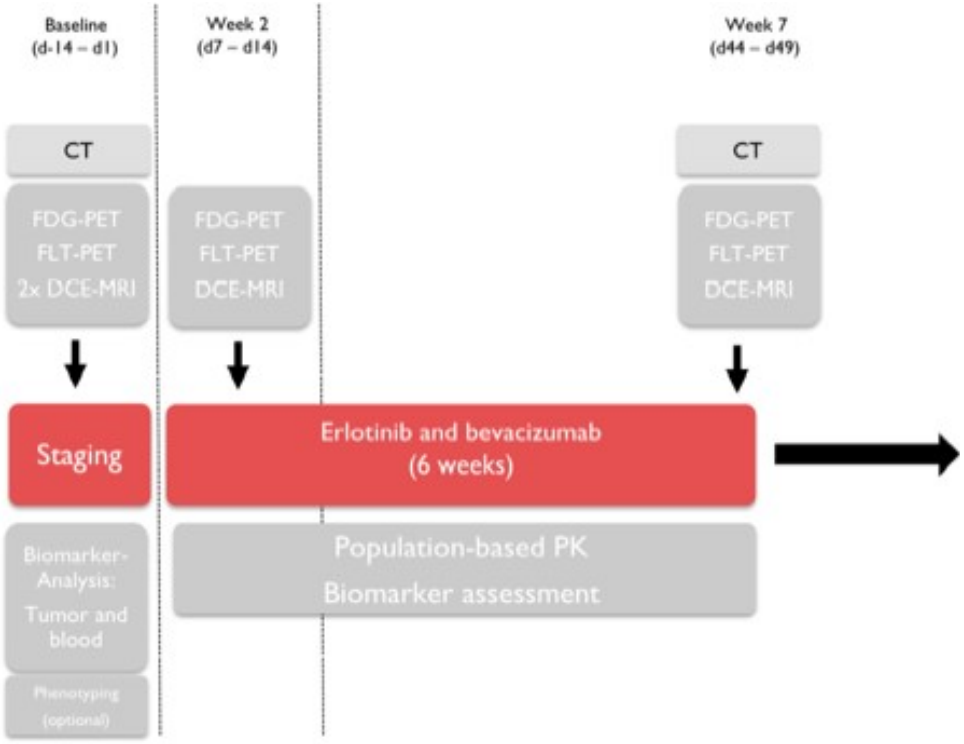


Figure 2: Flow-sheet of MIMEB

The trial was one-armed, non-randomized. 40 patients were statistically considered the total number of enrolled participants. The patients had NSCLC stage IV as primary diagnosis and did not receive prior systemic treatment. All patients started with 150 mg erlotinib/d and 15 mg/kg q3w. The tracers for imaging were FDG, FLT, and Gd-DPTA.

The conduction of the trial was supervised by the institutional Trial Steering Committee. The exact trial specific procedures and visits were predefined:

Screening and baseline procedures

Within a maximum of **21 days prior to d1**, all patients will be screened by the investigator. Each patient must provide a written informed consent to the trial procedures. The patient must be informed about the study verbally and by the patient information by the principal investigator or one of the co-investigators, before informed consent is obtained. After obtaining written informed consent, the patient will be screened for eligibility.

Between d-21 and d1 the following parameters will be assessed:

- Patient demography (incl. date of birth, gender, ethnic group, smoking history)
- Height, weight, vital signs including blood pressure and pulse rate, ECOG performance status
- Medical history (including concurrent illnesses)
- Documentation of adverse events
- Assessment of biomarkers
- MRI scan of the skull

Between d-14 and d1 (start of therapy), the following procedures will be performed:

- Laboratory assessment including:

- serum electrolytes (Na, K, Ca)
- serum liver function parameters (total bilirubin, AST, ALT, AP, gGT, LDH)
- WBC, Hb, Thrombocytes, ANC
- serum renal function parameters (creatinine, uric acid, urea)
- serum TSH, FT3, FT4 (only for screening)
- blood coagulation parameters (Quick, PTT, Fibrinogen) (only for screening)
- pregnancy test (HCG) for women with childbearing potential (only for screening)
- urine dipstick (only for screening)

- FDG-, FLT-PET and DCE-scan of involved areas (There are two DCE-MRI measurements to achieve as validated baseline data in the first 20 patients to be enrolled in this trial. For DCE-MRI, one specific lesion, preferable in the lung, will be analysed. Brain metastases will be analysed by FDG- and FLT-PET, DCE-MRI)
- Pharmacokinetic and biomarker sampling as mentioned above
- Optionally: CYP3A4 and CYP1A2 phenotyping

Week 2 imaging and study visit

Between d7 and d14, the following is to be assessed:

- FDG scan
- FLT scan
- DCE-MRI imaging
- Laboratory assessment as mentioned above
- Documentation of adverse events
- Documentation of concomitant medication
- Pharmacokinetic blood samples as mentioned above
- RR measurement and weight

Week 4 study visit

The study visit between **d22 and d28** will include:

- Laboratory assessment as mentioned above
- Documentation of adverse events
- Documentation of concomitant medication
- Pharmacokinetic blood samples as mentioned above
- Extended pharmacokinetic profile as mentioned above (d22)
- RR measurement and weight

Week 7 imaging and study visit

Between d44 and d49, the following is to be assessed:

- CT scan according to clinical care standards
- FDG scan
- FLT scan
- DCE-MRI imaging

- Laboratory assessment as mentioned above
- Documentation of adverse events
- Documentation of concomitant medication
- Pharmacokinetic blood samples as mentioned above
- RR measurement and weight

Visits in cases of disease control after six weeks of therapy

In cases of disease control (CR/PR/SD after week 7 CT scan), the patient will have **study visits every six weeks**, count on from the date of the last CT scan. These visits include:

- DCE-MRI imaging
- Laboratory assessment as mentioned above
- Documentation of adverse events
- Documentation of concomitant medication
- A single pharmacological blood sample
- RR measurement and weight

Visits in cases of progression

Patients who suffer from progressive disease (at any timepoint) will be excluded from the study thereafter. Their **Follow-up visits on day 14 and 28 after end of treatment** will include:

- Laboratory assessment as mentioned above
- Documentation of adverse events
- Documentation of concomitant medication
- Pharmacokinetic blood samples as mentioned above
- RR measurement and weight

Table 1: Visit schedule of the trial

	Screening	d-14 – d1	d7 – d14	d22 – d28	d44 – d49	6-weekly visits	Follow-up
Patient demography	X						
Height, vital signs, ECOG	X						
Biomarker assessment (tissue)	X						
AEs, blood pressure, weight	X		X	X	X	X	X
MRI skull	X						
Laboratory assessment		X	X	X	X	X	X
FDG-PET and FLT-PET		X	X		X		
DCE-MRI		X	X		X	X	
Concurrent medication	X	X	X	X	X	X	X
Biomarker (blood) and PK sampling**	X	(x)	(x)	(x)	(x)	(x)	(x)
Extended PK profiling				X			

* if possible

** if possible, see below

8.2 Trial population

8.2.1 Inclusion criteria

- Patients with histologically or cytologically proven non-squamos NSCLC stage IIIB with pleural effusion or stage IV
- ≥ 18 years of age
- Performance status ECOG 0-2
- Estimated life expectancy of at least 12 weeks
- Subjects with at least one measurable or nonmeasurable (CT or MRI) lesion according to RECIST
- Adequate bone marrow, liver and renal function as assessed by the following laboratory requirements to be conducted within 7 days prior to screening:

- Hemoglobin ≥ 9.0 g/dL
- Absolute neutrophil count (ANC) ≥ 1,500 /mm^3
- Platelet count ≥ 100 000/µL
- Total bilirubin ≤ 2 x ULN
- ALT, AST and alkaline phosphatase (AP) ≤ 2,5 x ULN
- PT-INR/PTT < 1.5 x ULN
- Creatinine clearance (CrCl) ≥ 60 ml/min calculated by either Cockcroft-Gault or by 24 hours urine collection

- Written informed consent (after adequate explanation of the trial) to participate in the trial and to adhere to trial procedures, as well as consenting to data protection procedures
- No clinical or radiological sign of interstitial lung disease, no interstitial lung disease in the past
- Patients must be able to take oral medication
- In case of female patients with childbearing potential:
    - negative serum or urine HCG in women with childbearing potential
    - effective method of contraception (Pearl-Index not greater than 1%)
    - at least 12 months after last menstruation

8.2.2 Exclusion criteria

- Patient has received prior chemotherapeutic regimens for advanced disease. Prior chemotherapy given as neoadjuvant or adjuvant therapy for early stage disease, completed at least 12 months prior to diagnosis of advanced stage disease, will not be considered as exclusion criterion.
- Patient has received prior EGFR-targeted therapy
- Squamous-cell carcinoma (SCC) histology, SCLC histology or mixed histology
- Evidence of tumor invading or abutting major blood vessels
- Patient has signs or symptoms of acute infection requiring systemic therapy (acute or within the last 14 days)
- Uncontrolled diabetes mellitus with HbA1c > 7,5% or elevated blood glucose levels levels of > 200 mg/dL
- History of uncontrolled heart disease (congestive heart failure > NYHA class 2; active Coronary Arterial Disease (CAD), (MI more than 6 months prior to study entry is allowed); cardiac arrythmias requiring anti-arrythmic therapy (except, when controlled by beta blockers or digoxin) and/or uncontrolled hypertension (> 150/100 mmHg)
- Impairment of gastrointestinal function or gastrointestinal disease that may significantly alter the absorption of erlotinib and (e.g. ulcerative disease, uncontrolled nausea, vomiting, diarrhea, malabsorption syndrome or small bowel resection, total parenteral nutrition with lipids)
- History of HIV infection or previously sero-positive for the virus
- History of Hepatitis B or/and C or previously sero-positive for the Hepatitis B or/and C virus
- Patients with seizure disorder requiring CYP3A4-inducing anti-epileptics
- History of organ allograft
- Patients with evidence or history of bleeding diathesis
- History of thrombotic disorders within the last 6 months prior to enrolment
- Fine needle biopsy or open biopsy within 1 week prior inclusion
- Clinically symptomatic leptomeningeal or brain metastases (patients with clinically stable brain metastases may be enrolled)
- Impaired wound healing, non-healing wounds, ulcers, fractures or any condition that provokes uncontrolled bleeding
- Preexisting neuropathia ≥ grade 2
- History of grade ≥2 hemoptysis (bright red blood of at least 2.5 ml)
- Patients undergoing renal dialysis

- Past or current history of cancer other than the entry diagnosis EXCEPT cervical carcinoma in situ, treated basal cell carcinoma, superficial bladder tumors [Ta, Tis & T1] or any cancer curatively treated > 3 years prior to study entry.
- Any person being in an institution on assignment of the respective authority
- Urine protein qualitative value of > 30 in urinalysis or > +1 in proteinuria testing by dipstick
- Any medical, mental or psychological condition which in the opinion of the investigator would not permit the patient to complete the study or understand the patient information
- Concomitant or intented anticoagulation therapy
- Planned surgical or dental invasive intervention (e.g. tooth extraction, planned surgeries) during the course of the study
- Any serious medical condition with organ impairment
- Hypersensitivity to bevacizumab or erlotinib or any of their ingredients
- Major surgery or significant traumatic injury within the last 4 weeks before inclusion
- Parallel participation in another clinical trial or participation in another clinical trial within the last 30 days or 7 half-life's, whatever is of longer duration, prior study start
- Pregnancy, breast feeding
- Claustrophobia
- Known allergic reaction to Gadolinium
- Heart pacemaker
- Ferromagnetic and electronic implants in special locations (e. g. cerebral)
- Cochlea implants
- known allergic reaction to non-ionic iodinated computed tomography contrast agents
- known hyperthyroidism

8.2.3 Discontinuation of patients from treatment or analyses

Predefined discontinuation criteria were defined according to the protocol as follows:

The patient can subsequently discontinue trial participation at his/her own wish at any time.

Specific reasons for discontinuing a patient from the trial are:

- Voluntary discontinuation by the patient, who is at any time free to discontinue his/her participation in the study without negative effects to further treatment
- Safety reasons as judged by the investigator or the sponsor

- Severe non-compliance or situations which would jeopardize compliance to the protocol as judged by the investigator or sponsor
- Incorrect enrollment of the patient (i. e., the patient does not meet the required inclusion/exclusion as defined by the criteria)
- Patient lost to follow-up
- Disease progression
- Unacceptable toxicity
- Serious intercurrent disease
- Death
- Pregnancy

Exclusion of the documentation was generally not planned.

8.3 Trial medication

8.3.1 Antineoplastic drugs

Table 2: Study medication

Medication	Administered
Erlotinib 150 mg p.o.	d1-d43
Bevacizumab 15 mg/kg KG i.v.	d1, d22, d43

8.3.2 Erlotinib

Trade name: Tarceva[R]

International non-proprietary name: Erlotinib

Dosage: 25 mg, 100 mg, 150 mg

Pharmaceutical form: Tablet

Marketing authorisation holder: Roche Registration Limited

Indications: Tarceva[R] monotherapy is indicated for the treatment of patients with locally advanced or metastatic NSCLC after failure of at least one prior chemotherapy regimen (FDA/EMEA). It is also indicated for the first-line treatment in combination with gemcitabine of patients with locally advanced, unresectable, or metastatic pancreatic cancer (FDA/EMEA) and for the first-line treatment of EGFR-mutated NSCLC patients.

Reference document:	The reference document is the German SPC of Tarceva[R] in its current version.
Fabrication:	Tarceva[R] as commercially available was fabricated and provided by Roche.
Labeling:	Due to GCP-V §5 (8), Tarceva[R] was provided as its correspondent trade ware. Further labelling (e. g. anonymisation) was not necessary, for no randomization or use of placebo was performed.
Storage:	Tarceva[R] was provided by Roche and stored at room temperature in the study office of LCGC.
Batch Numbers:	B2003B03; B2006B01; B2011B01; B2028B01; B2068B01

8.3.3 Bevacizumab

Trade name:	Avastin[R]
International non-proprietary name:	Bevacizumab
Dosage:	100 mg, 400 mg
Pharmaceutical form:	Vail for infusion
Marketing authorisation holder:	Roche Registration Limited
Indications:	Avastin[R] is indicated for the first-line treatment of patients with locally advanced or metastatic NSCLC in combination with carboplatin and paclitaxel or gemcitabine and cisplatinum (except tumors with predominant squamous cell histology). It is also indicated for the first-line treatment in combination with 5-FU-based chemotherapy of patients with metastatic colorectal cancer and, in combination with paclitaxel, for the first-line therapy of patients with metastatic breast cancer. It is further indicated in combination with interferon alfa in the first line treatment of advanced renal cell carcinoma.
Reference document:	The reference document is the German SPC of Avastin[R] in its current version.
Fabrication:	Avastin[R] as commercially available will be fabricated and provided by Roche.
Labeling:	Due to GCP-V §5 (8), Avastin[R] was provided as its correspondent trade ware. Further labelling is not necessary, for no randomization or use of placebo is performed.

Storage:	Avastin^{R} was provided by Roche and stored at 2-8°C in a temperature controlled, locked in the outpatient department of the Department of central pharmacy for cytotoxic agents (Zentrale Zytostatika Zubereitung, ZZZ) of University Hospital of Cologne.
Batch Numbers:	H0105B02;H0112B02;B2017B01;H0101B02;H0103B02; B2001B01;H0109B01;H0011; B7001; B2017; H0005B01

8.3.4 Diagnostic drugs

For PET procedures, fluoro-D-glucose (FDG) and fluoro-thymidine (FLT) were used. FDG was provided as its correspondent trade ware. FLT was produced and provided by the Max-Planck-Institute of Neurological Research, Multimodality Imaging Group (authorisation holder).

For DCE-MRI imaging, gadolinium-DPTA was as available as tradeware. Please refer to the attached "REQUEST FOR AUTHORISATION OF A CLINICAL TRIAL ON A MEDICINAL PRODUCT FOR HUMAN USE TO THE COMPETENT AUTHORITIES AND FOR OPINION OF THE ETHICS COMMITTEES IN THE COMMUNITY" ("Modul 1") for further details.

8.3.5 Concomitant medication

Concomitant medication was documented during the whole trial assessment. Because of the antiangiogenic properties of bevacizumab, anticoagulation therapy was prohibited by exclusion criterion.

Erlotinib is mainly metabolized by hepatic cytochrome p450 enzymes, with CYP3A4 playing a major role in erlotinib clearance, and, to a less extend, CYP1A2. Table 3 summarizes the potential effect of substrates, inducers or inhibitors of these enzymes. The investigators took care of potential interactions.

Table 3: Erlotinib drug interactions

Erlotinib should be administered with care when co-administered with the following drugs:

Drug	Effects
CYP3A4 inducers: Phenytoin, carbamazepine, rifampicin, barbiturates, corticoids etc. **CYP1A2 inducers:** Tobacco, omeprazole, insulin etc.	Blood **erlotinib concentration may be decreased**, which may lead to decreased effects of the drug. Co-administration of **tobacco** resulted in an 50% decrease in the AUC of erlotinib.
CYP3A4 inhibitors: Azole antifungal agent (i.e. itraconazole), macrolide (i.e. erythromycin), ritonavir, indinavir, diltiazem and verapamil etc. **CYP1A2 inhibitors:** Ciprofloxacin, amiodarone, fluvoxamine, cimetidine etc.	Blood **erlotinib concentrations may be increased**, which may lead to increased incidence and severity of adverse drug reactions. Co-administration of **itraconazole** resulted in an 80% increase in the AUC of erlotinib.
warfarin, cumarin (Marcumar)	There were reports on **INR elevations and bleeding events** on co-administration with warfarin. When erlotinib is used concomitantly with warfarin, the patient should be monitored for **prothrombine time or INR** at regular intervals.

8.3.6 Compliance

Erlotinib was given to the patient by the treating physician at d1. The treatment was ambulatory. If dose reduction was necessary, the patient would get an immediate visit and would receive the lower dosage according to the dosage scheme. Unused tablets were returned to Roche. The drug accountability was documented by LCGC. Empty blisters of erlotinib medication were counted and monitored as well to assure drug accountability and pharmacovigilance.

Bevacizumab was administered intravenously and supervised by qualified LCGC staff at d1, d22 and d43. This took place in the outpatient clinic of the Department I of Internal

Medicine, University Hospital of Cologne, or, regarding the physical state of the patient, in a ward setting. Any unused portion left in a vail was discarded, as the product contains no preservatives. The drug accountability was conducted and documented by LCGC in cooperation with the pharmacy.

The documentation of compliance and the assessed drug count is archieved in the pharmacy. In general, no protocol deviations regarding the used substances could be detected throughout termination of the trial.

8.4 Objectives for efficacy and safety

8.4.1 Assessments for primary objectives

As this trial was set out in order to define imaging biomarkers and biological biomarkers which may discriminate between clinical benefit, imaging and biomarker focus came into focus. The PET analyses were performed as follows:

For FDG-PET, patients had to refrain from ingestion for 6 hours prior to the PET examination; only the ingestion of carbohydrate-free drinks was permitted. The current glucose level was measured before starting PET. An uncontrolled diabetes mellitus with Hb1Ac >7,5% or elevated blood glucose levels of > 200 mg/dL were considered exclusion criteria. Optionally regular insulin (e.g. 2 units) was allowed to be administered if blood sugars were unexpectedly high, provided that monitoring and emergency care are assured. For FLT-PET, no special preparation was necessary.

18F-FDG and 18F-FLT were synthesized at the Max-Planck-Institute for Neurological Research at the University of Cologne following standard operating procedures which were submitted to BfArM and approved in 2007, or provided as commercial available (only FDG). A dose of 370 MBq FDG and 300 MBq FLT were injected intravenously for whole-body examination by PET. Scanning commenced 60 minutes after injection of tracers. PET instruments with full-ring detectors were applied. Emission data were corrected for attenuation, scatter and random and will be reconstructed to plains of 3.125 mm width and 128x128 voxels. For quantification of tracer uptake, standard methods, such as standardized uptake values (SUV) were applied.

In the protocol, there is stated that optionally, dynamic PET scans could be assessed. This was not performed due to logistic issues. Table 4 gives the a-priori assumptions about the effective dosages of radioactivity per patient, whereas Table 5 lists the exact activities used in the individual patients in this trial. Due to legal issues, Table 5 was provided to the BfS in unanonymised form (file no. Z5-22463/2-2010-013). All PET-procedures took place and was analysed by the Institute of Nuclear Medicine, University Hospital of Cologne.

Table 4: Preplanned effective dosages for PET-procedures

Total activity for the respective RP	ED due to application of the RP	ED due to application of ALL the RP's	ED of all TM's	Allover ED
Max. 370 MBq FDG Day -10 –Day 1	7,0 mSv	7 mSv	0,1 mSv	**7,1 mSv**
Max. 300 MBq FLT Day -10 –Day 1	9,9 mSv (f) 8,4 mSv (m)	16,9 mSv (f) 15,4 mSv (m)	0,1 mSv	**17,0 mSv (f)** **15,5 mSv (m)**
Max. 370 MBq FDG Day 7 – Day 14	7,0 mSv	23,9 mSv (f) 22,4 mS (m)	0,1 mSv	**24,0 mSv (f)** **22,5 mSv (m)**
Max. 300 MBq FLT Day 7 – Day 14	9,9 mSv (f) 8,4 mSv (m)	33,8 mSv (f) 30,8 mSv (m)	0,1 mSv	**33,9 mSv (f)** **30,9 mSv (m)**
Max. 370 MBq FDG Week +7	7,0 mSv	40,8 mSv (f) 37,8 mSv (m)	0,1 mSv	**40,9 mSv (f)** **37,9 mSv (m)**
Max. 300 MBq FLT Week +7	9,9 mSv (f) 8,4 mSv (m)	50,7 mSv (f) 46,2 mSv (m)	0,1 mSv	**50,8 mSv (f)** **46,3 mSv (m)**

Table 5: Applied activity for all patients

ID	Applied activity (MBq)							
	18F-FDG				18F-FLT			
	Appl. 1	Appl. 2	Appl. 3	Sum (RP 1)	Appl. 1	Appl. 2	Appl. 3	Sum (RP 2)
1-04	350 (18.03.2010)	366 (01.04.2010)	355 (04.05.2010)	1071	352 (19.03.2010)	158 (31.03.2010)	333 (05.05.2010)	844
1-	401	346	271	1018	392	367	278	1037

12	(27.07.2010)	(17.08.2010)	(17.09.2010)		(28.07.2010)	(13.08.2010)	(15.09.2010)	
1-10	376 (29.06.2010)	350 (13.07.2010)	370 (17.08.2010)	1096	330 (30.06.2010)	290 (14.07.2010)	286 (18.08.2010)	906
1-09	300 (25.06.2010)	370 (08.07.2010)	362 (10.08.2010)	1032	333 (23.06.2010)	370 (07.07.2010)	269 (13.08.2010)	972
1-08	243 (16.06.2010)	306 (29.06.2010)	371 (03.08.2010)	920	257 (18.06.2010)	326 (30.06.2010)	358 (04.08.2010)	941
1-07	374 (04.05.2010)	315 (18.05.2010)	370 (22.06.2010)	1059	322 (07.05.2010)	337 (19.05.2010)	338 (23.06.2010)	997
1-06	351 (27.04.2010)	282 (11.05.2010)	390 (15.06.2010)	1023	312 (28.04.2010)	354 (07.05.2010)	139 (16.06.2010)	805
1-03	372 (18.03.2010)	356 (01.04.2010)	358 (04.05.2010)	1086	342 (19.03.2010)	205 (31.03.2010)	343 (12.05.2010)	890
1-13	327 (19.08.2010)	305 (26.08.2010	316 (04.10.2010)	948	422 (18.08.2010)	n.d.	n.d.	422
1-02	386 (19.01.2010)	367 (01.02.2010)	380 (09.03.2010)	1133	373 (20.01.2010)	277 (29.01.2010)	350 (10.03.2010)	1000
1-01	284 (19.01.2010)	300 (28.01.2010)	306 (05.03.2010)	892	325 (20.01.2010)	332 (03.02.2010)	254 (10.03.2010)	911
1-11	374 (13.07.2010)	371 (23.07.2010)	269 (02.09.2010)	1014	214 (14.07.2010)	310 (21.07.2010)	n.d.	524
1-05	375 (26.03.2010)	369 (06.04.2010)	289 (10.05.2010)	1033	263 (24.03.2010)	379 (07.04.2010)	331 (12.05.2010)	973
1-14	334 (16.09.2010)	273 (28.09.2010)	300 (02.11.2010)	907	324 (15.09.2010)	276 (29.09.2010)	345 (05.11.2010)	945
1-20	334 (20.01.2011)	260 (03.02.2011)	269 (10.03.2011)	863	280 (19.01.2011)	198 (02.02.2011)	295 (09.03.2011)	773
1-17	290 (19.11.2010)	326 (30.11.2010)	361 (04.01.2011)	977	330 (17.11.2010)	284 (01.12.2010)	304 (05.01.2011)	918
1-18	269 (10.12.2010)	348 (17.12.2010)	n.d.	617	338 (08.12.2010)	358 (22.12.2010)	n.d.	696
1-16	262 (04.11.2010)	324 (18.11.2010)	274 (21.12.2010)	860	246 (05.11.2010)	261 (17.11.2010)	292 (22.12.2010)	799
1-19	231 (14.12.2010)	204 (23.12.2010)	192 (03.02.2011)	627	276 (15.12.2010)	276 (22.12.2010)	177 (02.02.2011)	729

1-15	284 (21.10.2010)	248 (04.11.2010)	258 (09.12.2010)	790	310 (22.10.2010)	183 (05.11.2010)	310 (08.12.2010)	803
1-21	307 (18.01.2011)	295 (03.02.2011)	211 (08.03.2011)	813	303 (19.01.2011)	189 (02.02.2011)	310 (09.03.2011)	802
1-22	268 (01.03.2011)	305 (17.03.2011)	203 (26.04.2011)	776	233 (02.03.2011)	181 (16.03.2011)	342 (27.04.2011)	756
1-23	270 (01.03.2011)	294 (11.03.2011)	230 (26.04.2011)	794	281 (02.03.2011)	272 (16.03.2011)	314 (20.04.2011)	867
1-26	319 (27.05.2011)	282 (03.06.2011)	275 (13.07.2011)	876	314 (25.05.2011)	280 (01.06.2011)	180 (20.07.2011)	774
1-28	316 (09.06.2011)	270 (24.06.2011)	286 (21.07.2011)	872	297 (01.06.2011)	213 (29.06.2011)	280 (27.07.2011)	790
1-24	243 (25.02.2011)	208 (17.03.2011)	180 (21.04.2011)	631	287 (02.03.2011)	167 (16.03.2011)	253 (20.04.2011)	707
1-25	200 (26.05.2011)	218 (03.06.2011)	279 (15.07.2011)	697	256 (25.05.2011)	235 (01.06.2011)	307 (20.07.2011)	798
1-27	287 (19.05.2011)	203 (03.06.2011)	222 (14.07.2011)	712	275 (25.05.2011)	305 (01.06.2011)	117 (13.07.2011)	697
1-29	286 (24.06.2011)	229 (05.07.2011)	174 (09.08.2011	689	n.d.	n.d.	n.d.	0
1-31	24.06.2011	15.07.2011	26.08.2011		06.07.2011	13.07.2011	24.08.2011	
1-33	327 (05.07.2011)	271 (29.07.2011)	273 (02.09.2011)	871	254 (20.07.2011)	271 (27.07.2011)	276 (31.08.2011)	801
1-34	206 (22.07.2011)	253 (02.08.2011)	310 (09.09.2011)	769	309 (27.07.2011)	183 (03.08.2011)	204 (07.09.2011)	694
1-36	261 (05.08.2011)	233 (12.08.2011)	250 (16.09.2011)	744	190 (03.08.2011)	221 (17.08.2011)	220 (21.09.2011)	631
1-37	12.08.2011	295 (26.08.2011)	325 (30.09.2011)		200 (10.08.2011)	94 (24.08.2011)	280 (05.10.2011)	574
1-39	322 (21.10.2011)	300 (03.11.2011)	270 (09.12.2011)	892	n.d.	n.d.	n.d.	0
1-40	318 (18.11.2011)	275 (29.11.2011)	270 (03.01.2012)	863	325 (16.11.2011)	262 (30.11.2011)	221 (04.01.2012)	808
1-41	301 (02.12.2011)	329 (22.12.2011)	305 (24.01.2012)	935	249 (07.12.2011)	322 (21.12.2011)	313 (25.01.2012)	884

1-42	261 (12.12.2011)	200 (05.01.2012)	n.d.	461	181 (21.12.2011)	n.d.	n.d.	181
1-44	308 (14.02.2012)	354 (06.03.2012)	359 (13.04.2012)	1021	360 (22.02.2012)	304 (07.03.2012)	277 (11.04.2012)	941
1-46	295 (28.02.2012)	198 (20.03.2012)	259 (20.04.2012)	752	340 (29.02.2012)	n.d.	n.d.	340

The DCE-MRI quantification was performed in the Institute of Radiology, University Hospital of Cologne, and the Multimodality Imaging Group of the Max-Planck-Institute of Neurological Research. The acquisition sequence had to be suitably linear with Gd-DPTA concentration. The quantification was made by using dynamic parameters such as SD, RE and dR. It was performed without assuming a relaxivity value. For analysis of the DCE-MRI images, K_{trans} and k_{ep} were measured. For baseline imaging, the first DCE-MRI imaging was repeated in a subset of patients before the start of study treatment. We provided a separate working manual for the assessment of DCE-MRI-imaging, written in advance in order to warrant reproducibility.

The patients first received standard MRI-reconstructions (i. e., T1 and T2 weighted imaging) in order to localize the tumour. Then, Gd-DPTA was injected by an investigator for DCE-measuring. Breath-correction and primary reconstruction was done using software purchased from Siemens.

Table 6: Overview of the exploratory imaging procedures at the according timepoints

Timepoint	Days	Modality
Baseline	d-14 – d1 (prior to therapy)	FDG-PET, FLT-PET, 2x DCE-MRI
Week 2	d7 – d14	FDG-PET, FLT-PET, DCE-MRI,
Week 7	d44-d49	FDG-PET, FLT-PET, DCE-MRI

Mutational analyses using next-generation sequencing (NGS) techniques were performed if material was still available as recently described (König et al., J Thorac Oncol. 2015 Jul;10(7):1049-57). Figure 3 is extracted of this publication and shows schematically the steps for amplicon preparation.

Figure 3: Library preparation for NGS

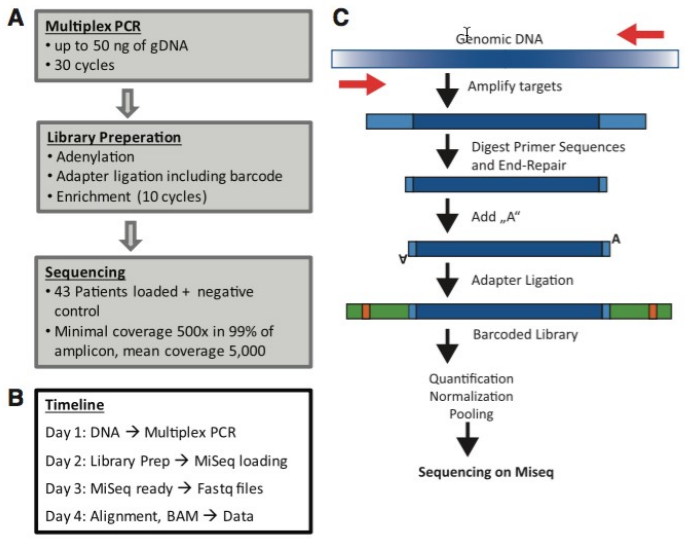


FIGURE 1. Schematic presentation of the library preparation. *A* and *C*, Genomic DNA (up to 50 ng) is amplified with the AmpliSeq Custom Panels. Primers are digested with FuPa, an "A" is added at the 3´ end to allow adapter ligation. After bead clean up and size selection, the successfully ligated products are amplified by PCR with a final bead clean up. Samples are quantified by Qubit, normalized to 10 nM, and pooled to be then sequenced on v2 MiSeq Cartridge. A table outline (*A*) and a cartoon (*C*) are depicted. *B*, Mutational analysis can be completed within 4 working days starting with the DNA extraction until data analysis. PCR, polymerase chain reaction.

Microvessel density was assessed by scanning the area of highest vessel density at low power magnification. Individual microvessels were then counted on a 200× field (0.80 mm^2 per field). Any endothelial cell cluster, which was positive for CD31 and contained a visible lumen was considered a single countable microvessel. For each section, three fields were counted, and the mean value calculated. This value was converted to the MVD count, expressed as vessels/mm^2. This method was introduced by and performed in accordance to Weidner et al., N Engl J Med, 324 (1) (1991), pp. 1–8.

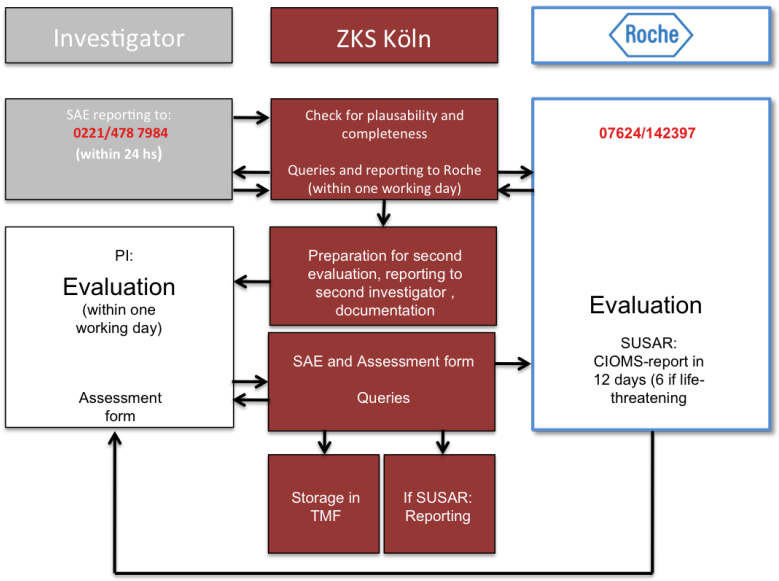
For assessment of VEGF-A and VEGFR expression, all sections were examined by light microscopy for the presence of expression and cellular distribution of the proteins (between the cell membrane, cytoplasm and nucleus) in the endometrial adenocarcinomas. Cell staining intensity was scored as negative (0), weak (1+), moderate (2+) and strong (3+). Tumors showing antigen expression showed intratumoral heterogeneity in the intensity of staining. For each case, the percentage of cells with the predominant staining intensity was estimated. For statistical purposes, cases which were moderately (2+) or strongly (3+) positive in more than 10% of cells were designated overall positive, while cases in which expression was weak (1+) or seen in less than 10% of cells were considered overall negative. These methods were adapted from Wang et al., Cytokine. 2014 Aug;68(2):94-100.

For evaluation of immunostaining, specimens were evaluated by two observers who were unaware of the clinical features and outcomes of patients. For VEGF, VEGFR1, and VEGFR2, the extension was scored as the percentage of positive cells (0% to 100%), and the intensity of staining was assessed by comparison with a known external positive control (0, below the level of detection; 1, weak; 2, moderate; and 3, strong). Scores were calculated by multiplying the staining intensity and extension at each intensity level, as previously described.25 For VEGFR1 and VEGFR2, the combined expression of cytoplas-

mic and membrane staining was assessed. The sum of the VEGF, VEGFR1, and VEGFR2 scores rendered the combined variable termed VEGF signaling score (VSS). The microvessel density (MVD) scores were evaluated using a 25-point Chalkley eyepiece graticule, as previously described.26 Median values were used as the cutoff. For VSS, the cutoff point was 400 in the EUELC and CHU series and 460 in the MD Anderson series. Automatic immunohistochemical and semiquantification procedures were cross validated in 10% of the patients by an independent investigator.

Clinical study visits took place in the outpatient ward of Department I of Internal Medicine, University Hospital of Cologne. Adverse events were documented using Common Toxicity Criteria for Adverse Events (CTC AE v4.0). Survival times were calculated from start of treatment until event (progression or death). In censored cases, the time point of the last contact within the trial was documented. Concomitant medication was documented at every contact. Laboratory values were provided by the Institute of Clinical Chemistry of the University Hospital of Cologne. In case of abnormal laboratory findings, the values were graduated and documented according to CTC AE v4.0, and clinical significance was estimated by the investigator.

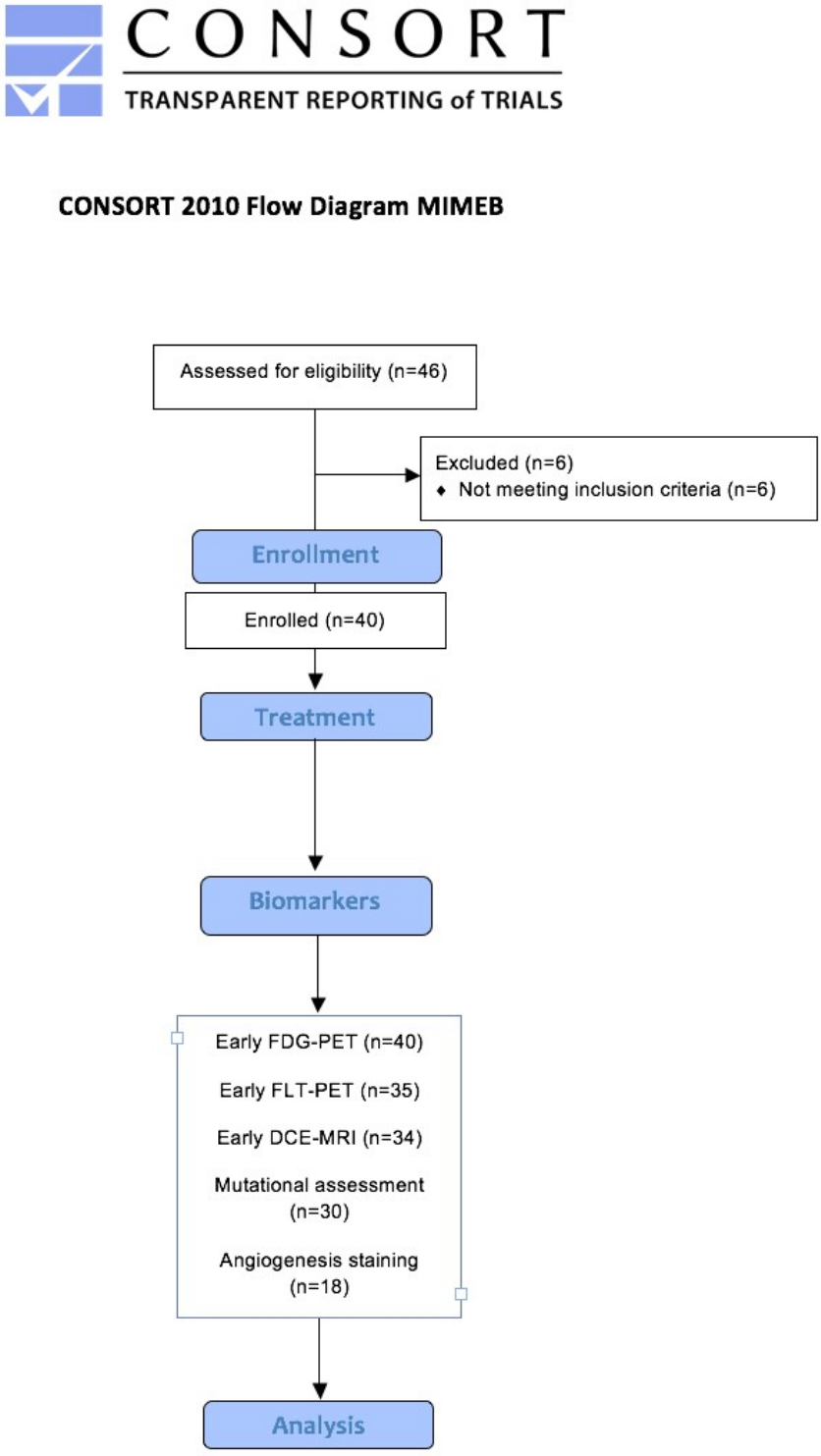
Figure 4: Reporting of SAEs in the trial



CT scans were performed in a routine fashion in the Institute of Radiology of the University Hospital of Cologne. After changing to PET-CT scanners, CTs could also be performed in addition to the PET acquisition. In these cases, CT scans were analyzed by a radiologist afterwards.

Pharmacokinetic assessments and phenotyping as proposed in the protocol were not performed, as it was not possible to set-up the fitting infrastructure for these analyses. Figure 4 shows a CONSORT-sheet describing the enrolled patients and the performed analyses.

Figure 5: CONSORT diagram for final analyses



8.4.2 Accuracy of endpoints

The endpoints for safety and efficacy were standard procedures and followed RECIST v1.1 and CTC AE v4.0.

For PET imaging, we first determined the highest value within the tumor (SUV_{max}). SUV_{peak} was estimated for each lesion by using a 1.2-cm diameter fixed sized circle centered around the tumor area with the highest uptake. This method was used within the ERLOPET trial as well (Zander et al., J Clin Oncol. 2011 May 1;29(13):1701-8). Further analyses of the data gained in ERLOPET proved that these SUV values represent the most robust and reliable methods in order to detect differences (Kahraman et al., J Nucl Med. 2011 Dec;52(12):1871-7, Scheffler et al., PLoS One. 2013;8(1):e53081). For FDG-PET, these methods have been proposed to detect metabolic response (PERCIST: Wahl et al., J Nucl Med. 2009 May;50 Suppl 1:122S-50S). For FLT-PET, such a recommendation does not exist.

The hypothesized effect of adding bevacizumab to the treatment with erlotinib is tumor-vessel normalization (Chatterjee et al., J Clin Invest. 2013 Apr;123(4):1732-40; Chatterjee et al., Cancer Res. 2014 May 15;74(10):2816-24). We therefore chose to take k_{trans} as the surrogate marker in DCE-MRI best fitting to the hypothesis, as it quantifies the efflux of the contrast agent out of the tumor vessels. This setting was completely exploratory, as DCE-MRI was considered not applicable for lung tumors due to breathing artefacts. The feasibility of performing DCE-MRI scans in patients with lung cancer.

In 2011, it became known that *EGFR* mutations are the best predictors of clinical benefit in patients treated with the EGFR inhibitor gefitinib (Fukuoka et al., J Clin Oncol. 2011 Jul 20;29(21):2866-74). Subsequently, EGFR inhibitors became the first-line treatment of choice in patients harboring these mutations. Hence, molecular testing as proposed in the protocol have become a standard procedure, at least with regards to *EGFR*. With regards to other mutations, few is known about the efficacy of erlotinib in these cohorts.

With new PD1/PD-L1 antibodies, there are new therapeutic options for patients with NSCLC. We added the highly discussed PD-L1 staining to the staining models focusing on angiogenesis.

We chose non-progression and PFS as the endpoint of prediction, as it has been shown that there are clinical courses of patients not achieving a response but have nevertheless benefit in terms of prolonged PFS. Therefore, the primary endpoints focus on these parameters.

Overall survival for first-line treated patients is strongly dependent on subsequent treatment; reductions in imaging predicting OS are therefore considered prognostic rather than predictive. The efficacy of a drug is best reflected by response and PFS; a fact that is strengthen by the fact that the FDA nowadays accepts PFS advantages in first-line treated patients for approval.

8.4.3 Primary endpoints

Percentaged changes (proposed reductions) in the SUVs of PET images after one week of therapy as compared with the baseline assessment were compared with the response

outcome (i. e., PD vs non-PD) in the first restaging procedure after six weeks of combined therapy. The related changes were analyzed using a responder-operator-characteristics (ROC) curve and the corresponding area under the curve (AUC).

The maximal Youden-indices found in these analyses were then taken as cut-off values for Kaplan Meier analyses regarding PFS. Further, we tested predefined cut-off values (20%, 30% reductions in PET activities) regarding their potential role in discriminating patients with benefit from therapy as seen in prolonged PFS.

8.5 Data quality

This trial was monitored by the ZKS Cologne. The monitoring manual and the monitoring reports documenting advance and proegress are attached to this report.

Complete comparison of the data entry form with the source data was performed in a randomly chosen part of the patients. The existence of the written informed consent was checked for any patient, as well as the inclusion and exclusion criteria. The exact amount and the exact kind of the monitoring is described in the monitoring manual. Within this manual, all study-specific issues were elucidated, and the minimum amount of monitoring activities were defined by forms. The amount of the monitoring was evaluated by using the ADAMON criteria, which has been established by ZKS Cologne and was granted by the BMBF.

Between the visits, the responsible monitor held regularly contact to the study center (LCGC).

All investigators confirm agreed with the regular visits in the study centres. The primary purposes and aims of these visits were:

- the evaluation of the study progress,
- the controlling of adherence with the study protocol,
- the discussion of problems, including AEs,
- the examinations of the CRFs regarding accuracy and completeness,
- the validation of the CRFs with the original data,
- the checking of study medication handling

The monitor had the right to compare the CRFs with the original data, having regard to the Data Protection Act (the monitor is bound to professional discretion). The investigators offered the monitor a direct access to the original data and sources.

Each investigator was responsible at his site that the trial is performed in accordance with GCP, AMG and the protocol. He or she ensured that data was collected as stated in the protocol and documented correctly in the responding CRF. Monitors assisted investigators in reviewing the complete, legible, clearly arranged and recallable data.

During the termination of this trial, no trial-specific audits took place. Nevertheless, LCGC has been audited four times during the report period. Audit reports are attached to this report.

The study-relevant data was documented in the prepared CRFs promptly by the responsible investigator or a study assistant.

Source data were original data (ECG print, pulmonary function print, laboratory print etc.) and data from quantitative imaging tools (MPI-Tool for PET imaging, for example) and their prints.

All data received in the University Hospital of Cologne, were source data.

Within the CRFs, the staging of the tumor according to the TNM-system was source data. The results of the clinical examination might have been source data as well.

PET- and MRI-imaging results had to be reported in summary in the CRF. The original reports, traces and films retained by the investigator for future reference. The sponsor provided a study specific template for MRI evaluation, which was considered as source data.

The investigator reviewed all pages within the CRF for accuracy and consistency with the protocol, and sign and date the CRF sign-off page(s) upon completion.

The patient consent forms designated for the clinical investigator were also kept in the ISF. All information on which the entries in the CRF were based are available in the patient files e. g. results of laboratory investigations.

If data entry had to be corrected, the correction was signed and dated by the responsible person. Corrections to data on the CRFs were made by lining out the incorrect data with a single line and writing the correct data near to those crossed out.

A multilevel plan for data validation and data management was provided by the ZKS Cologne and is attached to this report. In case of missing or inconsistent data, a data manager contacted the treating physician or person responsible for documentation in the department by telephone or in writing (queries).

In the documentation constitution (ZKS Cologne), all documentation forms and CRFs were registered and checked for completeness. The data was admitted and aligned doubly by two independent data admissioners in a validated study database. Additionally, there were checks of plausibility. All discrepancies in plausibility were clarified with the study centre in written form.

The central IT-infrastructure was offered by the ZKS Cologne.

The study database based on a validated study software (MACRO). The system was validated before data admission. All changes of the data were documented and saved in an audit trail. It also had a study-specific user and role concept. The database is integrated in a general IT-infrastructure system with firewall and backup system. The data is saved daily.

All CRFs, written informed consents and essential trial documents were archived in accordance with GCP-Ordinance §13(10) for at least 35 years due to regulations of the BfS. The patient identification list is stored separated from documentation forms. Case records and raw data are to be retained for at least 35 years following the end of the trial. The investigator ensures that a correct assignment of the CRFs to the corresponding patient files and raw data is possible at any time.

A copy of the final completed CRFs is retained by the investigator, who ensures that it is stored with other trial documents, such as the protocol, the investigator's brochure and any protocol amendments, in a secure place.

Advance and progress of the trial was discussed once weekly with the PI and the investigators who signed their presence in a presence list. Investigator's training was performed continuously, as all investigators shared one room (study center). All investigators had a "Prüfarztkurs" prior to trial participation.

The author of this report was funding member of the SOP working group of the ZKS Cologne and the quality safety group of the Department I of Internal Medicine, University Hospital of Cologne. The SOP group realized the goal to provide SOPs for any aspect of GCP-conformity in the termination of clinical trials, whereas the quality safety group implemented these SOPs in the Department I of Internal Medicine. This was finished during the termination of the present trial, and the trial used the specific SOPs.

8.6 Statistics according to the protocol

8.6.1 Planned analyses

The evaluation of imaging data will be measured by ROC analysis using SUV and percentual changes of SUVs at different time points of FDG-/FLT-PET. Statistical association of molecular status and clinical response (Fisher's exact test) will also be performed. The final statistical report will be done by PD Dr. M. Hellmich and is included also in the final report of the study.

Raw data of PET will be analyzed by visual control and semiquantitative (Standardized Uptake Value, SUV) methods using the MPI-tool. PET-tomograms will be analyzed in all three planes (coronar, sagital, transversal). Data will be correlated with computed tomography. Data will be stored for direct comparison in the follow-up. DCE-MRI data will be analyzed by evaluating Ktrans, Ve, Kep, BF, BV, PS and MTT.

Raw data of CT will be analysed according to the RECIST criteria.

The ROC analysis will be performed nonparametrically according to DeLong et al., Biometrics, 1988. 44(3): p. 837-45.

Based on the results of clinical trials with erlotinib and bevacizumab in NSCLC patients in Western Europe and the US, we expect at about 20% of patients with response (CR, PR) to treatment with erlotinib and bevacizumab, about 40% with stable disease and about 40% with progressive disease. Furthermore, we expect about 90% of patients with mutated EGFR responding to therapy. KRAS mutations (15-50% of patients) and EGFR mutations (2-25% of patients) are expected to be nearly exclusive.

Based on published PET data, especially from patients with GIST treated with imatinib, we expect a nearly 100% association of clinical response (CR, PR in CT after 6 weeks of treatment) and reduction in SUV in PET analyses after one week of therapy.

We also expect a predefined cut-off value for metabolic response (a reduction of at least 20% or 30% in sSUV for FDG-PET and FLT-PET after one week of therapy) to be predic-

tive for progression-free survival. We will calculate Kaplan Meier (curve) estimations regarding these cut-off values as well as at the value maximising the Youden index.

For DCE-MRI, a predictive/prognostic value at baseline imaging might be assumed. We will therefore exploratively assess multiple cut-off values regarding differences in PFS and OS as well as RR. The same will be done for FDG- and FLT-PET.

For the expected patient population, similar response rates as described in published studies for Western World patients are expected.

Response criteria are according to the Response Evaluation Criteria in Solid Tumors (RECIST) (http://www3.cancer.gov/bip/RECIST.htm).

The analysis of the data will be based on the following definitions:

- *Complete Response (CR):* Disappearance of all target lesions
- *Partial Response (PR):* At least a 30% decrease in the sum of the longest diameter (LD) of target lesions, taking as reference the baseline sum LD
- *Progressive Disease (PD):* At least a 20% increase in the sum of the LD of target lesions, taking as reference the smallest sum LD recorded since the treatment started or the appearance of one or more new lesions
- *Stable Disease (SD):* Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum LD since the treatment started
- *Non-progression:* CR + PR + SD but not PD

Analyses will be performed on three study populations:

Intention-to-treat population (primary analysis): All patients included in the study will be analysed (contingent on availability of endpoints).

Per-protocol population (secondary analysis): Only patients fully compliant with the protocol (especially regarding imaging procedures) will be analyzed.

The full analysis set will be taken as *safety analysis population (tertiary analysis)* as all patients will receive the study medication.

To evaluate the accuracy of imaging findings of the parameters mentioned above in FDG-/FLT-PET and DCE-MRI after one week of treatment for early prediction of RECIST-based non-progression (CR+PR+SD) after 6 weeks of therapy in patients with NSCLC stage IIIb/IV treated first line with erlotinib and bevacizumab.

The predictive value/accuracy of FDG-/FLT-PET and DCE-MRI regarding non-progression in NSCLC patients treated with erlotinib and bevacizumab after 6 weeks will be evaluated by ROC analysis of percentual changes of (semi-)quantitative parameters (see parameters listed above; area under the ROC curve, AUC); H_0: $AUC_{FDG/FLT\ PET/DCE\text{-}MRI}$ 0.5, H_A: $AUC_{FDG/FLT\ PET/DCE\text{-}MRI} > 0.5$ (one-sided level 2.5% , not adjusted for multiple testing).

To evaluate the accuracy of imaging findings in FDG-/FLT-PET and DCE-MRI after one week of treatment for early prediction of PFS in patients with NSCLC treated first line

with erlotinib and bevacizumab. Cut-offs maximizing the Youden index for each method will be used for Kaplan-Meier (curve) estimations.

Additionally, predefined percentual reductions of (semi-)quantitative imaging parameters will be analysed:

- reductions of at least 10%
- reductions of at least 20%
- reductions of at least 30%
- reductions of at least 40%
- reductions of at least 50%

Further reduction values can be analysed depending on ROC analyses.

To evaluate differential results concerning the predictive value of the PET analysis between male and female patients, a separate analysis for both groups will be performed. Additional subgroup analyses will be performed regarding the mutational status of the patient, histology, the ECOG status, clinical signs and symptoms as well as baseline molecular marker characteristics and smoking history.

Interim analyses will be performed after 10, 20 and 30 patients. O'Brien-Fleming boundaries (calculated with ADDPLAN 5 MC) will be used to assess statistical significance, i.e. critical limits 4.049, 2.863, 2.337 and 2.024 (one-sided significance levels 0.000026, 0.0021, 0.0097, 0.0215) will be employed.

The interim analysis after 20 patients will include the evaluation of data reliability/reproducability of DCE-MRI.

8.6.2 Sample-size calculation

Assumptions for response prediction:

(1) 20% responder, 40% progressors; thus 60% non-progressors

(2) Primary variable is the 'area under the ROC curve (AUC, ϑ)'

(3) Type-I-error 0.043; type-II-error 0.20

(4) True accuracy (AUC) is at least 0.70 (alternative hypothesis; some common diagnostic tests have an AUC greater than 0.70).

The sample size in the following table was calculated using formulae (6.3) and (6.6) in Zhou et al., Statistical Methods in Diagnostic Medicine. 2002, New York: Wiley. Thus, assuming $\vartheta_1 = 0.75$, about 40 patients are needed to reject the null hypothesis H_0: $\vartheta = 0.5$ in favour of the alternative H_1: $\vartheta \neq 0.5$. The sample size of 40 will be sufficient to yield a 95% confidence interval for ϑ with approximate width 2 0.18 = 0.36.

Table 6 shows the sample size for different considerations of the AUC.

Table 7: Statistical considerations I

AUC (ϑ)	# Non-progressors	# Progressors	# Total
0.70	39	27	66
0.75	25	17	42
0.80	17	11	28

Assumptions for PFS prediction:

(1) total sample size=40

(2) alpha=0.05, two-sided

(3) median PFS non-progressors=6.5 months

(4) median PFS progressors=1.5 months

(5) accrual period=12 months; additional follow-up period=3 months)

Table 8: Statistical considerations II

Non-progressors, n	progressors, n	Power*
16	24	0,976
12	28	0,957
8	32	0,895

*according to Dupont et al., Control Clin Trials, 1990. 11(2): p. 116-28

8.7 Changes in design or statistics

The mentioned interim analyses were not performed, as DCE-MRI data was not available before 2013.

Further, some of the additional analyses, i. e. pharmacokinetics or phenotyping were not performed due to the lack of a functional infrastructure regarding these issues at study start and the missing possibility to set this up when this trial had started. As this was in favor for the enrolled patients (less stress, less invasive measurements), we did not amend these changes.

Some of the parameters listed in the protocol to be determined in DCE-MRI were not performed until today, as we did not prioritize them over k_{trans}.

At ASCO 2011, there was an interim report of the German INNOVATIONS trial comparing the combination used in MIMEB with standard chemotherapy in first-line NSCLC St. IV patients (in the meantime published: Thomas et al., Eur Respir J. 2015 Jul;46(1):219-29.) Here, the colleagues demonstrated a clear advantage of chemotherapy first-line vs erlotinib and bevacizumab in terms PFS and OS in genetically unselected patients. This was in contrast to the second-line data which we used as the rationale for MIMEB (Herbst et al., J Clin Oncol. 2007 Oct 20;25(30):4743-50). Even worse, the colleagues reported a response rate of only 25% in patients with *EGFR*-mutations, suggesting a negative effect on EGFR-targeting by the addition of bevacizumab. We therefore discussed the issue in a grand round and decided to proceed with the trial, as a) data was still preliminary (not cleaned), b) had some surprising findings, reporting for example an OS for chemotherapy which was far beyond of those reported in phase III trials c) the primary objective of our trial was not to show a general efficacy of the combination, but to identify non-invasive methods in order to early identify patients benefitting from therapy.

We did not amend to include PD-L1 staining in the expression analyses.

Individual deviations from proposed endpoints regarding the imaging analyses at week 7 were analyzed by the investigators and not considered of harm for neither the patients nor the primary endpoints. Nevertheless, the proposed time intervals were widened in order not to lose information derived from the imaging procedures.

Beside the above-mentioned issues, no deviation from the initial plan took place. Analyses were performed in exact the same manner as proposed in the protocol. Due to the exploratory character of this trial, there will be additional analyses from the data-set beyond this final report.

9 Trial population

9.1 Patients enrolled

As preplanned, we enrolled 40 patients for the ITT analyses. One patient was excluded subsequently after start of therapy because his tumor turned out to be a malign melanoma (he continued with the medication afterwards due to a partial response). Five patients received a screening number and signed informed consent, but did not meet inclusion criteria before receiving medication. Therefore, there were 46 patients screened for the trial, and 40 of them were enrolled correctly (see figure 5). The descriptive statistics for the 40 patients was done by Prof. Dr. Martin Hellmich and Dr. Hildegard Christ and is attached entirely (MIMEB_Descriptive_Statistics 2016-03-23).

Of the 40 patients enrolled, 24 (60.0%) were male and 16 (40.0%) were female. One female patient was black with Nigerian origin; one female patient was Turkish. The remaining patients were considered Caucasian/white.

Until date of this report, 39 patients have died. One female patient was officially taken out of the study due to an intended therapy pause, but is still in close contact to the investigators.

For details, see MIMEB_Descriptive_Statistics 2016-03-23.

9.2 Protocol deviations

All protocol deviations are listed in Table 18 of "998_Analysis_of_endpoints" (attached). The check for inclusion and exclusion criteria per patient is provided by the attached "Listing_Incexcl" document.

In four patients (1-01, 1-24, 1-25, 1-27), there were laboratory deviations in baseline assessments. Each case was discussed by the PI and the investigators, and all patients remained on study treatment. The note-to-files are attached to this report.

In two post hysterectomy female patients (1-31, 1-33), pregnancy exclusion was not documented properly in the CRFs. These patients remained on study treatment.

In two patients, the gap between screening laboratory assessment and start of medication was too long due to organizational issues. As the assessed values were in range and there was no clinical suspicion that this might have changed, the patients remained on study treatment.

Patient 1-15 had received chemotherapy prior to study inclusion. This has been addressed in a note-to-file (attached) that this treatment was initiated in a lower tumor stage, the metastases which led to the stage were detected during this treatment. We therefore argued that the patient did not receive proper systemic stage IV treatment until enrollment to the trial. The patient remained on study treatment.

Patient 1-17 had a history of thrombosis which came into knowledge of the investigators during treatment. Because of profound clinical benefit for this patient, the sponsor decided to maintain the patient in the trial.

Patient 1-42 had progressive disease very quickly after initiation of therapy. In his week 2 FDG-PET, massive pleural effusion leading to the response of a PD, subsequent imaging procedures have not been performed afterwards. We discussed an exclusion of this patient from the analysis as the primary objectives (prediction of non-progression/prediction of PFS, both in the future) were compromised by the fact that the assessment of these parameters were already achieved by the PET-scan itself. Nevertheless, as the FDG week 2 changes were documented, the patient remained in the ITT analysis.

Missing or deviating MRI procedures were documented properly but not considered protocol deviations, as a secondary objective of the trial was the assessment of feasibility of DCE-MRI diagnostics in lung cancer. Details of the imaging time points and results are given in "Listing_mrifdg" (attached).

Medication was provided according to the protocol, and dose changes were performed in accordance with the IBs of the drugs in the current versions. Dose modifications are listed in Section 002 of "MIMEB_Descriptive_Statistics".

10 Efficacy analyses

10.1 Patient collectives

We here provide analyses of the ITT collective of the trial, which equals the "safety population. The following analyses in compliance with the protocol are attached and taken into account for this report:

- 998_Analysis_of_endpoints: Endpoint analyses for the whole ITT population focusing on SUV_{max} values.
- 998b_Analysis_of_endpoints: Endpoint analyses for female patients in the ITT population focusing on SUV_{max} values.
- 998c_Analysis_of_endpoints: Endpoint analyses for male patients in the ITT population focusing on SUV_{max} values.
- 998d_Analysis_of_endpoints: Endpoint analyses for the whole ITT population focusing on SUV_{peak} values.
- 998e_Analysis_of_endpoints: Endpoint analyses for female patients in the ITT population focusing on SUV_{peak} values.
- 998f_Analysis_of_endpoints: Endpoint analyses for male patients in the ITT population focusing on SUV_{peak} values.
- MIMEB_Descriptive_Statistics: Descriptive statistics for the whole ITT population.

10.2 Demographics and descriptive statistics

For details, please confer to "MIMEB_Descriptive _Statistics. As mentioned in 9.1, 24 of the 40 analyzed patients were male and 16 were female. All patients were enrolled in Germany and treated in the University Hospital of Cologne. Table 9 summarizes the baseline characteristics of the patients:

Table 9: Baseline characteristics of ITT population

Characteristics	No.
Age, years	
median	60
range	30-76
Sex	
female	16
male	24

Ethnic Group	
caucasian/white	38
black	1
turk	1
others	0
ECOG PS	
0	31
1	8
2	1

All patients had non-squamous cell NSCLC at first diagnosis. Uncontrolled diabetes mellitus was an exclusion criterion because of FDG-PET procedures, nevertheless, blood glucose levels were recorded immediately before FDG application. None of the patients therefore had diabetes. A listing of medical history and baseline characteristics is given per patient in "Listing_Screenstudmed" (attached).

Chronic or acute renal failure was an exclusion criterion because of the use of Gadolinium-DPTA for DCE-MRI assessment; few cases with preexisting renal failure (as defined by CrCl) developed systemic nephrogenic fibrosis (SNF). See note-to-file for patient 1-01.

Concomitant medication was recorded and documented throughout the course of the trial. Beside anticoagulation, no medication was prohibited explicitly. "Listing_Conmed" lists all concomitant medication documented during trial termination per patient.

10.3 Compliance

Compliance was assessed by counting the given erlotinib tablets as well as the returned ones and counting emptied blisters of trial medication. For bevacizumab, medication was prepared for intravenous administration in the outpatient ward of the Department I for Internal Medicine of the University Hospital of Cologne. As it for the intravenous nature of this therapy, compliance was directly warranted by supervising the infusion.

"Listing_Screenstudmed" shows the study medication administered individually. Reasons for dose reductions or discontinuations are listed there.

10.4 Results of efficacy analyses and tabular presentation of individual patient data

10.4.1 Efficacy analysis

Individual patient data regarding the primary objectives of this trial can be found in table form in Table 17 of the respective "998-998f_Analysis_of_endpoint documents".

Primary endpoint was the prediction of non-progression after six weeks of therapy. Table 10 lists the frequency of outcomes according to RECIST.

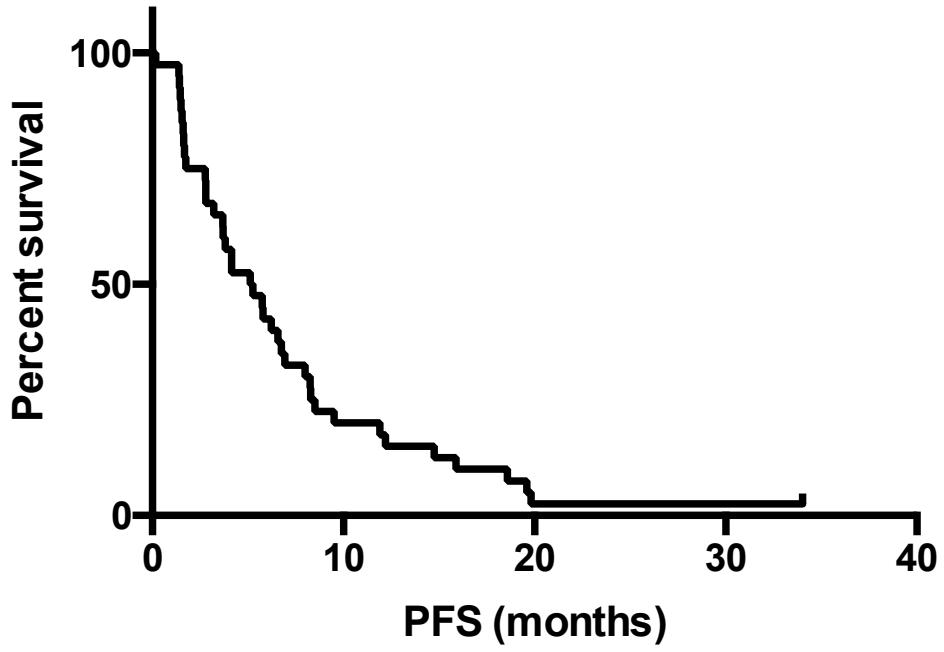
Table 10: Response in the ITT population

Response (RECIST)	No. (%)
Progressive Disease (PD)	16 (40.0)
Stable Disease (SD)	17 (42.5)
Partial Response (PR)	6 (15.0)
Complete response (CR)	1 (2.5)

Hence, 24 patients (60.0%) had non-progression (nonPD) as defined earlier, whereas 16 patients (40.0%) had PD.

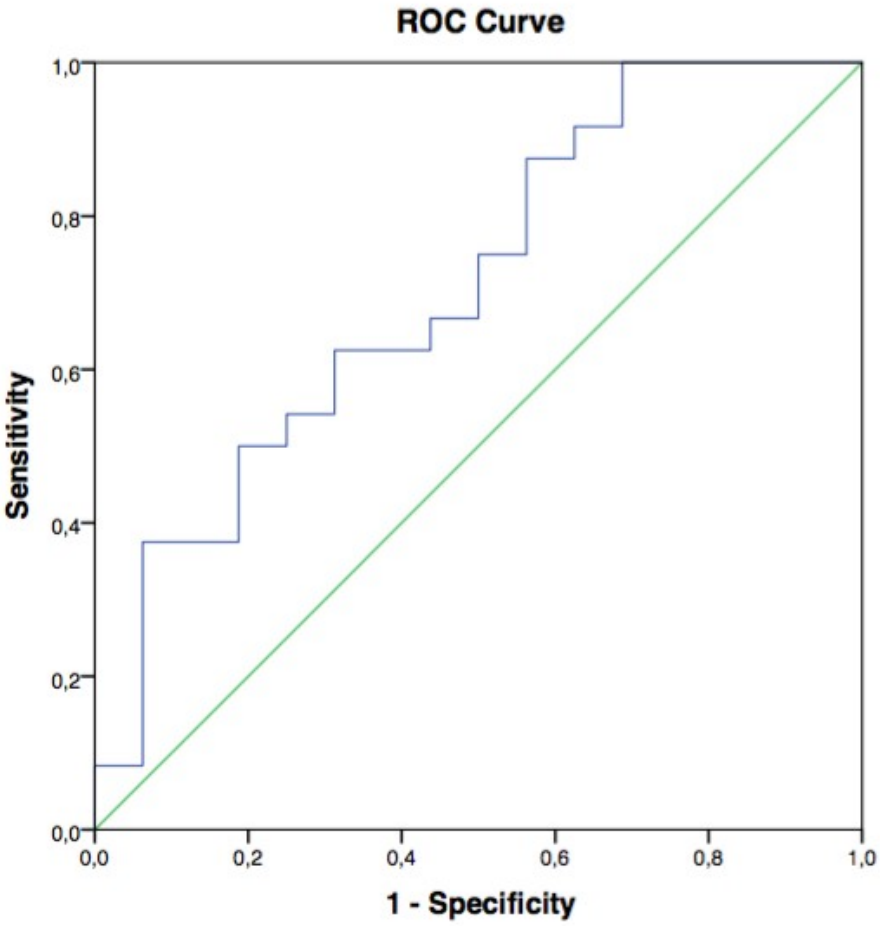
PFS was measured using Kaplan Meier statistics. Median PFS was 5.13 months (95% CI, 2.73-7.54). Figure 5 shows the Kaplan Meier curve for the ITT population.

Figure 5: PFS in the ITT population



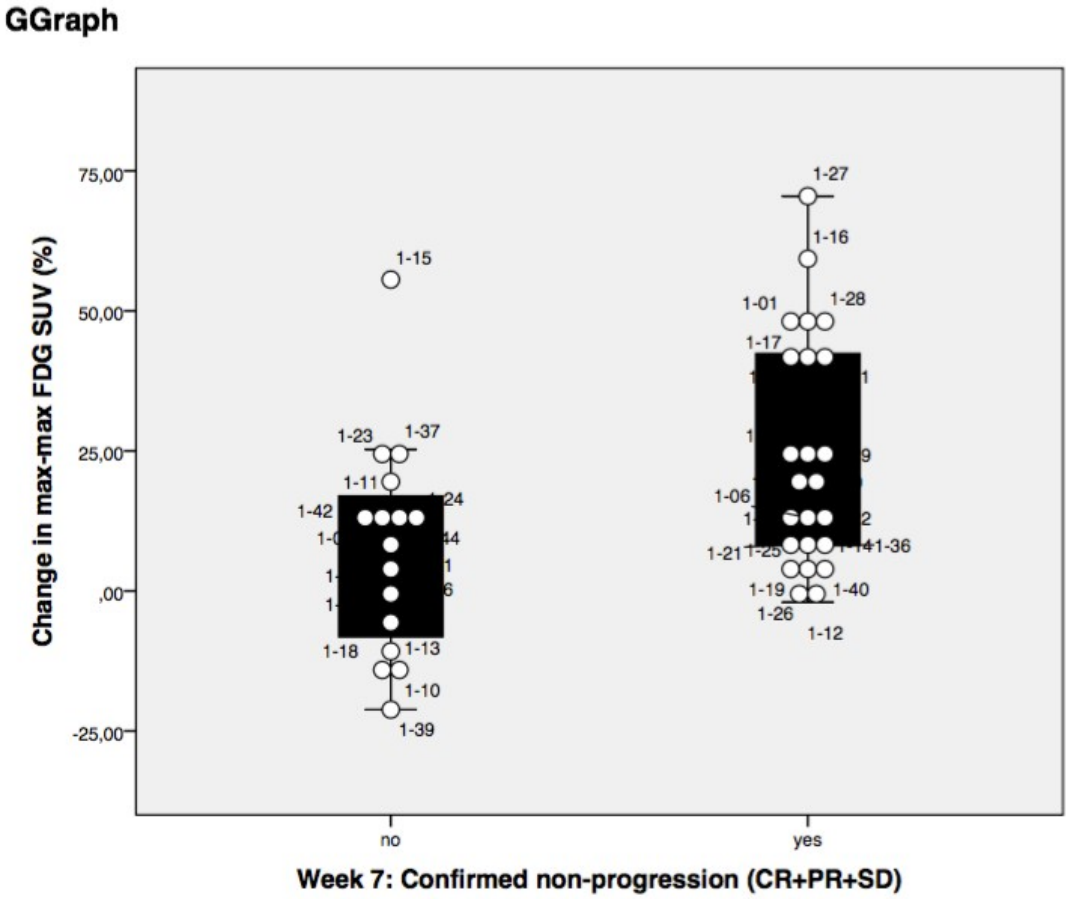
Prediction of non-progression was performed the same manner as in ERLOPET. Accordingly, FDG-PET (as measured with SUV_{max}) showed significant potential in predicting nonPD after one week of therapy, leading to an AUC of 0.71 (p=0.027). Figure 6 shows the ROC curve of the analysis.

Figure 6: ROC curve of early FDG-PET vs nonPD



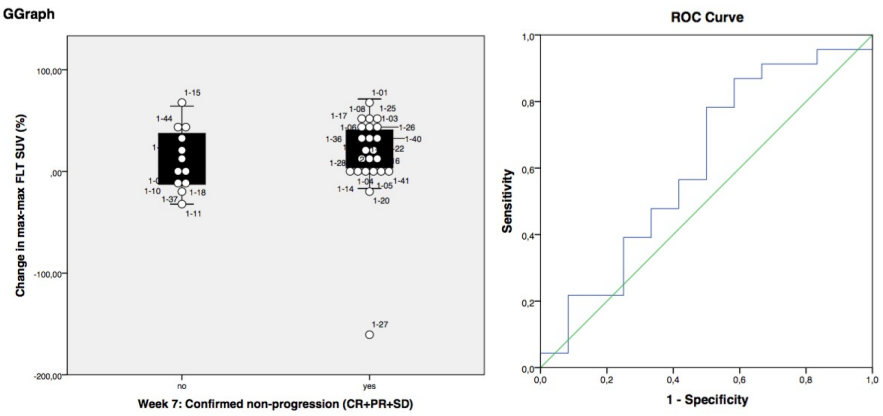
The differentiation between PD and nonPD was the basis for Figure 7, which demonstrates the changes in SUV_{max} from baseline examination until after one week of therapy. Whereas patients with primary PD had a mean reduction of SUV_{max} of 7.7% (SD, 19.2%), patients with nonPD had a reduction of 24.4% (SD, 20.6%) (t-test, p=0.013). These findings were in line with the ERLOPET data with erlotinib monotherapy, suggesting a modest effect of adding bevacizumab on the tumor metabolism.

Figure 7: Box-plots demonstrating the difference in early FDG-SUV$_{max}$ reductions in patients with or without confirmed nonPD after six weeks of therapy



As in ERLOPET, early FLT-PET was not able to distinguish between PD and nonPD. Figure 8 a) and b) show the corresponding graphs to the FDG analyses above.

Figure 7: Early FLT-PET analyses



Notably, reviewing both graphs, FLT by using the early SUV_{max} values failed to predict nonPD. The AUC under the ROC curve was 0.609, with a p-value of 0.297. Reductions in the PD group had a mean of -11.1 (SD, 30.2%), whereas the nonPD group had a mean reduction of 15.7% (SD, 44.1%), with a p-value in t-test of 0.76.

DCE-MRI procedures showed a marked reduction of ktrans in the majority of patients, as shown in Figure 8 (paired t-test, p<0.001). Patients with nonPD had a mean reduction of 50.0% (SD, 45.1%), whereas patients with PD had a mean reduction of 37.3% (SD, 32.7%, p=0.369). The AUC was 0.68, with a p-value of 0.098 (Figure 9).

Figure 8: Early k_{trans} reductions per patient

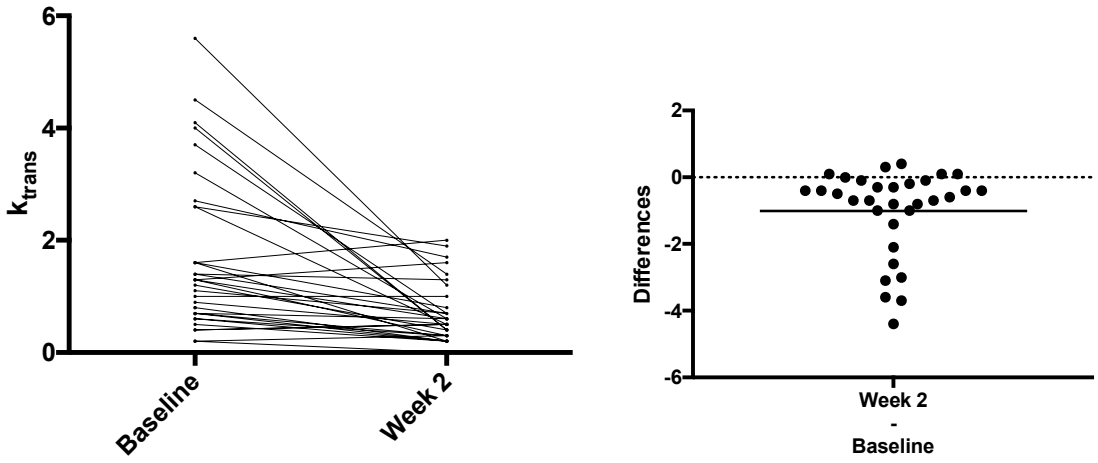
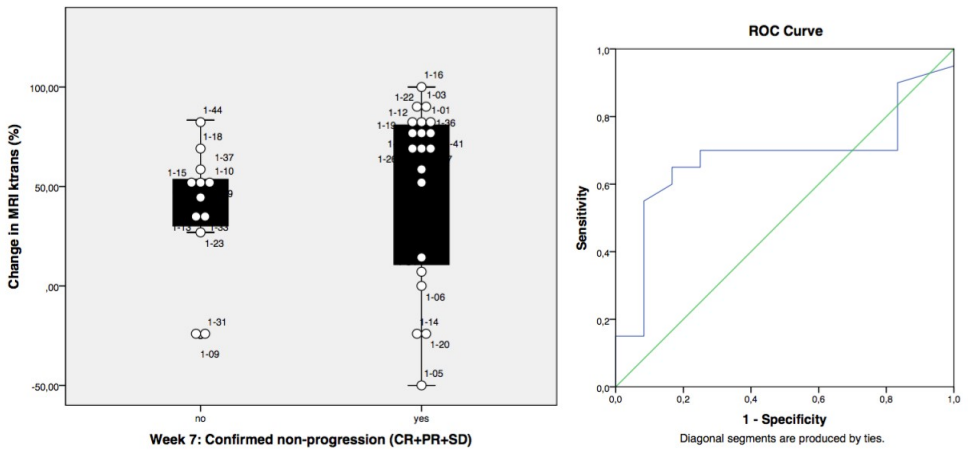
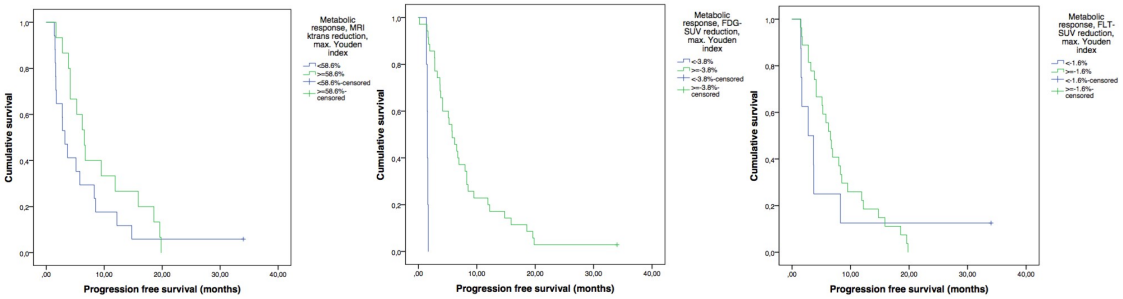


Figure 9: Early DCE-MRI analyzes, Box plots and ROC



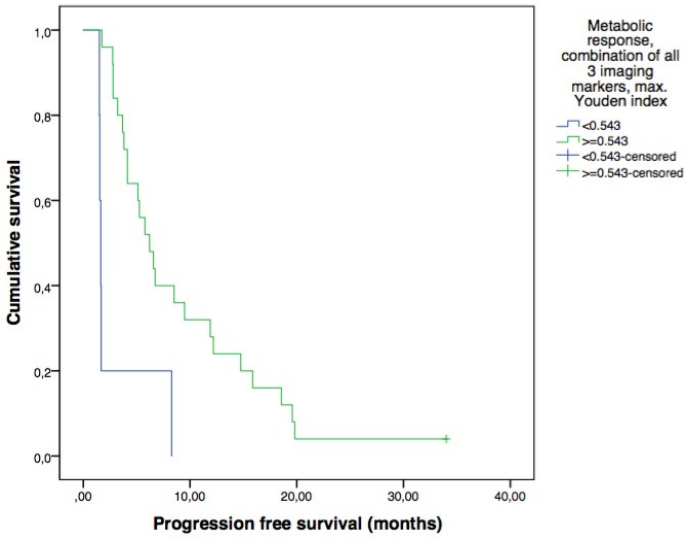
Regarding PFS prediction, non of the predefined cut-off values was able to find significant results. Taken the value of the maximum Youden-Index as a cut-off, early FDG-PET was able to significantly predict both PFS (p<0.001) and OS (p=0.002). Figure 10 shows the Kaplan Meier curves for all three early imaging methods with their respective Youden indices regarding PFS(DCE-MRI: p=0.194, FLT-PET: p=0.520):

Figure 10: Early imaging analyzes according to max. Youden-Index



The combination of all three imaging biomarkers led also to a significant predictive value (p=0.001):

Figure 11: Combining early imaging analyzes according to max. Youden-Index



The results of the tissue-based biomarker assessment are listed per patient in the attached "Listings_Pathoresults", and analyses on *KRAS*- and *EGFR*-mutated patients are provided at the end of the "998_Analysis_of_endpoints" files.

10.4.2 Statistical methods

All statistical methods were used as predefined in the trial's protocol. In accordance with the predefined statistics, we used Kaplan Meier statistics in order to define survival. Response according to RECIST and mutation results were considered nominal parameters. Reductions in the parameters assessed by imaging (percent) were considered metric variables. Based on this, ROC curves were calculated as described earlier.

Statistical calculations were done with the software SPSS Statistics 23 (IBM Corp., Armonk, NY, USA).

10.4.2.1 Control of covariates

Covariate analyses were performed as outlined in the protocol.

10.4.2.2 Missing values and drop-outs

Missing values were excluded from the analyses. All data gained within the trial were taken into account, i. e., no post-trial drop-out was performed.

10.4.2.3 Interim analyses and data monitoring

In the protocol, interim analyzes were mentioned, as within the then ongoing ERLOPET trial, (not preplanned) interim analyzes gave the impression that the primary objective might have been proven in less patients than the preplanned 40 patients. Nevertheless, at the end of the trial, this did not turn out to be true, and interim analyzes regarding the endpoints in the present trial were not performed due to a delay in imaging procession.

10.4.2.4 Multiple tests

Even though there were two predefined primary endpoints in this trial, the underlying statistics considered both endpoints for calculation of the participants. For both endpoints, 40 patients were considered the best fitting number.

10.4.2.5 „Per Protocol" analysis

Because of the small number of patients, we considered every patient who has received at least one mode of imaging after one week of therapy as suitable for the analyses. As there was no drop-out, we analyzed the whole ITT group as relevant. A separated per protocol analysis was not done.

10.4.2.6 Subgroup analyses

As preplanned in the protocol, subgroup analyses had been performed regarding gender and mutational status. The results are given in the respective "998_Analysis_of_endpoints" files

10.4.3 Summary of efficacy analyzes

The trial confirmed findings of the ERLOPET trial regarding the different PET tracers and their accuracy in predicting outcome, and added insights into tumor biology and the effects of bevacizumab on tumor vascularization - an effect which has been proposed so far, but was not shown in human patients so far. Because of the "universal" drops in these parameters regardless from mutational status or therapy outcome, we consider this to be a pharmacodynamic parameter of bevacizumab effect.

Some surprising outliers (e. g. patient 1-20 harbored a *PIK3CA* mutation [*EGFR* mutation was excluded using all possible analysis pipelines] and gained a complete response with 34 months PFS (still censored); patient 1-42 had PD in the first FDG scan after some view days despite the presence of an *EGFR* mutation. Both had nearly the same reduction in early FDG-PET and are mentioned as examples here) led to problems regarding the prediction of PFS according to predefined cut-off values. In ERLOPET, reduction of the SUVs of 30% could significantly discriminate between short and long PFS. In the present trial, these cut-off values were of limited value, unlike the maximum Youden-Index was. Noteworthy, in ERLOPET, the 30% cut-off value equaled the calculated Youden-Index in that trial (29.5% reduction). We will carefully discuss these findings - the drop of the Youden-Index by the addition of bevacizumab - in the future.

A complete analysis of mutational and immunohistochemical biomarkers still has to be performed. As proposed in the protocol, KRAS and EGFR mutational status has been analyzed concerning predictive values. Noteworthy, KRAS mutations did not a significant role in predicting PD or nonPD. Neither did EGFR mutational state, clearly due the the aforementioned outlier. For further details, see attachments.

Subgroup analysis provided strong differences between the gender groups (see attached files). For example, (see 998b_Analysis of Endpoints), in women, the combination of all biomarkers led to an AUC of 0.926, whereas in men, an AUC 0f 0.714 did not significantly predict nonPD. In contrast, in men, FLT-PET reductions greater than the maximum Youden-Index were significantly predictive for PFS ($p<0.001$), whereas in women, this parameter had a p-value in Log Rank of 0.914. Further analyses including this issue in combination with molecular findings are warranted.

11 Safety analyses

11.1 Exposure of trial medication

Please confer to the aforementioned attached files (Listing_Screenstudmed, MIMEB_Descriptive_statistics), Table 5 and Figure 5. All patients received at least base-line imaging procedures, erlotinib, one infusion of bevacizumab and one early FDG-PET.

11.2 Adverse events

11.2.1 Summarization of adverse events

An overview about the AEs within the trial is given in "997_Analysis_of_adverse_events". In total, we documented 508 AEs, whereof the majority (308, 60.1%) where grade I, according to CTC AE. 156 (30,7%) where graded as °II. 40 (7.9%) where grade III, three (1%) were life-threatening (°IV), and one (0.2%) of AEs was fatal (hemoptysis, considered related to bevacizumab).

Related to the treated patients, 38 (95%) suffered from at least °I AE(s). 36 patients (90%) had at least °II, while half of the trial population suffered severe (= °III) AEs. The remaining four patients (10%) suffered life-threatening (°IV) or fatal (°V, see above) AEs.

Of the forty patients, 28 patients (70.0%) had AEs considered related to bevacizumab. 36 patients (90%) had AEs related to erlotinib.

We counted 42 SAEs in 24 patients (60.0%) patients, whereof most were attributed to the underlying disease. Neither the AEs nor the SAEs were unexpected.

In general, and as opposed the chemotherapy first-line treatments, the treatment was considered well tolerated and manageable in an outpatient ward setting.

11.2.2 All adverse events

A complete list of all adverse events as named by the investigators is attached as "List_AEs".

11.2.3 Analysis of adverse events

Most of the adverse events were related to the underlying terminal disease. The typical erlotinib adverse reactions, rash and diarrhea, occurred in 32 (80.0%) and 26 (65.0%) patients, respectively. Hypertension as a common adverse reaction of bevacizumab was documented in 8 (25%) patients, hemoptysis was seen in two patients (5%).

In general, the frequencies of therapy-related AEs were in accordance with those seen in other trials. Relationship was assessed using the IBs of both products.

11.2.4 Listing of all AEs per patient

A complete list with all AEs per patient is provided by "Listing_Aesae", as attached.

11.3 Deaths, SAEs

11.3.1 Deaths

At the time of reporting, all patients except one (1-20) have died from their underlying disease. There was no unexpected death or death which was provoked or accelerated by trial specific procedures or by the trial medication. The dates of deaths are given in the "998_Analysis_of_endpoints" file.

11.4 Evaluation of laboratory values

11.4.1 Listing of laboratory values assessed per patient

A list of all laboratory parameters assessed in the trial per patient is provided in the attached "Listing_Chemhemat". Values which were out of range (for ranges, ring-trial certifications of the Institute of Clinical Chemistry can be made available) were considered by the investigators as "significant" or "not significant".

11.4.2 Evaluation of laboratory values over time

11.4.2.1 Per patient

As laboratory assessment was part of the adverse events analysis and neither primary nor secondary endpoint, please confer to "Listing_Aesae" in which all deviations from range are found and evaluated according to CTC AE. All laboratory data assessed per patient during the whole course of the trial is given in "Listing_Chemhemat".

11.4.2.2 Individual noticeable problems

We did not reveal noticeable problems regarding the laboratory values during trial termination. See "Listing_Aesae" and "Listing_Chemhemat" as well as "MIMEB_Descriptive_statistics" and "Listing_Screenstudmed", in which the treatment modifications are listed.

11.5 Vital signs and clinical parameters

Vital signs reported in the study were blood pressure, pulse rate, weight, height, and ECOG performance status, all of which were assessed at any study visit. The findings are shown in the attached "Listing_Vitalsigns" file.

As stated, hypertension was documented in 8 patients. Loss of weight was not analyzed as it was biased by the course of the underlying disease. "Listing_Screenstudmed" lists the preexisting illnesses per patient. In total, no evidence of so far not known findings were made during termination of this trial.

11.6 Summary of safety analyses

Safety analyses were secondary objectives taking into account the combination of both drugs with the extended imaging procedures. Both investigational drugs as well as the substances used for imaging have been well described regarding their adverse events, and all five substances used in this trial have proven in the past to be safe. Nevertheless, there was limited data available on the safety of the combination of the trial medications erlotinib and bevacizumab, both which were approved for the treatment of NSCLC at the initiation of the trial, but not in combination. In the present trial, we did not find any hints of a marked toxicity beyond the known adverse events of the single substances. Because of the small sample size, it is not known if the addition of bevacizumab leads to a higher rate of skin-related AEs, as rash (of all grades) occurring in 80% of patients seems to be slightly higher than in erlotinib mono trials.

Potential risks regarding the contrast-agents used for imaging procedures were minimized by the inclusion and exclusion criteria, i. e., diabetes (FDG) was as well an exclusion criterion as renal failure (Gd-DPTA), and bleeding disorders which might aggravate bevacizumab side-effects were also excluded. Laboratory abnormalities were predefined, e. g. patients with liver metastases had a greater tolerability of transaminase elevation.

Dose reductions were in line with standard of care, and concomitant medication related to erlotinib, bevacizumab or FDG (insulin) was administered according to local and community standards, or to clinical guidelines affecting this issue.

The trial population consisted of terminally ill patients with NSCLC. Therefore, it is understandable that most AEs and SAEs documented within the trial were due to or related to the underlying disease. The high percentage of deaths occurring from the disease was anticipated.

12 Summary and discussion

The trial was very difficult to conduct in comparison with other trials, because of the amount of experimental/non-established methods. Nevertheless, given the fact that all 40 patients correctly enrolled in this trial were available for statistical analyzes is worth to be mentioned.

The primary objectives of the trial did not focus on the efficacy and/or safety of the combination of erlotinib and bevacizumab, and the patient collective was as close as possible to an all-comer scenario. The safety analyzes showed a good tolerability in the patients, and no new safety concerns were raised. The finding that FDG-PET was (again) able to predict non-progression after six weeks underlines its position as the tracer of choice in cancer assessment. Proliferation as assessed by FLT-PET did not discriminate benefit from non-benefit at all; its use in further trials with NSCLC as entity is not imaginable at the moment. DCE-MRI was feasible in assessing data on tumor vascularization, and changes regarding a reduction of the contrast-agent efflux from the vessels good be shown in nearly all patients, providing evidence that recent cellular models on how anti-angiogenesis should work are right in proposing the vascular normalization being the main key player in that process. Nevertheless, here was a problem of the trial design: as all patients started paralleled with both trial drugs, one could only guess which drugs is responsible for the detected effect. A monotherapy run-in phase would have solved this problem.

The fact that none of the procedures could predict PFS was disappointing, as this finding in ERLOPET was considered of great clinical value. There were some outliers in the trial regarding PFS, that this might contribute to the small n.

When this trial started (and even after LPLV), genetic testing of NSCLC patients was highly anticipated, but far from reality. The main idea of this trial was to identify patients with benefit from the therapy non-invasively without or beyond mutational status known. This had completely changed since then, with growing knowledge about specific mutations and there optional treatment.

13 Attached files

- **997_Analysis_of_adverse_events_2016-04-05**
- **998_Analysis_of_endpoints_2016-04-05**
- **998b_Analysis_of_endpoints_2016-04-05**
- **998c_Analysis_of_endpoints_2016-04-05**
- **998d_Analysis_of_endpoints_2016-04-05**
- **998e_Analysis_of_endpoints_2016-04-05**
- **998f_Analysis_of_endpoints_2016-04-05**
- **Listing_Aesae_2016-03-17**
- **Listing_Chemhemat_2016-03-17**
- **Listing_Conmed_2016-03-17**
- **Listing_Continuos_2016-03-23**
- **Listing_Inclexcl_2016-03-17**
- **Listings_mrifdg_2016-03-23**
- **Listing_Pathoresults_2016-03-22**
- **Listing_Screenstudmed_2016-03-17**
- **Listing_Vitalsigns_2016-03-17**
- **MIMEB_Descriptive_statistics_2016-03-23**

14 Further Attachments

14.1 Trial information

14.1.1 Protocol, Module 1

14.1.2 CRF

14.1.3 Informed consent

14.1.4 List of investigators and responsible persons

14.1.5 Audit reports