PET-adapted treatment for newly diagnosed advanced Hodgkin lymphoma (AHL2011): a randomised, multicentre, non-inferiority, phase 3 study

Rene-Olivier Casasnovas, Reda Bouabdallah, Pauline Brice, Julien Lazarovici, Hervé Ghesquieres, Aspasia Stamatoullas, Jehan Dupuis, Anne-Claire Gac, Thomas Gastinne, Bertrand Joly, et al.

PET-adapted treatment of patients with advanced Hodgkin lymphoma (AHL2011): final results of a randomised, multi-centre, phase 3 study

Rene-Olivier Casasnovas MD[1], Reda Bouabdallah MD[2], Pauline Brice MD[3], Julien Lazarovici MD[4], Prof Hervé Ghesquieres MD[5], Aspasia Stamatoullas MD[6], Jehan Dupuis MD[7], Anne-Claire Gac MD[8], Thomas Gastinne MD[9], Bertrand Joly MD[10], Krimo Bouabdallah MD[11], Emmanuelle Nicolas-Virelizier MD[12], Prof Pierre Feugier MD[13], Prof Franck Morschhauser MD[14], Richard Delarue MD[15], Hassan Farhat MD[16], Philippe Quittet MD[17], Alina Berriolo-Riedinger MD[18], Adrian Tempescul MD[19], Véronique Edeline MD[20], Hervé Maisonneuve MD[21], Luc-Matthieu Fornecker MD[22], Prof Thierry Lamy MD[23], Prof Alain Delmer MD[24], Peggy Dartigues MD[25], Prof Laurent Martin MD[26], Prof Marc André MD[27], Prof Nicolas Mounier MD[28], Prof Alexandra Traverse-Glehen MD[29], Prof Michel Meignan MD[30]

1. Department of hematology, University Hospital F. Mitterrand and Inserm UMR 1231, Dijon, France
2. Department of hematology, Institut P. Calmette, Marseille, France
3. Department of hematology, APHP, Hopital Saint Louis, Université Paris Diderot, Paris, France
4. Department of hematology, Gustave Roussy, Université Paris-Saclay, Villejuif, France
5. Department of hematology, Hospices Civils de Lyon, Centre Hospitalier Lyon-Sud, et Université Claude Bernard Lyon-1, Pierre Bénite, France
6. Department of hematology, Centre H. Becquerel, Rouen, France
7. Department of hematology, Hopital H. Mondor, Creteil, France
8. Department of hematology, Institut d'hématologie de basse normandie, Caen, France
9. Department of hematology, University Hospital of Nantes, Nantes, France
10. Department of hematology, Hospital Sud Francilien, Corbeille-Essonnes, France
11. Department of hematology, University Hospital of Bordeaux, Bordeaux, France
12. Department of hematology, Centre L. Bérard, Lyon, France
13. Department of hematology, University Hospital of Nancy, Vandoeuvre les Nancy, France
14. Department of hematology, Univ. Lille, CHU Lille, EA 7365 - GRITA - Groupe de Recherche sur les formes Injectables et les Technologies Associées, Lille, France
15. Department of hematology, Hopital Necker, Paris, France
16. Department of hematology, Hopital Mignot, Versailles, France
17. Department of hematology, University Hospital of Montpellier, Montpellier, France
18. Department of nuclear medicine, Centre G.F. Leclerc, Dijon, France
19. Department of hematology, University Hospital of Brest, Brest, France
20. Department of nuclear medicine, Hopital R. Huguenin, Institut Curie, St-Cloud, France
21. Department of hematology, Hopital departemental de Vendée, La Roche sur Yon, France
22 Department of hematology, University hospital of Strasbourg, Strasbourg, France
23. Department of hematology, University hospital of Rennes, Rennes, France
24. Department of hematology, University Hospital of Reims, Reims, France
25. Department of pathology, Institut G. Roussy, Villejuif, France
26. Department of pathology, University Hospital F. Mitterrand, Dijon, France
27. Department of hematology, CHU UCL Namur, Université catholique de Louvain, Yvoir, Belgium
28. Department of hematology, University Hospital of Nice, Nice, France
29. Department of pathology, Hospices Civils de Lyon, Centre Hospitalier Lyon Sud, et Université Claude Bernard Lyon-1, Pierre Bénite, France
30. LYSA Imaging, University Hospital H. Mondor, Creteil, France

A part of this study has been presented at the ASH meeting 2015 (Orlando – USA), the ASCO meeting 2018 (Chicago - USA), and the EHA meeting 2018 (Stockholm – Sweden)

Abstract: 585 words (300 words)
Text: 5667 (4500 words) / Research in context: 562
Tables: 4, Figures: 3
30 references (30 references)

Abstract
Background,
Escalated BEACOPP improves progression free survival (PFS) in patients with advanced Hodgkin lymphoma (HL) compared to ABVD but at the cost of increased risk of hematological toxicity, secondary myelodysplasia/leukemia and infertility. We thus tested interim PET to guide treatment after upfront BEACOPPescalated aiming at reducing BEACOPPescalated exposure in early responding patients without loss of disease control compared to a not PET-monitored BEACOPPescalated treatment.
Methods,
AHL2011 is a randomized, multicenter study in 16 to 60 years old patients with ECOG performance status <3 and previously untreated advanced HL (Ann Arbor Stage III, IV or IIB with a mediastinum/thorax ratio ≥0.33 or extranodal localization) that compares standard treatment delivering 6xBEACOPPescalated (every 21 days in mg/m^2: Bleomycin (10) and vincristine (1.4) intravenously on day 8, Etoposide (200) intravenously on days 1 to 3, Doxorubicin (35) and Cyclophosphamide (1250) intravenously on day 1, Procarbazine (100) given orally on days 1 to 7, and prednisone (40) given orally from day 1 to 14) to a PET-driven strategy after 2xBEACOPPescalated (PET2), aimed at delivering 4xABVD (every 28 days in mg/m^2: Doxorubicin (25), Bleomycin (10), Vinblastine (6), Dacarbazine (375) intravenously on days 1 and 15) for PET2 negative and 4xBEACOPPescalated for PET2 positive patients. The randomization of patients was done centrally using the permuted block method and stratified according to Ann Arbor stage (IIB vs III-IV) and international prognosis score (IPS: 0-2 vs ≥3).

PET2 central review using Deauville criteria guided the allocation of treatment in the PET-driven arm. The primary objective was to demonstrate a non-inferiority of the PET-driven arm in terms of PFS, defined as the time from randomization to first progression, relapse and either death, whatever the cause, or last follow-up, with an absolute difference <10% in the 5-year PFS estimates in both intent to treat and per-protocol analysis. This study is registered with ClinicalTrial.gov, number NCT01358747.

Findings,
823 patients were registered in 90 centres, 410 in the PET-driven and 413 in the 6xBEACOPPescalated arms. Based on PET2 results, 346/84% patients were assigned to receive ABVD and 51/12% additional cycles of BEACOPPescalated in the PET-driven arm. With a median follow-up of 50.34 months (IQR:42.9-59.3), the 5-year PFS was similar in the 6xBEACOPPescalated and the PET-driven arms in both the intent-to-treat (86.2% (95%CI:81.6-89.8) vs 85.7% (95%CI:81.4-89.1); HR=1.084 (95%CI:0.737-1.596); p=0.65) and the per-protocol analysis (86.7% (95CI%:81.9-90.3) vs 85.4% (95CI%:80.7-89); HR=1.144 (95CI%:0.758-1.726); p=0.73). The most common grade 3-4 adverse events in the safety population including 412 patients in the 6xBEACOPPescalated and 407 patients in the PET-driven arms were leukopenia 381/92% vs 387/95%, neutropenia 359/87% vs 366/90%, anemia 286/69% vs 114/28%, thrombocytopenia 271/66% vs 163/40%, febrile neutropenia 145/35% vs 93/23%, infections 78/19% vs 43/11% and gastro-intestinal disorders 41/10% vs 41/10% respectively. 192/47% and 114/28% serious adverse events related to treatment were reported in the 6xBEACOPPescalated and PET-driven arms respectively and were mainly infections (84/20% vs 50/12%) and febrile neutropenia (21/5% vs 23/6%). 6/1.4% (2 cases of septic shocks, 2 cases of pneumopathy, 1 case of heart failure, 1 case of acute myeloblastic leukemia) and 2/0.5% patients (1 case of septic shock, and 1 case of acute myeloblastic leukemia) respectively, died from serious adverse events deemed related to study treatment.

Interpretation,

PET performed after 2xBEACOPPescalated can safely guide subsequent treatment and supports the use of response-adapted strategy delivering 4xABVD for patients with negative PET2 without impairing the disease control. This PET-driven strategy after upfront BEACOPPescalated provides a better disease control than reported after upfront ABVD and could become the safer treatment alternative than ABVD in routine practice.

Funding,

French government PHRC program

Research in context
Evidence before this study

At the time of planning this trial two randomized phase III studies in patients with advanced Hodgkin lymphoma (HL) comparing upfront ABVD and BEACOPP-based treatment have shown that BEACOPPescalated significantly reduces the risk of treatment failure but at the cost of a higher toxicity compared to ABVD. Since AHL2011 study was launched two studies and one metanalysis were published confirming the benefit of BEACOPPesc over ABVD. To reduce patient exposure to BEACOPPescalated without compromising disease control, PET performed after 2 cycles of chemotherapy, a time point which it has a strong prognostic value, could allow to guide chemotherapy dose intensity. We searched Medline up to august 22, 2018, for full papers reporting prospective trials evaluating PET-guided therapy in advanced Hodgkin lymphoma, with the search terms "Hodgkin", "lymphoma", "advanced or Stage III IV", "PET2 or interim PET". PET-guided strategies after upfront ABVD, which delivers BEACOPP in 16 to 19% PET2 positive patients and pursues ABVD in PET2 negative patients has been shown to moderately improve disease control in PET2 positive patients while inferior progression free survival was seen in PET2 negative patients compared to patients treated with upfront BEACOPPescalated. One phase III PET-guided study after upfront BEACOPPescalated (HD18) that tested the reduction to 4 instead of 6 or 8 cycles of BEACOPPescalated in PET2 negative patients has been reported. Due to the PET2 positivity criteria used only 48% of patients were eligible for deescalated treatment with a non-inferior outcome compared to those receiving 6 or 8 cycles of BEACOPPescalated.

Added value of this study

To our knowledge AHL2011 is the first multi-centre study that compares head-to-head in advanced stage Hodgkin's lymphoma a PET-driven strategy after 2xBEACOPPescalated delivering, 4xABVD for PET2 negative patients and 4xBEACOPPescalated for PET2 positive patients, to a not PET-monitored standard treatment delivering 6xBEACOPPescalated.

The proportion of patients in the PET-driven arm eligible for a de-escalated treatment on the basis of a negative PET2 defined by a Deauville score <4 is impressive and reached 84% compared to 48% in the HD18 study, without loss of disease control compared to the 6xBEACOPPescalated arm. Delivering 4xABVD in PET2 negative patients is associated with less grade 3-4 hematological toxicity than reported with 2xBEACOPPesc in the HD18 study

The PET4 assessment implemented in AHL2011 study to secure the treatment de-escalation strategy brings additional prognostic information to PET2, and PET4 positivity identifies a subset of patients with a particularly poor outcome. Indeed, the full interim PET assessment allows an innovative stratification of patients in 3 subsets with significantly different outcomes.

Implication of all the available evidence

Cumulative evidence from 6 trials show that PET-driven strategies after upfront BEACOPPescalated provide a an improved disease control than after upfront ABVD and could become a better treatment option than ABVD. The de-escalation of treatment in PET2 negative patients increases the tolerability of BEACOPP-based treatment

and delivering 4xABVD in these patients appears the safest treatment option in routine practice. The interim PET interpretation criteria using Deauville score<4 for negativity are required to obtain an accurate identification of patients eligible for deescalated treatment and using a SUVmax liver threshold allows a more precise definition of PET positivity and could be recommended to apply the present PET-driven strategy. The full interim PET staging including PET2 and PET4 allows an accurate monitoring of patient treatment and could thus be considered for use in the routine management of patients with advanced HL.

Introduction

ABVD (doxorubicin, bleomycin, vinblastine, dacarbazine) chemotherapy is widely used as standard treatment for Hodgkin lymphoma (HL). However, escalated BEACOPP (BEACOPPescalated) regimen (bleomycin, etoposide, doxorubicin, cyclophosphamide, vincristine, procarbazine, and prednisone) developed by the German Hodgkin study group[1] (GHSG), which delivers more drugs at a higher dose intensity appears to improve patient's outcome. BEACOPPescalated provides a 10-years failure free and overall survival (OS) of 82% and 86% respectively[2]. Four studies have shown that BEACOPP improves the disease control with a 15% 3-year progression free survival (PFS) benefit[3-6] in patients with advanced HL compared to ABVD. A meta-analysis[7] has shown an OS benefit in favor of BEACOPPescalated over ABVD, although formal proof in a randomized trial is not available.

A drawback however is that increased anti-lymphoma activity of BEACOPP compared to ABVD is associated with a marked and frequent albeit manageable immediate hematologic toxicity and a higher risk of secondary myelodysplasia and leukemia[2,8]. Also, gonadal toxicity which is a real concern in young patients and increases with the number of cycles delivered and the age of patients, is also higher when using the BEACOPPescalated regimen[9].

The toxicity issues incited us to identify: (1) -Early response patients to BEACOPP treatment in whom a dose intensity decrease strategy after upfront BEACOPP could be beneficial both in terms of treatment safety and cure rate and (2) -Patients in whom it would be beneficial to prolong BEACOPP treatment at higher dose intensity than that provided by the ABVD regimen.

Interest in using early fluorodoxyglucose positron emission tomography (PET) to better predict response to treatment and to drive therapy is emerging for HL patient management. PET performed after 2 courses of chemotherapy (PET2) has been shown to predict patients outcome in terms of progression free survival (PFS)[10,11] and to reach a negative predictive value of 98% in BEACOPPescalated treated patients[12]. Thus, PET2 may allow identification of a population of early responding patients who are suitable an ABVD-based conventional dose chemotherapy after 2 cycles of upfront BEACOPPescalated[12,13].

In this setting, we designed and performed the AHL 2011 study to evaluate an experimental PET-driven strategy after 2xBEACOPPescalated, aimed at delivering 4xABVD for PET2 negative patients and 4xBEACOPPescalated for PET2 positive patients, respectively, compared to a treatment regimen with 6xBEACOPPescalated in patients with advanced HL.

Methods

Study design and participants

This open-label, multicenter randomized phase 3 study was designed by the Lymphoma Study Association (LYSA) scientific committee and conducted in 90 centers from Belgium and France.

Eligible patients were 16 to 60 years old with an ECOG performance status <3, a minimum life expectancy of 3 months and previously untreated, histologically proven, classical HL according to WHO 2008 criteria and an Ann Arbor Stage III, IV or IIB with a mediastinum/thorax ratio ≥0.33 or extranodal localization. Patients were required to have negative HIV, HCV, and HTLV serology, and normal liver (bilirubin less than 2.5 normal level), renal (creatinine ≤150 µmol/L) and hematological functions (leukocyte count ≥2000 per µl an platelet count ≥100000 per µl) unless abnormalities were related to HL. Patients with severe cardio-pulmonary (left ejection ventricular fraction <50% or respiratory insufficiency prohibiting bleomycin use) or metabolic disease (uncontrolled diabetes mellitus) interfering with normal application of protocol treatment were not eligible for inclusion. All patients provided written informed consent before enrolment. The study was approved by the French and Belgium Health authorities, the Dijon Hospital ethics committee for French centers and by the institutional review boards of each participating site in Belgium and was done in accordance with the Declaration of Helsinki and the International Conference on Harmonisation Guidelines for Good Clinical Practice.

Randomization and masking

Patients were enrolled by center using LYSARC e-Rando system. Patients were randomly assigned to receive in 1:1 ratio a PET-driven strategy, or a standard treatment not monitored by PET. The randomization of patients was done centrally using the permuted block method and stratified according to Ann Arbor stage (IIB vs III-IV) and international prognosis score (IPS: 0-2 vs ≥3). The randomization list was generated by LYSARC. Patients and investigators were not masked to treatment allocation.

Patients received 2 cycles of upfront BEACOPPescalated. In the PET-driven arm, on the basis of the blinded central PET review results, patients with positive PET2 received 4 additional cycles of BEACOPPescalated and those with negative PET2 received 4 cycles of ABVD. In the standard arm, patients received 4 additional cycles of BEACOPPescalated whatever the PET2 result (figure 1 and appendix p1). PET was implemented in both arms to evaluate response after 4 cycles of chemotherapy and secure the de-escalation strategy. PET4 positive patients were considered as treatment failure and treated at the discretion of the investigator (figure 1).

Procedures

BEACOPPescalated was repeated every 21 days and included in mg/m^2: Bleomycin (10) and vincristine (1.4) given IV on day 8, Etoposide (200) given IV on days 1 to 3, Doxorubicin (35) and Cyclophosphamide (1250) given IV on day 1, Procarbazine (100) given orally on days 1 to 7, and prednisone (40) given orally from day 1 to 14. ABVD was repeated every 28 days and included in mg/m^2: Doxorubicin (25), Bleomycin (10), Vinblastine (6), Dacarbazine (375) given IV on days 1 and 15. Granulocyte colony-stimulating factor administration was mandatory at day 9 of each cycle of BEACOPPescalated in both randomization arm until neutrophil count reached >1000 per µl and optional after ABVD cycles. Chemotherapy dose reductions were permitted according to the rules detailed in the protocol (appendix pp 30-32). Prophylactic sulfamethoxazole 800mg/trimethoprime

160mg, 3 days a week and valacyclovir 1000mg/day were mandatory to prevent opportunistic infections for the duration of the 6 cycles of chemotherapy.

A baseline PET scan (PET0) was mandatory with at least one evaluable hypermetabolic lesion. Two PET examinations scheduled 3 weeks after the second (PET2) and 2 or 3 weeks after the 4th induction cycle (PET4) for patients receiving ABVD or BEACOPPescalated, respectively, were required for full assessment. Each patient was scanned on the same camera for baseline and subsequent PET scans. A whole-body acquisition from groin to head was started 60 ± 10 minutes after a 5MBq/Kg 18F-FDG injection.

PET0, PET2, and PET4 images were sent through a web platform[14] for a blinded independent central review, to 3 expert reviewers. PET2 and PET4 were binary interpreted as positive or negative as per Deauville criteria[15,16] and the final result was based on at least 2 concordant responses. To improve the interobserver reproducibility of interim PET interpretation[17-19] the increased activity relative to the liver background which defines positive residual uptake in the Deauville criteria should be at least 40% and 100% higher than the liver, instead of "moderate" or "marked" increased uptake assessed visually, for a Deauville score 4 and 5 respectively[20]. The central review PET result was sent back to the investigator, together with the per-protocol recommended treatment allocation for the patients randomized in the PET-driven arm.

Chest X ray was mandatory at baseline to estimate the mediastinum/thorax ratio. CT was mandatory at baseline, after 4 cycles of chemotherapy, and at the end of treatment and then every 6 months during the follow-up period. Bone marrow aspiration was mandatory at baseline and for confirmation of complete remission in case of baseline involvement. Laboratory monitoring of hematological parameters was mandatory before each cycle of chemotherapy and at least twice a week during the treatment period. Response to treatment was assessed with Cheson 2007 criteria[21].

Study treatment had to be discontinued in case of lymphoma progression, toxicity of study treatment or concomitant illness or protocol violation that precluded further study treatment, start of a new treatment for Hodgkin lymphoma, withdrawal of consent or refusal to continue treatment. PET2 or PET4 positivity were not considered as a PFS event unless a disease progression according to Cheson 2007[21] criteria was documented.

Outcomes

The primary endpoint was PFS by investigator assessment, defined as the time from randomization to first progression, relapse and either death, whatever the cause, or last follow-up. The secondary endpoints included toxicity, PET2 and PET4 responses and event free (EFS) disease free (DFS) and overall survival. Treatment-emergent adverse events were assessed after each cycle of chemotherapy and graded according to the National Cancer Institute Common Terminology Criteria for Adverse Events, version 3 and the treatment-related toxicities were reported according to the randomization arm. PET2 and PET4 responses were defined as a Deauville score 1 to 3. EFS was defined as the time from randomization to the first documented disease progression, relapse, initiation of a new anti-lymphoma therapy, death from any cause or last follow-up. DFS was defined as the time from complete response attainment to the date of first documented disease progression, relapse or death related to lymphoma, toxicity including secondary cancer, unknown cause, or last follow-up. OS was defined as the time from randomization to death from any cause, or last follow-up.

Other secondary endpoints including: - Comparison of the fertility parameters of patients under 45 year-old before and after treatment in both arms, - Assessment of the SUVmax reduction between baseline PET and PET2 or PET4 and analysis of its impact on response rate, EFS, DFS, PFS and OS in both arms, - Identification of

biological parameters related to the Hodgkin lymphoma cells and the tumor microenvironnement, influencing PET2 and PET4 responses, and - Assessment of the influence of the genetic polymorphisms of cytokine and cytokine receptor involved in the lymphoma process, and enzyme genes involved in the drug metabolism, on PET2 and PET4 responses and PFS, will be reported elsewhere.

Statistical analysis

The aim of the study was to demonstrate the non-inferiority of the PET-driven arm relative to the not PET-monitored arm in terms of PFS. The sample size calculation used an exponential model and was based on an estimated 85% 5-year PFS in the 6xBEACOPPescalated arm and a non-inferiority margin of 10% was considered clinically relevant, corresponding to a 5-year PFS>75% in the PET-driven arm with a hazard ratio (HR)=1.77. Indeed, since the PET-driven arm is expected to provide at least 76% 5y-PFS, this result would be better than the 70% 5-year PFS reported with ABVD[3,4]. With a one-sided significance level of 0.025 and 80% power, a total of 97 PFS events were required for the final analysis. No drop-outs accounted for the power calculation. The plan was to randomly assign a total of 810 patients. An interim analysis of the primary endpoint for futility following a Lan-DeMets sequential design was planned after 50% of the scheduled events needed for the final analysis. The interim analysis was conducted in 2015 (data cutoff date: July 1st, 2014) and showed no significant difference of PFS between the 2 arms, leading the data and safety monitoring committee to recommend pursuing the study.

The analysis of outcome including PFS and OS was done in an intent to treat basis (ITT set) including all randomized patients and stratified on randomization factors. Sensitivity analysis were conducted and included an unstratified analysis and a per protocol analysis (per-protocol set) excluding all patients with major protocol deviations. Per protocol analysis was considered a more conservative analysis to support the non-inferiority objective. Major protocol violations included: unconfirmed HL diagnosis, at least one inclusion or exclusion criteria not respected, first cycle of chemotherapy not received or not received at full dose, PET2 or PET4 not performed at the right time, central review of PET2 or PET4 not performed, treatment assignment according to PET result not followed.

Estimates of survival were calculated according to the Kaplan-Meier method with 95% CIs. The survival distributions were compared with the log-rank test and Cox proportional hazard regression models were used to estimate the hazard ratios (HR) and associated 95% CIs. The date of point was October 31, 2017.

Non-inferiority of PFS was established if the upper limit of the two-sided 95% CI for HR was above 1.77. In addition, non-inferiority was tested in a post hoc analysis using the Com-Nougue test[22].

In order to compare the relative influence on PFS and OS of the full PET-driven strategy to the baseline patients' characteristics found to impact outcome in univariate analysis, a Cox proportional hazard regression model was fitted including the interim PET profile and IPS as explanatory variables.

Response rates and PET results after 2 or 4 cycles in both arms were expressed with 95% exact Clopper Pearson Confidence Interval limits and compared using the Chi-squared test.

The analysis of toxicity was done on the safety population which includes all patients who were formally randomized and had received at least one dose of treatment.

We judged differences to be significant if p values were <0.025 for PFS and OS analysis according to treatment arm to respect the one-sided hypothesis, and <0.05 for all other analysis.

The relative dose intensity of the BEACOPPescalated regimen drugs was estimated as: (Administered dose/Expected dose) / (Observed administration duration/Expected administration duration)x100.

All outputs were produced using SAS version 9.3 (SAS Institute, Cary, NC, USA). This study is registered with ClinicalTrial.gov, number NCT01358747.

Role of the funding source

The trial sponsor was the University Hospital of Dijon, France, which was funded by the French government PHRC program. The funder and the sponsor had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The sponsor delegated to the LYSARC (LYSA Clinical Research) the clinical operations including randomization, monitoring procedures, organization of central PET and pathology review, reporting of serious adverse events, data gathering, entry and validation, statistical analysis, and production of the report. ROC, NM and MM had full access to all the data in the study and ROC had final responsibility for the decision to submit for publication.

Results

From May 19, 2011 to April 29, 2014, 823 patients were registered and randomized according to the protocol (ITT set) including 413 patients in the 6xBEACOPPescalated arm and 410 in the PET-driven arm, respectively. The characteristics of the patients (table 1) were well balanced between both arms: median age was 30 years (range:16–60; IQR:24-41), 63% of patients were male, 68% had B symptoms, 88% had stage III or IV disease and 11% a stage IIB with risk factors, 58% had an IPS of ≥3. Seven hundred and twenty-eight patients (88%) had a centralized pathology review of the diagnosis biopsy with 96% of these showing a confirmed diagnosis of HL. Most cases of misdiagnosis were grey zone lymphoma with features intermediate between diffuse large B-cell lymphoma and HL. Three hundred and forty-two (83%) patients in the 6xBEACOPPescalated arm and 359 (88%) in the PET-driven arm completed the planned treatment. In the 6xBEACOPPescalated versus PET-driven arm, 10 and 12 patients, respectively, discontinued treatment for progression, and 27 and 4 patients respectively, discontinued therapy because of treatment-related toxicity (figure 1).

Seven hundred ninety-nine patients (97%) had evaluable PET2 and a central review was performed in 795 cases (99%). PET2 negativity was reached in 87% of patients according to central review and was similar in both arms (table 2). Then, in an ITT analysis, 346 of the 410 patients randomized in the PET-driven arm (84%) were assigned to receive 4xABVD and 51 (12%) to four additional cycles of BEACOPPescalated. Fourteen patients (3.5%) did not receive the allocated treatment due to the clinician's decision: 9 patients were given BEACOPPescalated instead of ABVD while 5 patients received ABVD instead of BEACOPPescalated.

At the time of the analysis, with a median follow-up of 50.4 months (range:0-72; IQR:42.9-59.3), a PFS event occurred in 103 (12.5%) of 823 patients: 41 (10%) of 413 and 47 (11.5%) of 410 patients progressed or relapsed and 2 (0.5%) and 4 (1%) died from lymphoma in the 6xBEACOPPescalated and the PET-driven arm respectively. Eight (1%) of 823 patients died from toxicity of the study treatment (6 vs 2), 4 (0.5%) from toxicity of additional treatment (3 vs 1), 2 (0.25%) from concurrent illness (1 vs 1) and 5 (0.6%) for other or unknown reasons (2 vs 3) in the 6xBEACOPPescalated and the PET-driven arm respectively. In an ITT analysis the estimated 5-year PFS was similar in both arms (86.2% (95%CI:81.6-89.8) and 85.7% (95%CI:81.4-89.1)), with a stratified HR=1.084 (95%CI:0.737-1.596, p=0.65) and an unstratified HR=1.066 (95%CI:0.725-1.569),

p=0.63), from which the 95%CI superior bound were lower than the preplanned 1.77 HR (figure 2). So far, the median PFS was not reached in both arms. To support the primary endpoint result a per protocol analysis was performed retaining 739 patients (90%) with no major protocol deviation of the 823 in the ITT set. 41 (9.9%) of 413 and 43 (10.4%) of 410 patients were excluded from the pet protocol analysis in the 6xBEACOPPescalated and PET-driven arms respectively for: -unconfirmed HL diagnosis in 26 cases (20 vs 5), -at least one inclusion or exclusion criteria was not respected in 12 cases (9 vs 3), -the first cycle of chemotherapy was not received or not received at full dose in 12 cases (5 vs 7), PET2 or PET4 was not performed at the right time in 11 cases (7 vs 4), central review of PET2 or PET4 was not performed in 7 cases (3 vs 4), treatment assignment according to PET2 or PET4 result was not followed in 16 cases (2 vs 16). The per-protocol analysis provided consistent results with 5-year PFS (86.7% (95CI%:81.9-90.3) vs 85.4% (95CI%:80.7-89); HR=1.144 (95CI%:0.758-1.726); p=0.74). The post hoc Com-Nougue non-inferiority test gave a similar conclusion by rejecting the null hypothesis (p=0.0047). OS was also similar in both randomization arms with, in the ITT population 95.2% 5-year OS (95%CI:91.1-97.4) in the 6xBEACOPPescalated and 96.4% (95%CI:93.3-98.1) in the PET-driven arms, respectively (HR=0.936 (95%CI:0.427-2.051); p=0.91) (figure 2) and in the per-protocol population 5-year OS (95.6% (95CI%:91.2-97.8) vs 95.9% (95CI%:92.5-97.8); HR=1.248 (95CI%:0.53-2.88); p=0.69). The event-free survival and disease-free survival estimates in the ITT population were also similar in both randomization arms (appendix p 1-2): 5-year EFS was 76.8% (95CI%: 71.7-81) vs 78.6% (95CI%: 73.9-82.6) (HR=0.925 (95CI%: 0.686-1.248); p =0.31) and 5-year DFS 89.9% (95CI%: 85.1-93.2) vs 90% (95CI%: 86-92.9) (HR=1.099 (95CI%: 0.667-1.711); p = 0.66) in the 6xBEACOPPescalated and PET-driven arms respectively.

Seven hundred and sixty-six (93%) of 823 patients had evaluable PET4, and a central review was performed in 759 cases (99%). PET4 negativity according to central review was achieved in 716 (94%) of 759 patients and was similar in both arms. In 654 (86%) of 716 PET4 negative patients, PET negativity was already obtained after 2xBEACOPPescalated (PET2-/PET4-). In 62 (8.6%) of 716 patients PET positivity after 2xBEACOPPescalated converted to PET4 negativity after 2 additional BEACOPPescalated (PET2+/PET4-). Among the 43 patients (5.7%) who had a positive PET4, and were removed from the study, 13 (1.6%) had a previous negative PET2, including 6 patients in the PET-driven arm.

Interim PET positivity was associated to a higher risk of relapse or progression, regardless of the randomization arm: patients with positive PET2 had a lower PFS compared to PET2 negative patients (5-year PFS: 70.7% (95%CI:60.7-78.6) vs 88.9% (95%CI:85.7-91.4); HR=3.59 (95%CI:2.32-5.56); p<0.0001). The 43 (5.6%) of 766 patients with positive PET4 had a particularly poor outcome with a 46.5% 5-year PFS (95%CI:31.2-60.4) compared to PET4 negative patients (89.6% (95%CI:86.5-92); HR=10.9 (95%CI:6.75-17.61); p<0.0001). The 5-year PFS estimates according to PET4 results were similar in both randomization arms (51.9% (95%CI:31.9-58.5) vs 37.5% (95%CI:25.4-59.8); p=0.4751) (table 2). Twenty (46.5%) of the 43 patients with positive PET4 had progressive disease at the time of PET examination and received salvage chemotherapy. The remaining patients continued BEACOPPescalated except one who received radiotherapy on a residual mediastinal mass and one who proceeded to salvage therapy.

Analysis of the full PET-driven strategy allowed identification of 3 prognostic subsets of patients (figure 3): PET4 positive patients had a significantly lower 5-year PFS compared to PET2+/PET4- (75.4%, 95%CI:62.5–84.4) and PET2-/PET4- patients (90.9%, 95%CI:87.7–93.3, p<0.0001).

Baseline IPS≥3 (5-year PFS: 82.8% (95%CI:78.5-86.3) vs 90.3% (95%CI:85.8-93.4); HR=1.91 (95%CI:1.25-2.94); p=0.0025) male gender and bulk (≥ 10 cm) were also associated with lower 5-year PFS (84.6% (95%CI:81.1-87.5) vs 88.2% (95%CI:81.9-92.5); HR=0.577 (95%CI:0.37-0.89); p=0.013; and 83.6% (95%CI: 78.4-87.7) vs 87.8 (83.4-91.1); HR=1.6 (95%CI: 1.05-2.24); p=0.027, respectively) while ECOG, Age, Ann Arbor stage, B symptoms and mediastinal bulk (M/T≥0.33) had no impact on PFS. Multivariable analysis including the full PET-driven strategy, and IPS as covariates, shows that PET assessment retained independent prognosis value from IPS for PFS (table 3).

PET2 positivity was associated to a lower 5-year OS (92.4% vs 96.7%; HR=3.727 (95%CI:1.5–9.24); p=0.0029) while PET4 positivity was not associated to a significant increased risk of death (5-year OS: 93.6% vs 96.8%; HR=2.569 (95%CI:0.58–11.28); p=0.19). Indeed, with the current follow-up the full PET-driven strategy combining PET2 and PET4 has marginal influence on OS (figure 3) as 5-year OS estimate of PET2-/PET4- patients was 97.1% (95%CI:94.2-98.5) compared to 93.5% (95%CI:83.6–98) and 93.6% (95%CI:75.6–98.4;p=0.039), in PET2+/PET4- and PET4+ patients respectively (figure 3). No standard clinical or biological factors including IPS was found to significantly impact the risk of death.

A total of 819 (99.5%) of the 823 patients who entered the study received at least one dose of the planned treatment (safety population) including 412 (99.8%) of 413 and 407 (99.3%) of 410 patients in the 6xBEACOPPescalated and the PET-driven arms, respectively. The median relative dose intensity of each drug composing the BEACOPPescalated regimen was similar in both arms and reached 95% or higher of the planned dose for each cycle (appendix, p3) and the planned full dose of each drug was maintained in at least 85% of patients (appendix, p4). Overall, 467 (57%) of the 819 patients of the safety set required at least one dose reduction including 264 (64%) of 412 patients in the 6xBEACOPPescalated arm and 203 (50%) of 407 patients of the PET driven arm respectively. 32 patients discontinued treatment due to study treatment-related toxicity and the chemotherapy discontinuation which was mainly related to hematological toxicity and infections, occurred more frequently in the 6xBEACOPPescalated arm (28/7%) compared to the PET-driven arm (4/<1%).

The most common treatment-emergent adverse events of any cause or grade in the 819 patients in the safety population were hematological toxicity, gastro-intestinal disorders, general disorders as fatigue or fever and infections (table 4). The most common grade 3-4 adverse events in the 412 and 407 patients in the safety population in the 6xBEACOPPescalated and the PET-driven groups were leukopenia 381/92% vs 387/95%, neutropenia 359/87% vs 366/90%, anemia 286/69% vs 114/28%, thrombocytopenia 271/66% vs 163/40%, febrile neutropenia 145/35% vs 93/23%, infections 78/19% vs 43/11% and gastro-intestinal disorders 41/10% vs 41/10% respectively. 192/47% and 114/28% serious adverse events related to treatment were reported in 6xBEACOPPescalated and PET-driven arms respectively and were mainly infections (84/20% vs 50/12%) and febrile neutropenia (21/5% vs 23/6%). 6/1.4% (2 cases of septic shocks, 2 cases of pneumopathy leading to acute distress syndrome, 1 case of heart failure, 1 case of acute myeloblastic leukemia) and 2/0.5% patients (1 case of septic shock after the first cycle of BEACOPPescalated, and 1 case of acute myeloblastic leukemia) respectively, died from serious adverse events deemed related to study treatment.

To date 15 cases of secondary primary malignancies were reported including 10 (2.4%) and 5 (1.2%) cases in patients of the 6xBEACOPPescalated and PET-driven groups respectively. Secondary primary malignancies included: - in patients in the 6xBEACOPPescalated arm: acute myeloid leukemia (4 cases), non-

Hodgkin lymphoma (1 case), breast cancer (2 cases), cutaneous basal cell carcinoma (2 cases), lung cancer (1 case). - in patients in the PET-driven arm: acute myeloid leukemia (1 case), non-Hodgkin lymphomas (2 cases), renal cancer (1 case), thyroid cancer (1 case).

To date, 73 pregnancies were reported including 28 (6.8%; 95%CI:4.5-9.8) in the 6xBEACOPPescalated arm and 45 (11%; 95%CI:8-14.7) in the PET-driven arm respectively (p=0.036). Among these, assisted reproduction was required in 6 (21%) and 3 cases (7%), respectively.

Discussion

To our knowledge AHL2011 is the first large phase III study aimed at randomly comparing head-to-head PET-driven versus conventional non-monitored treatment in advanced HL showing that after 2xBEACOPPescalated the PET-driven strategy provides similar patients outcomes. We find that reducing treatment intensity in patients achieving an early metabolic response is safe and does not impair disease control as evidenced by non-inferiority of PFS between both randomization arms. Indeed, the primary endpoint of the study was met with 5-year PFS of 86.2% vs 85.7% in the 6xBEACOPPescalated and the PET-driven arms, respectively. As 84% of patients achieved a negative PET2, the study demonstrates that a high dose intensity chemotherapy beyond the 2 first cycles of BEACOPPescalated is not required in most patients. Indeed, 97% of PET2 negative patients received ABVD in this study. Furthermore, this PET-driven strategy allowed to reduce patient exposure to BEACOPPescalated and consequently the rate of early treatment-related toxicity (0.4 decrease in serious adverse events), and the risk of treatment discontinuation related to toxicity. Long-term toxicity may also be lowered as evidenced by: -lower incidence of secondary primary malignancy in patients receiving ABVD compared to those continuing BEACOPPescalated[2,8] even if the current follow-up is too short to conclude and -significantly more pregnancies reported in the PET-driven arm compared to the 6xBEACOPPescalated arm while further fertility parameters remain to be analyzed.

The Deauville criteria for interim PET interpretation applied here were relevant to select patients eligible for treatment by ABVD versus those who should continue to receive BEACOPPescalated. Deauville scores 4 and 5 are suitable for identifying patients with different outcome and use of a SUVmax liver threshold allows more precise definition of these scores. This limited modification of the Deauville score does not alter the Deauville scale per se. Rather it more precisely defines scores 4 or 5 as recommended[18,19], with an agreement for score 4 of 82% between the standard and per protocol Deauville criteria. This limits the risk of false positive results and may help nuclear medicine physicians and clinicians to make decision since it is easy to get SUV values for the liver and the residual tumor mass. We recommend this modification to define PET2 positivity for PET-driven BEACOPPescalated dose de-escalation in routine practice. In particular, using Deauville scores 4 and 5 for defining PET positivity minimized the risk of false positive results observed in the HD18[23] study in which Deauville score 3 was considered positive. Indeed, all other previously published PET-guided studies[24-27] used Deauville score >3 as a cut-off giving a PET2 positivity rate <20% (appendix p2) and, as in the present study, found an inferior outcome for PET2 positive patients. The inappropriate Deauville score cutoff used in the HD18 study has three main consequences: -PET2 does not show prognostic significance (appendix p2), -only 48% of ITT patients are eligible to reduce BEACOPPescalated treatment to 4 cycles of instead of 6 or 8 cycles[23,] -and the apparent good outcome of PET2 positive patients is misleading as this group includes a mix of true and false PET2 positive patients. Thus, the PET-guided strategy proposed by the GHSG is not adapted for patients

reaching a Deauville score 3 after 2xBEACOPPescalated. In a post hoc analysis the GHSG has shown that patients with Deauville score 3 and 1-2 have similar outcomes[28] and have further stated that patients with Deauville score 3 who represent about 25% of the whole HD18 cohort, could benefit from a de-escalated strategy. Unfortunately, we do not know if the deescalation to 4xBEACOPPescalated would have been possible in these patients without impaired outcome.

In our study, 28 of 100 (28%) PET2 positive patients did not achieve a complete metabolic response after 2 additional BEACOPPescalated cycles and 13 PET2 negative patients converted to positive PET4. PET4 brings additional prognosis value to PET2 results and identifies a 6% subset of patients with a very poor outcome. Conversely, double negative interim PET observed in 79% of patients (ITT set) is associated to a particularly favorable outcome. Therefore, the full PET-driven strategy including PET2 and PET4 has a strong independent prognostic value and improved power for risk stratification of patients with advanced HL independently of the IPS. This PET2/PET4 strategy shows increased power for identification of high versus lower risk for disease progression and death in advanced HL. It could be further developed for identification of patients with unfavorable PET profile for whom new treatment options could be proposed, versus those who have the maximal probability to be definitively cured with a safer treatment such as 2xBEACOPPescalated plus 4xABVD. Our work also suggests that PET4 is probably more suitable for patient management than end of treatment PET. Firstly, the confirmation of PET2 negativity after 4 cycles of chemotherapy is associated with excellent outcome. Secondly, the 62 patients with positive PET2 who subsequently converted in negative PET4 maintain acceptable probability of favorable outcome without treatment modification. Thirdly, PET4 permits early identification of patients with progressive disease who need salvage therapy.

The PFS of patients receiving 6xBEACOPPescalated in the 6xBEACOPPescalated arm is close to the 91% and 89.4% 5-year PFS reported in the GHSG studies[23,29]. Interestingly, this disease control rate was obtained in the present study while radiotherapy was not permitted, unlike the HD15 and HD18 studies in which 11% and 13% of the randomized patients received radiotherapy, respectively. Thus, the patient outcome observed in the 6xBEACOPPescalated arm of the present study looks robust enough to rule out achievement of a primary endpoint because of use of a weak comparator arm. Further supporting this, the relative dose intensity of chemotherapy in the 6xBEACOPPescalated arm was satisfactory and the proportion of patients with stage IV or IPS>3 was higher in AHL2011 compared to most of the other PET-driven studies published so far except the SWOG study (appendix p2).

Due to more accurate interim PET criteria, the de-escalated PET driven strategy developed in the present study was applicable in a higher proportion of patients than in the HD18 study (84% vs 48%). In addition, the toxicity observed in PET2 negative patients who received 2xBEACOPPescalated plus 4xABVD seems lower than patients who received 4xBEACOPPescalated in the HD18 study with less grade≥3 anemia (24% vs 39%), thrombocytopenia (36% vs 57%). Conversely, the risk of leukopenia, febrile neutropenia or sepsis was similar in PET2 negative patients in both studies. Lastly, no excess of pulmonary toxicity related to bleomycin with 2xBEACOPPescalated plus 4xABVD was reported (table 4).

Our results also compare favorably with the Echelon-1 study[30], even if the populations of patients enrolled in both studies were not completely similar: specifically, in AHL2011 patients were younger (30y vs 36y), presented more frequently with stage IIB with bulk or extranodal localization (12% vs 0%), slightly less

frequent stage IV disease (60% vs 64%) but more frequent B symptoms (68% vs 59%) and IPS>3 (31% vs 26%). Thus, despite the statistically significant improvement of modified PFS (mPFS) with AVD-Brentuximab Vedotin (AVD-BV) compared to ABVD the 82.1% 2y-mPFS achieved with AVD-BV was disappointing. In addition, the toxicity of AVD-BV included more serious adverse events (43%), grade>3 infections (18%) and treatment discontinuation related to toxicity (4.2%) than in the PET-driven arm of the present study (28%, 10% and 1.2% respectively).

PET-driven strategies were also developed after upfront ABVD[24-27]. The PET2 negativity rate after ABVD was a little lower than after BEACOPPescalated ranging from 80% to 84% while it reached 87.4% in the present study (appendix p2). In addition, the 90.4% 5-year PFS reached in PET2 negative patients after upfront BEACOPPescalated compares favorably to patients achieving a negative PET2 after ABVD (3-year PFS ranging from 79% to 87%)[24-27], resulting in more patients with a better outcome when using upfront BEACOPPescalated. PET2 positivity rate after ABVD in 3 studies that enrolled patients with comparable features to those enrolled in our AHL2011 study was associated with a consistent 64% or lower 3-year PFS despite a switch to BEACOPPescalated, so inferior to the 70.7% 5-year PFS reported here after upfront BEACOPPescalated. The RATHL study[27] provided better results in PET2 positive patients (67.5% 3-year PFS) but patients had a more favorable baseline profile with 42% of stage II, and only 37% of IPS≥3.

Altogether, PET-driven strategies after ABVD showed inferior results compared to upfront BEACOPPescalated with a lower chance to control the disease in the more frequent PET2 positive patients despite an intensified treatment (appendix p2). In aggregate this suggests that the dose intensity of upfront treatment matters for improved outcome in patients with more unfavorable features.

The limitations of the study included a PFS non-inferiority design with a predefined wide margin that could reach 10% between the two randomization arms. At the time the study was launched this margin seemed relevant as in the worst case, non-inferiority of the PET-driven arm compared to 6xBEACOPPescalated would be declared with a 5-year PFS of 76%, ie higher than the 70% 5-year PFS reported with standard ABVD and with a balance effectiveness/toxicity probably better than 6xBEACOPPescalated. However, the study shows that the observed 5-year PFS difference is much lower (-0.5%; 95%CI: -6.07-5) than the predefined margin and consequently no meaningful difference was detected between the randomization arms. Another limitation of the study is the choice of ITT rather than per-protocol as primary analysis of the main endpoint, which is a more conservative approach for demonstrating PFS estimates equivalence between both arms in a non-inferiority design. Primary ITT analysis was preferred because it was difficult to anticipate the rate of patients excluded in the per-protocol analysis since treatment assignment depended on a centrally reviewed interim PET and the compliance of investigators to apply this strategy was unknown. The PET-driven strategy was well applied in most cases and only 14 (3.5%) of the 397 patients who had PET2 central review in the PET-driven arm did not followed the per-protocol treatment assignment. However, the per-protocol analysis was prespecified as a sensitive analysis and was shown to support the ITT analysis results allowing to draw reliable conclusions of the results.

In summary, PET performed after 2xBEACOPPescalated can safely guide subsequent treatment and supports the use of a response-adapted strategy to deliver 4xABVD for patients with negative PET2 without impairing the disease control. This PET-driven strategy increases the tolerability of BEACOPP-based treatment

in most patients with advanced stage HL and provides similar outcome compared to 6xBEACOPPescalated. PET4 brings additional prognostic information to PET2. Full interim PET staging using the modified Deauville score describe here allows accurate monitoring of patient treatment and thus could be considered as a strategy for the routine management of patients with advanced HL.

Contributors

ROC, RB, PB, JL, HG, AS, JD, ACG, TG, BJ, KB, ENV, PF, FM, RD, HF, PQ, AT, HM, LMF, TL, AD, and MA performed the clinical management of patients at participating study centres. PD, LM, and ATG did the pathology central review. ABR, VE, and MM did the PET central review. NM supervised the statistical plan and the statistical analysis. ROC is the principal investigator, led the design of the study, and drafted the report. All authors contributed to data interpretation, reviewed the draft, and approved the final version of this report.

Declaration of interests

ROC reports grants, personal fees and non-financial support from Roche, grants, personal fees and non-financial support from Gilead, grants, personal fees and non-financial support from Takeda, personal fees and non-financial support from BMS, personal fees and non-financial support from MSD, personal fees from Abbvie, personal fees and non-financial support from Celgene, outside the submitted work.

PB reports personal fees from Takeda, personal fees from BMS, grants from Takeda Millenium, personal fees from MSD, non-financial support from Roche, outside the submitted work

AS reports personal fees from Takeda, outside the submitted work

ENV reports personal fees from Keocyt, personal fees from Sanofi, outside the submitted work.

FM reports personal fees from Roche/genentech, personal fees from Gilead, personal fees from janssen, personal fees from celgene, outside the submitted work

RD reports personal fees from Roche, personal fees from Janssen, personal fees from Karyopharm, personal fees from Celgene, personal fees from Takeda, personal fees from Gilead, personal fees from Sanofi, personal fees from BMS, outside the submitted work

MM reports personal fees for lectures for Roche China, outside the submitted work.

RB, JL, HG, JD, ACG, TG, BJ, KB, PF, HF, PQ, ABR, AT, VE, HM, LMF, TL, AD, PD, LM, MA, NM, ATG, have nothing to disclose.

Acknowledgments

The authors thank the LYSARC team which was greatly committed in the management and the operational support of this study and all the investigators

The authors thank Pr Mary Callanan for English assistance

Data sharing

Data sharing will not be possible for the AHL2011 data base. According to European General Regulation for Data Protection 2016/679 applicable from May 25, 2018, patients have to be informed about data transfer, especially when transfer is done out of European Union. Since AHL2011 patients were not informed about it, we are not allowed to transfer data.

References

1 Diehl V. Advanced Hodgkin's disease: ABVD is better, yet is not good enough! *J Clin Oncol* 2003; **21**: 583–5.

2 Engert A, Diehl V, Franklin J, *et al.* Escalated-dose BEACOPP in the treatment of patients with advanced-stage Hodgkin's lymphoma: 10 years of follow-up of the GHSG HD9 study. *J Clin Oncol* 2009; **27**: 4548–54.

3 Federico M, Luminari S, Iannitto E, *et al.* ABVD compared with BEACOPP compared with CEC for the initial treatment of patients with advanced Hodgkin's lymphoma: results from the HD2000 Gruppo Italiano per lo Studio dei Linfomi Trial. *J Clin Oncol* 2009; **27**: 805–11.

4 Viviani S, Zinzani PL, Rambaldi A, *et al.* ABVD versus BEACOPP for Hodgkin's lymphoma when high-dose salvage is planned. *N Engl J Med* 2011; **365**: 203–12.

5 Mounier N, Brice P, Bologna S, *et al.* ABVD (8 cycles) versus BEACOPP (4 escalated cycles ≥ 4 baseline): final results in stage III-IV low-risk Hodgkin lymphoma (IPS 0-2) of the LYSA H34 randomized trial. *Ann Oncol* 2014; **25**: 1622–8.

6 Carde P, Karrasch M, Fortpied C, *et al.* Eight Cycles of ABVD Versus Four Cycles of BEACOPPescalated Plus Four Cycles of BEACOPPbaseline in Stage III to IV, International Prognostic Score ≥ 3, High-Risk Hodgkin Lymphoma: First Results of the Phase III EORTC 20012 Intergroup Trial. *J Clin Oncol* 2016; **34**: 2028–36.

7 Skoetz N, Trelle S, Rancea M, *et al.* Effect of initial treatment strategy on survival of patients with advanced-stage Hodgkin's lymphoma: a systematic review and network meta-analysis. *Lancet Oncol* 2013; **14**: 943–52.

8 Diehl V, Franklin J, Pfreundschuh M, *et al.* Standard and increased-dose BEACOPP chemotherapy compared with COPP-ABVD for advanced Hodgkin's disease. *N Engl J Med* 2003; **348**: 2386–95.

9 Behringer K, Mueller H, Goergen H, *et al.* Gonadal function and fertility in survivors after Hodgkin lymphoma treatment within the German Hodgkin Study Group HD13 to HD15 trials. *J Clin Oncol* 2013; **31**: 231–9.

10 Gallamini A, Hutchings M, Rigacci L, *et al.* Early interim 2-[18F]fluoro-2-deoxy-D-glucose positron emission tomography is prognostically superior to international prognostic score in advanced-stage Hodgkin's lymphoma: a report from a joint Italian-Danish study. *J Clin Oncol* 2007; **25**: 3746–52.

11 Hutchings M, Loft A, Hansen M, *et al.* FDG-PET after two cycles of chemotherapy predicts treatment failure and progression-free survival in Hodgkin lymphoma. *Blood* 2006; **107**: 52–9.

12 Dann EJ, Bar-Shalom R, Tamir A, *et al.* Risk-adapted BEACOPP regimen can reduce the cumulative dose of chemotherapy for standard and high-risk Hodgkin lymphoma with no impairment of outcome. *Blood* 2007; **109**: 905–9.

13 Avigdor A, Bulvik S, Levi I, *et al.* Two cycles of escalated BEACOPP followed by four cycles of ABVD utilizing early-interim PET/CT scan is an effective regimen for advanced high-risk Hodgkin's lymphoma. *Ann Oncol* 2010; **21**: 126–32.

14 Meignan M, Itti E, Bardet S, *et al.* Development and application of a real-time on-line blinded independent central review of interim PET scans to determine treatment allocation in lymphoma trials. *J Clin Oncol* 2009; **27**: 2739–41.

15 Meignan M, Gallamini A, Meignan M, Gallamini A, Haioun C. Report on the First International Workshop on Interim-PET-Scan in Lymphoma. *Leuk Lymphoma* 2009; **50**: 1257–60.

16 Barrington SF, Mikhaeel NG, Kostakoglu L, *et al.* Role of imaging in the staging and response assessment of lymphoma: consensus of the International Conference on Malignant Lymphomas Imaging Working Group. *J Clin Oncol* 2014; **32**: 3048–58.

17 Biggi A, Gallamini A, Chauvie S, *et al.* International validation study for interim PET in ABVD-treated, advanced-stage hodgkin lymphoma: interpretation criteria and concordance rate among reviewers. *J Nucl Med* 2013; **54**: 683–90.

18 Biggi A, Bergesio F, Chauvie S, *et al.* Concomitant semi-quantitative and visual analysis improves the predictive value on treatment outcome of interim 18F-fluorodeoxyglucose / Positron Emission Tomography in advanced Hodgkin lymphoma. *Q J Nucl Med Mol Imaging* 2017; published online July 27. DOI:10.23736/S1824-4785.17.02993-4.

19 Barrington SF, Kluge R. FDG PET for therapy monitoring in Hodgkin and non-Hodgkin lymphomas. *Eur J Nucl Med Mol Imaging* 2017; **44**: 97–110.

20 Itti E, Juweid ME, Haioun C, *et al.* Improvement of early 18F-FDG PET interpretation in diffuse large B-cell lymphoma: importance of the reference background. *J Nucl Med* 2010; **51**: 1857–62.

21 Cheson BD, Pfistner B, Juweid ME, *et al.* Revised response criteria for malignant lymphoma. *J Clin Oncol* 2007; **25**: 579–86.

22 Com-Nougue C, Rodary C, Patte C. How to establish equivalence when data are censored: a randomized trial of treatments for B non-Hodgkin lymphoma. *Stat Med* 1993; **12**: 1353–64.

23 Borchmann P, Goergen H, Kobe C, *et al.* PET-guided treatment in patients with advanced-stage Hodgkin's lymphoma (HD18): final results of an open-label, international, randomised phase 3 trial by the German Hodgkin Study Group. *Lancet* 2018; **390**: 2790–802.

24 Press OW, Li H, Schöder H, *et al.* US Intergroup Trial of Response-Adapted Therapy for Stage III to IV Hodgkin Lymphoma Using Early Interim Fluorodeoxyglucose-Positron Emission Tomography Imaging: Southwest Oncology Group S0816. *J Clin Oncol* 2016; **34**: 2020–7.

25 Gallamini A, Patti C, Viviani S, *et al.* Early chemotherapy intensification with BEACOPP in advanced-stage Hodgkin lymphoma patients with a interim-PET positive after two ABVD courses. *Br J Haematol* 2011; **152**: 551–60.

26 Gallamini A, Tarella C, Viviani S, *et al.* Early Chemotherapy Intensification With Escalated BEACOPP in Patients With Advanced-Stage Hodgkin Lymphoma With a Positive Interim Positron Emission Tomography/Computed Tomography Scan After Two ABVD Cycles: Long-Term Results of the GITIL/FIL HD 0607 Trial. *J Clin Oncol* 2018; **36**: 454–62.

27 Johnson P, Federico M, Kirkwood A, *et al.* Adapted Treatment Guided by Interim PET-CT Scan in Advanced Hodgkin's Lymphoma. *N Engl J Med* 2016; **374**: 2419–29.

28 Borchmann P, Goergen H, Kobe C, *et al.* Early Interim PET in Patients with Advanced-Stage Hodgkin's Lymphoma Treated within the Phase 3 GHSG HD18 Study. *Blood* 2017; **130**: 737–737.

29 Engert A, Haverkamp H, Kobe C, *et al.* Reduced-intensity chemotherapy and PET-guided radiotherapy in patients with advanced stage Hodgkin's lymphoma (HD15 trial): a randomised, open-label, phase 3 non-inferiority trial. *Lancet* 2012; **379**: 1791–9.

30 Connors JM, Jurczak W, Straus DJ, *et al.* Brentuximab Vedotin with Chemotherapy for Stage III or IV Hodgkin's Lymphoma. *N Engl J Med* 2018; **378**: 331–44.

Table 1: Patients characteristics

		6xBEACOPPescalated		6xBEACOPPescalated or 2xBEACOPPescalated plus 4xABVD	
		N = 413		N = 410	
Median age (range)		31 (16 – 60)		29 (16-60)	
Male (n - %)		263	64%	253	62%
ECOG (n - %)	0	203	49%	193	47%
	1	181	44%	184	45%
	2	27	7%	31	8%
	missing	2		2	
B symptoms (n - %)		282	68%	278	68%
Ann Arbor stage (n - %)	I	0		2	<1%
	II	44	11%	52	13%
	III	114	28%	115	28%
	IV	255	62%	241	59%
Stage IIA (n - %)		2	<1%	7	2%
Stage IIB (n - %)		42	10%	45	11%
	M/T≥ 0.33	41	98%	45	100%
	Extra nodal localization	6	14%	4	9%
Bulky mass (n - %)	≥10 cm	143	38%	134	37%
	<10 cm	233	62%	229	63%
	missing	37		47	
Bone marrow involved (n - %)		33	8%	32	8%
IPS group (n - %)	0-2	160	39%	183	45%
	≥ 3	250	61%	225	55%
	missing	3		2	
Pathology review					
	Nodular sclerosis HL	273	74%	264	74%
	Mixed cellularity HL	20	5%	22	6%
	Lymphocyte-depleted HL	2	<1%	3	<1%
	Lymphocytes rich HL	2	<1%	1	<1%
	Interfollicular HL	1	<1%	0	
	Unclassified HL	51	14%	61	17%
	Gray zone lymphoma	20	5%	3	<1%
	Anaplastic large cell lymphoma-ALK-	0		2	<1%
	EBV-associated B-cell lymphoproliferative disorder	1	<1%	0	
	Insufficient material	1	<1%	1	<1%
	Missing	42	10%	53	13%

Table 2: Metabolic response after 2 and 4 cycles of chemotherapy according to PET central review

	6xBEACOPPescalated			6xBEACOPPescalated or 2xBEACOPPescalated plus 4xABVD		
	n = 413		5y-PFS	n = 410		5y-PFS
	n	% (95%CI)	% (95%CI)	n	% (95%CI)	% (95%CI)
PET2						
Reviewed	398	96%		397	97%	
Negative	349	88% (79-97)	88.4% (83.3-92)	346	87% (78-96)	89.4% (84.9-92.6)
Positive	49	12% (9-16)	73.5% (58.7-83.6)	51	13% (10-16)	68.2% (53.4-79.2)
PET4						
Reviewed	383	93%		376	92%	
Negative	356	93% (84-100)	90.1% (85.3-93.3)	360	96% (86-100)	89.2% (84.8-92.3)
Positive	27	7% (5-10)	51.9% (31.9-58.5)	16	4% (2-7)	37.5% (25.4-59.8)

Table 3: Risk factors found to influence progression free survival

Risk factors		n (%)	5y-PFS % (95%CI)	Univariate analysis HR	Univariate analysis p	Multivariate analysis HR	Multivariate analysis p
PET2/PET4	PET2-/PET4-	654 (79%)	92.5 (90.1-94.3)				
	PET2+/PET4-	62 (7.5%)	75.4 (62.5-84.4)	**3.588**	**<0.0001**	**2.567**	**0.0046**
	PET4+	43 (5.2%)	46.5 (31.2-60.4)	**13.14**	**<0.0001**	**10.367**	**<0.0001**
Gender	female	516 (63%)	88.2 (81.9-92.5)				
	male	307 (37%)	84.6 (81.1-87.5)	**1.73**	**0.013**	**1.407**	**0.19**
ECOG	0	396 (48%)	87.8 (84-90.8)				
	1	365 (44%)	84.3 (78.7-88.5)	**1.137**	**0.53**	**0.847**	**0.44**
	2	58 (7%)	86.1 (74.1-92.8)	**1.294**	**0.50**	**1.274**	**0.34**
B symptoms	No	263 (32%)	88.6 (83.5-92.2)				
	Yes	560 (68%)	84.7 (80.8-87.9)	**1.407**	**0.12**	**1.056**	**0.037**
Ann arbor stage	IIB	87 (11%)	86.3 (76.3-92.3)				
	III-IV	725 (89%)	85.8 (82.5-88.5)	**0.997**	**0.43**	**1.254**	**0.32**
Bulk	<10cm	462 (63%)	87.8 (83.4-91.1)				
	≥10cm	277 (37%)	83.6 (78.4-87.7)	**1.601**	**0.027**	**1.246**	**0.79**
IPS	0-2	343 (42%)	91.9 (88.4-94.4)				
	≥3	475 (58%)	83.7 (79.9-86.9)	**1.915**	**0.0025**	**1.6**	**0.044**

Table 4: Treatment emergent adverse events in the safety population according to the randomization arm

	6xBEACOPPescalated				6xBEACOPPescalated or 2xBEACOPPescalated plus 4xABVD			
	n = 412				n = 407			
	Grade 1-2	Grade 3	Grade 4	Grade 5	Grade 1-2	Grade 3	Grade 4	Grade 5
Blood and lymphatic system disorders								
Anemia	402 (98%)	249 (60%)	37 (9%)	0	394 (97%)	107 (26%)	7 (2%)	0
Leukopenia	189 (46%)	138 (33%)	243 (59%)	0	273 (67%)	102 (25%)	285 (70%)	0
Neutropenia	157 (38%)	65 (16%)	294 (71%)	0	221 (54%)	60 (15%)	306 (75%)	0
Febrile neutropenia	0	129 (31%)	16 (4%)	0	0	85 (21%)	8 (2%)	0
Thrombocytopenia	342 (83%)	148 (36%)	123 (30%)	0	306 (75%)	99 (24%)	64 (16%)	0
Gastro-intestinal disorders								
Mucositis	101 (25%)	13 (3%)	3 (<1%)	0	91 (22%)	18 (4%)	1 (<1%)	0
Vomiting	161 (39%)	9 (2%)	1 (<1%)	0	141 (35%)	10 (2%)	0	0
Diarrhea	93 (23%)	6 (1%)	1 (<1%)	0	88 (22%)	7 (2%)	0	0
Other	280 (68%)	16 (3%)	1 (<1%)	0	291 (72%)	11 (3%)	1 (<1%)	0
General disorders								
Fatigue	262 (64%)	16 (4%)	1 (<1%)	0	228 (56%)	11 (3%)	0	0
Fever	132 (32%)	5 (1%)	3 (1%)	0	125 (31%)	1 (<1%)	1 (<1%)	0
Other	87 (21%)	5 (1%)	3 (1%)	0	96 (24%)	5 (1%)	0	0
Infections and infestations								
Sepsis	3 (<1%)	0	27 (7%)	2 (<1%)	2 (<1%)	0	14 (3%)	0
Lung infection	17 (4%)	12 (3%)	0	0	16 (4%)	4 (1%)	0	0
Other	118 (29%)	45 (11%)	4 (1%)	0	120 (30%)	23 (6%)	5 (1%)	1 (<1%)
Investigation								
AST and/or ALT increased	136 (33%)	12 (3%)	3 (<1%)	0	132 (32%)	9 (2%)	2(<1%)	0
Creatinin increased	14 (3%)	1 (<1%)	1 (<1%)	0	25 (6%)	0	0	0
Other	92 (22%)	16 (4%)	1 (<1%)	0	80 (20%)	13 (3%)	0	0
Nervous system disorders								
Peripheral neuropathy	85 (21%)	8 (2%)	0	0	87 (21%)	2 (<1%)	0	0
Other	66 (16%)	6 (2%)	0	0	66 (16%)	5 (1%)	0	0
Respiratory, thoracic and mediastinal disorder								
Pneumonitis	4 (1%)	3 (<1%)	0	0	5 (1%)	3 (<1%)	1 (<1%)	0
Other	121 (29%)	11 (3%)	1 (<1%)	2 (<1%)	108 (26%)	11 (3%)	2 (<1%)	0
Vascular disorders								
Thromboembolic event	20 (5%)	7 (2%)	1 (<1%)	0	29 (7%)	7 (2%)	1 (<1%)	0
Hypotension	18 (4%)	4 (1%)	2 (<1%)	0	12 (3%)	0	2 (<1%)	0
Other	23 (6%)	3 (<1%)	2 (<1%)	0	24 (6%)	2 (<1%)	1 (<1%)	0
Skin and subcutaneous disorders	122 (30%)	4 (1%)	0	0	125 (31%)	8 (2%)	0	0
Metabolism and nutrition disorder	58 (14%)	5 (1%)	0	0	40 (10%)	5 (1%)	0	0
Cardiac disorders								
Dysrhytmia	17 (4%)	1 (<1%)	0	0	14 (3%)	0	0	0
Other	24 (6%)	1 (<1%)	1 (<1%)	1 (<1%)	12 (3%)	4 (1%)	0	0
Renal and urinary disorders								
hematuria	6 (2%)	0	0	0	1 (<1%)	0	0	0
Other	20 (5%)	7 (2%)	0	0	15 (4%)	0	1 (<1%)	0
Immune system disorder	6 (2%)	1 (<1%)	0	0	9 (2.2%)	2 (<1%)	0	0
hepatobiliary disorders	5 (1%)	2 (<1%)	1 (<1%)	0	4 (1%)	0	1 (<1%)	0
Secondary malignancy possibly related to Hodgkin lymphoma treatment	0	2 (<1%)	7 (2%)	1 (<1%)	0	0	4 (1%)	1 (<1%)

Data are n (%). The table shows adverse events grade 1-2 that occurred in at least 10% of patients and all the grade 3-5 adverse events.

Figure 1: Disposition of patients

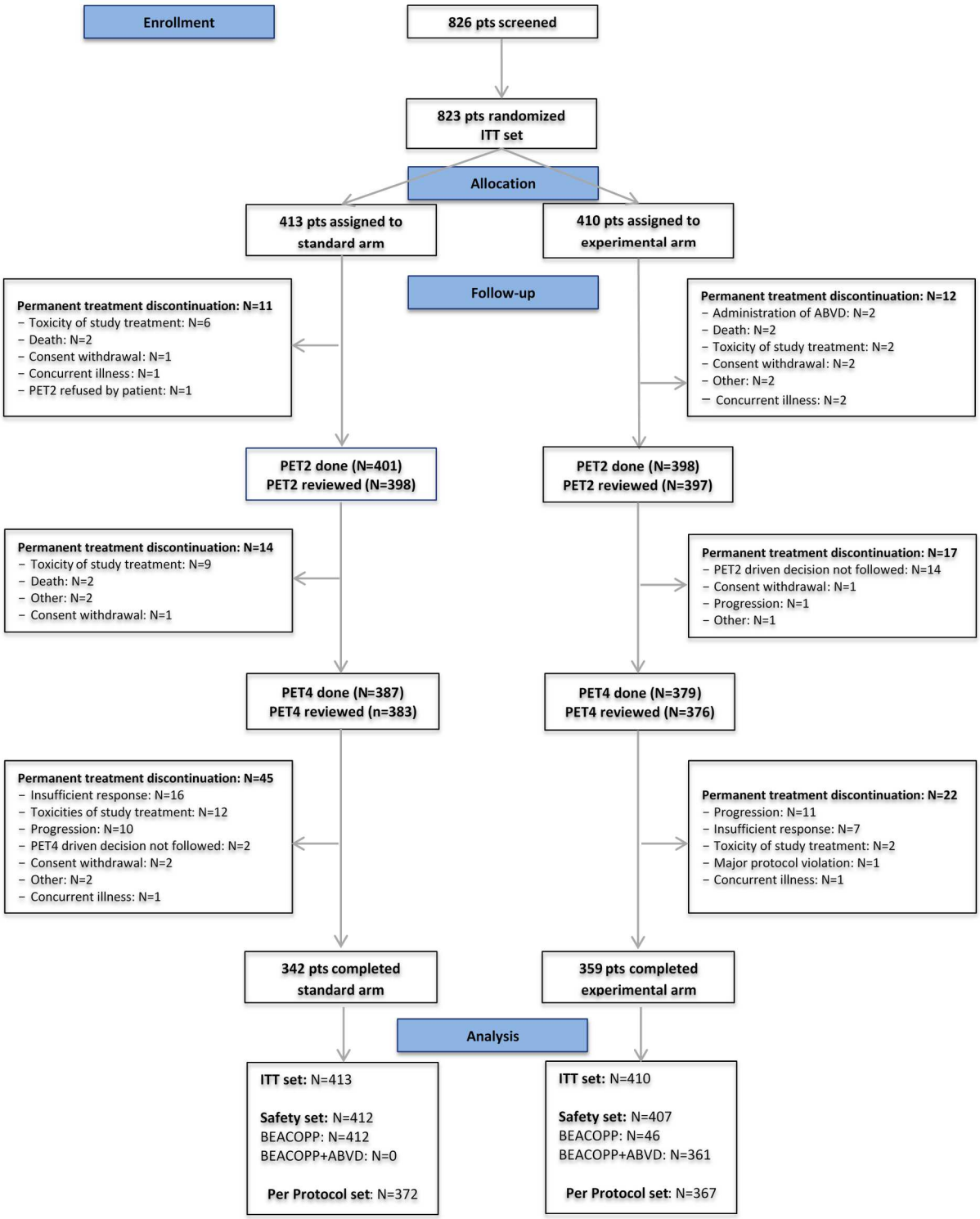
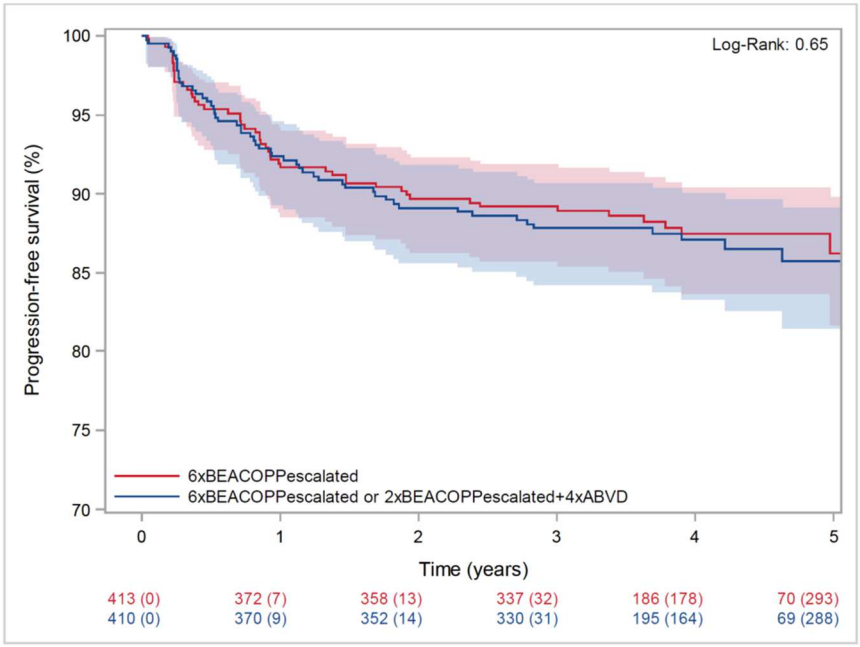


Figure 2: Progression free survival (A) and overall survival (B) according to the randomization arm in the ITT population

A



B

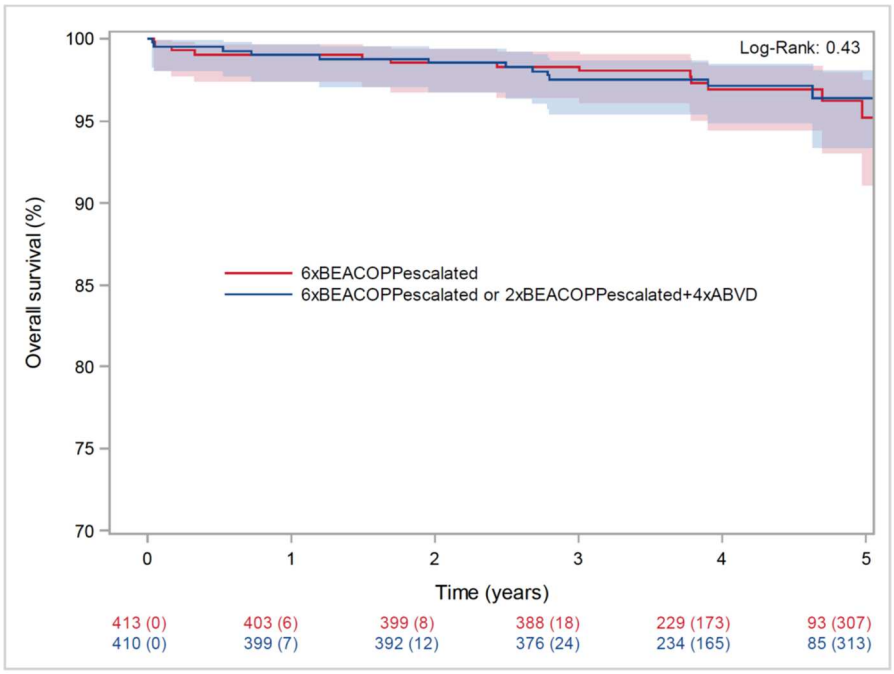
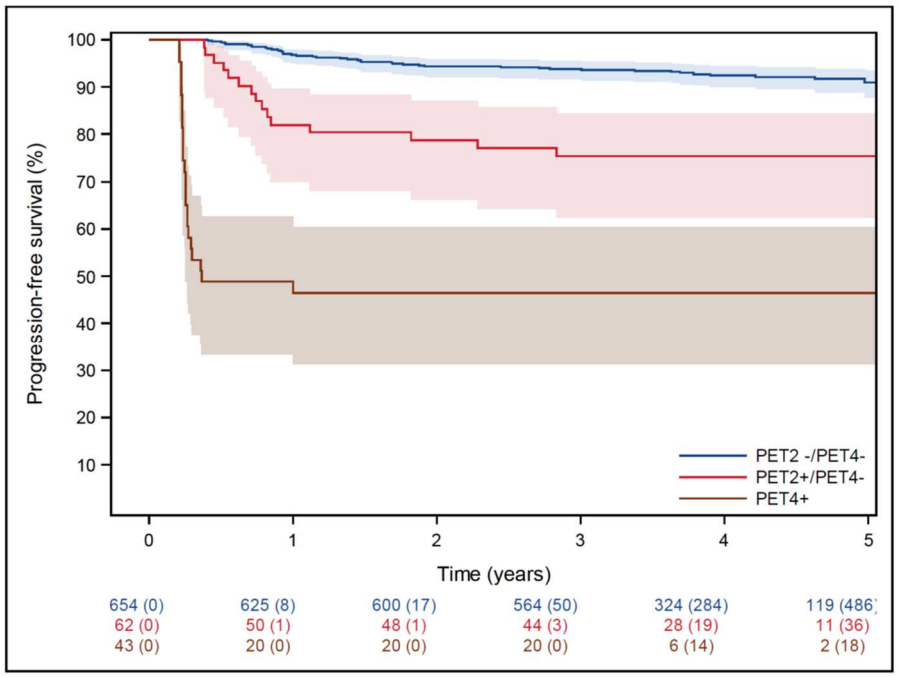


Figure 3: Progression free survival (panel A) and overall survival (panel B) according to the PET2 and PET4 results

A



B