The predictive role of interim positron emission tomography for Hodgkin lymphoma treatment outcome is confirmed using the interpretation criteria of the Deauville five-point scale

Andrea Gallamini,[1] Sally F. Barrington,[2] Alberto Biggi,[3] Stephane Chauvie,[4] Lale Kostakoglu,[5] Michele Gregianin,[6] Michel Meignan,[7] George N. Mikhaeel,[8] Annika Loft,[9] Jan M. Zaucha,[10] John F. Seymour,[11] Michael S. Hofman,[12] Luigi Rigacci,[13] Alessandro Pulsoni,[14] Morton Coleman,[15] Eldad J. Dann,[16] Livio Trentin,[17] Olivier Casasnovas,[18] Chiara Rusconi,[19] Pauline Brice,[20] Silvia Bolis,[21] Simonetta Viviani,[22] Flavia Salvi,[23] Stefano Luminari,[24] and Martin Hutchings[25]

[1]Research and Medical Innovation Department, Centre Antoine Lacassagne, Nice, France; [2]Division of Imaging, King's College London, PET Centre, Guy's & St. Thomas' Hospital, London, UK; [3]Nuclear Medicine Department, PET Center, Azienda Ospedaliera S. Croce e Carle, Cuneo, Italy; [4]Medical Physics Unit, Azienda Ospedaliera S. Croce e Carle, Cuneo, Italy; [5]Department of Radiology, Division of Nuclear Medicine, Mount Sinai Medical Center, New York, NY, USA; [6]Radiotherapy and Nuclear Medicine Unit, Istituto Oncologico Veneto IOV-IRCCS, Padua, Italy; [7]Nuclear Medicine Department, Centre Universitaire Hospitalier Henri Mondor, Creteril, Paris, France; [8]Clinical Oncology Department. Guy's & St. Thomas' Hospital, London, UK; [9]PET & Cyclotron Unit, Department of Clinical Physiology, Nuclear Medicine & PET, Rigshospitalet, Copenhagen University Hospital, Denmark; [10]Department of Oncology, Gdynia Oncology Centre & Department of Propedeutic Oncology, University of Gdansk, Poland; [11]Haematology Department, Peter MacCallum Cancer Centre, Melbourne, and University of Melbourne, Parkville Victoria, Australia; [12]Center for Cancer Imaging Peter Mac Callum Cancer Center, Melbourne, Australia; [13]Hematology Department, University of Florence, Careggi Hospital, Italy; [14]Cellular Biotechnology and Hematology Department, Sapienza University, Rome, Italy; [15]Hematology-Oncology Division, Center for Lymphoma & Myeloma, Weill Cornell Medical Center, New York, NY, USA; [16]Department of Hematology & Bone Marrow Transplantation; Rambam Medical Center, Haifa, Israel; [17]Hematology Department, University of Padua, Italy; [18]Hematology Department, Hopital Le Bocage, Dijon, France; [19]Hematology Department - Niguarda Ca' Granda Hospital, Milan, Italy; [20]Hematology Department Centre Hospitalier Universitaire St. Louis, Paris, France; [21]Hematology Department, S. Gerardo University Hospital, Monza, Italy; [22]Department of Medical Oncology, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy; [23]Department of Hematology, SS Antonio e Biagio Hospital, Alessandria, Italy; [24]Onco-Hematology Department, Modena University, Italy; and [25]Department of Hematology, Rigshospitalet, Copenhagen University Hospital, Denmark

 doi:10.3324/haematol.2013.103218
Manuscript received on December 22, 2013. Manuscript accepted on March 17, 2014.
Correspondence: gallamini.a@ospedale.cuneo.it

ON LINE SUPPLEMENTAL FILE

PET scan centralization and review

After anonymization, PET-0 and PET-2 scans were -uploaded from the participating PET centers to a dedicated website (https://magic5.to.infn.it/ivs) hosted by the National Institute of Nuclear Physics (INFN) in Turin, Italy. Details have been published previously [15]. Images were then transferred from the Duo workstations (Keosys, Nantes, France) located in Cuneo to a central server managed by Keosys [16]. The server was accessed by six international reviewers with recognized expertise in the field (AB, SB, MG, MH, LK, MM) who reported the scans using the Deauville 5PS [11,17]. In brief, PET-2 scans were scored by comparing the sites of uptake that were deemed to be involved by lymphoma on the baseline scan to the uptake in the normal mediastinal blood pool and the liver as follows:

Score 1, No uptake

Score 2, Uptake ≤ mediastinum

Score 3, Uptake >mediastinum and ≤ liver

Score 4, Uptake moderately increased above liver at any site

Score 5, markedly increased uptake above liver and/or new sites of disease

For the purpose of the analysis PET-2 scans with scores 1–3 were considered negative; scores 4–5 were considered positive. Reviewers scored the scans independently and blinded to the clinical outcome. It was decided prior to the review process that a scan would be defined as positive or negative where at least 4 reviewers agreed that a particular scan was positive or negative, respectively. True "discordant" cases were defined as cases where the reviewers were equally split in their opinions with 3 negative and 3 positive reports. For true discordant cases, a joint interpretation session was held with all the reviewers to reach final agreement. Additional clinical data were made available at that stage on request to clarify possible confounding factors in interpretation such as active clinical infection and the use of granulocyte colony stimulating factors.

Statistical analysis.

Progression-free survival (PFS) was defined as the time from diagnosis to either disease progression or relapse, or to death as a result of any cause, whichever occurred first. Overall Survival (OS) was defined as

previously reported [8]. Survival curves were calculated using the Kaplan Meier method [18]. Comparison between survival curves was carried out using Mantel-Haenszel, Log-Rank, Wilcoxon and Tarone-Ware tests. The association between clinical prognostic factors and the probability of treatment failure was assessed by log-rank and univariate regression analyses [19]. To investigate the contribution of individual prognostic factors to PFS, a multivariate analysis based on the Cox proportional hazards regression model was performed [20]. The level for significance was $p < 0.05$. All data analyses were performed using SPSS for Windows [21]. The concordance between pairs of reviewers with respect to binary results for PET interpretation, with a PET scan scored as 1,2 and 3 defined as negative and scored 4 and 5 defined as positive was measured using Cohen's Kappa for the 15 combinations of the 6 reviewers. Kappa values between 0.81 and 1.00 indicate a very good agreement, between 0.61 and 0.80 a good agreement [22].