Hematological Oncology
Hematol Oncol 2017; **35**: 69–78
Published online 7 August 2015 in Wiley Online Library
(wileyonlinelibrary.com) **DOI**: 10.1002/hon.2249

Original Research Article

Evaluation of the prognostic role of tumour-associated macrophages in newly diagnosed classical Hodgkin lymphoma and correlation with early FDG-PET assessment

Emanuele Cencini[1,2]*, Alberto Fabbri[1], Luigi Rigacci[3], Stefano Lazzi[4], Guido Gini[5], Maria Christina Cox[6], Salvatrice Mancuso[7], Elisabetta Abruzzese[8], Sofia Kovalchuk[3], Gaia Goteri[9], Arianna Di Napoli[10], Roberto Bono[7], Stefano Fratoni[11], Simonetta Di Lollo[12], Alberto Bosi[3], Lorenzo Leoncini[4] and Monica Bocchia[1,2]

[1]*Unit of Hematology, Azienda Ospedaliera Universitaria Senese, Italy*
[2]*University of Siena, Siena, Italy*
[3]*Hematology Department, University of Florence, Florence, Italy*
[4]*Department of Human Pathology and Oncology, University of Siena, Siena, Italy*
[5]*Hematology Department, Azienda Ospedaliera Universitaria 'Ospedali Riuniti', Ancona, Italy*
[6]*Hematology Department, Ospedale S.Andrea, Roma, Italy*
[7]*Hematology Division, Azienda Universitaria Policlinico, Palermo, Italy*
[8]*Hematology Department, Ospedale S.Eugenio, Roma, Italy*
[9]*Pathology Department, Università Politecnica delle Marche, Ancona, Italy*
[10]*Pathology Unit, Department of Clinical and Molecular Medicine, Sant'Andrea Hospital, La Sapienza University, Rome, Italy*
[11]*Pathology Department, Ospedale S.Eugenio, Rome, Italy*
[12]*Pathology Department, University of Florence, Florence, Italy*

*Correspondence to:
Dr. Emanuele Cencini Unit of Hematology University Hospital Viale Bracci, 16, 53100 Siena, Italy.
E-mail: cencioema@libero.it

Received 22 March 2015
Revised 7 July 2015
Accepted 9 July 2015

Abstract

In Hodgkin Lymphoma (HL), about 20% of patients still have relapsed/refractory disease and late toxic effects rate continue to rise with time. 'Early FDG-PET' and tissue macrophage infiltration (TAM) emerged as powerful prognostic predictors. The primary endpoint was to investigate the prognostic role of both early FDG-PET and TAM; the secondary endpoint was to test if early FDG-PET positivity could correlate with high TAM score.A cohort of 200 HL patients was analysed. Induction treatment plan consisted of two to six courses of ABVD and, if indicated, involved field radiation therapy. All patients repeated CT scan and FDG-PET after two cycles and after the completion of therapy. TAM in diagnostic specimens was determined by immunohistochemistry with a monoclonal antibody (anti-CD68 KP1). Overall, early FDG-PET was negative in 163 patients (81.5%) and positive in 37 patients (18.5%), showing a significant correlation with the achievement of CR ($p < 0.0001$). After a median follow-up of 40 months, progression free survival (PFS) was significantly better for PET negative patients ($p < 0.0001$). CD68 expression was low, intermediate or high in 26 (13%), 100 (50%) and 74 (37%) cases, without difference in the distribution between responders and non-responders. PFS analysis showed no significant difference in any score group. TAM score did not show any correlation with early FDG-PET result.This study confirms that early FDG-PET has a high prognostic power, while TAM score does not seem to influence the outcome; in contrast to our original hypothesis, it does not correlate with FDG-PET assessment.

Keywords: **Hodgkin's lymphoma; macrophages; CD68; PET; prognosis**

Introduction

Classic Hodgkin lymphoma (cHL) is a highly curable lymphoid malignancy that mostly affects young adults. Neoplastic cells, called Reed–Sternberg (R–S) cells, have a defective B-cell phenotype and represent just a minority of overall cell population, surrounded by reactive cells including T-lymphocytes, eosinophils, plasma cells and macrophages [1–3]. Treatment choice is made on the basis of clinical variables, such as Ann Arbor staging and the

presence of B symptoms, with the risk of both undertreat and over treat cHL patients. First-line therapy for early-stage disease is represented by a combined modality with ABVD (doxorubicin, bleomycin, vinblastine and dacarbazine) two to four cycles followed by involved-field radiotherapy (IF-RT); while for advanced-stage disease at least six cycles of ABVD without RT are recommended [4,5]. The number of ABVD cycles for early stage disease will depend on favourable or unfavourable risk categories, as established by EORTC criteria, that give a measure of disease tumour burden [6]. Despite satisfactory results, about 20% of patients still have relapsed/refractory disease while the rate of late toxic effects, often because of over treatment, continues to rise with time [7–9].

International prognostic score (IPS) was originally designed for advanced disease, is less suitable for early-stage disease and its power has weakened in the modern treatment era [10,11]. Moreover, none of the published prognostic factors can reliably identify patients in whom treatment is likely to fail, so there is a need of new markers to better stratify cHL patients at diagnosis and to consequently provide a truly tailored individual therapy [12].

About imaging prognostic factors, early FDG-PET assessment (after two cycles of chemotherapy, also called interim PET or PET2) in recent years demonstrated a high negative predictive value in limited-stage disease and a strong positive predictive value in advanced-stage disease [13,14]. In this field, a growing consent is emerging after the study of Gallamini and colleagues, showing a significantly lower 2-year progression-free survival (PFS) for patients with positive results (12.8% vs 95%, $p < 0.0001$), regardless of IPS [15]. Early FDG-PET seems to be very promising to assess chemosensitivity [16,17], avoiding overtreatment and late side effects. Currently its role in guiding risk and response-adapted therapy is under investigation in several studies. To date, revised response criteria for cHL define complete response (CR) when all previous lesions are PET negative, even if a residual mass persists by CT scan [18]. Notwithstanding there is no indication in current guidelines to tailor therapy administration on the basis of early FDG-PET results outside of clinical trials [5,19].

Tumour-associated macrophages (TAM), determined both by gene-expression profiling (GEP) and by immunohistochemistry (IHC) using the marker CD68, emerged as powerful prognostic predictors of disease-specific survival (DSS) in HL patients in an important paper by Steidl and colleagues, outperforming the IPS [20]. As previously hypothesized, these macrophages could be 'bad guys' in cHL, favouring tumour cells survival and leading to disease progression [12,21,22]. However, some consequent studies showed conflicting results about TAM association with survival and optimal cut-off [23–33].

Because of macrophages are the main component of inflammatory background surrounding RS cells and its invasion could be linked with a poor prognosis, we hypothesize, as previously published, that an abundance of TAM could inhibit cell death in response to cytotoxic treatments, leading to a more aggressive disease [22]. At our institution we speculated that the macrophage infiltration could be a consequence of the inflammatory microenvironment which is detected by PET and that could persist at early revaluation in patients who will finally fail to achieve a CR [34]. FDG uptake could reflect the amount of inflammation in terms of both metabolic activation and number of all inflammatory cells, thus we would like to investigate whether baseline CD68 expression could be a good predictor of early PET response and consequently of chemosensitivity.

In the study by Steidl and colleagues there was no mention to early FDG-PET and in our knowledge only 2 recent studies investigated this association with opposite results [35,36]. According to this background, in the present study the primary endpoint was to investigate, in a large cohort of patients with a long-term follow-up, the prognostic role of both early FDG-PET and TAM, while the secondary endpoint was to test if early FDG-PET positivity could correlate with high TAM score in diagnostic specimens.

Patients and methods

Patients

A cohort of 200 consecutive cHL patients with adequate available paraffin-embedded tissue diagnosed and treated at six Italian hematology institutions between March 2005 and December 2012 was retrospectively analysed. Diagnosis was made according to the World Health Organization classification of hematolymphoid neoplasms [37]; nodular lymphocyte predominant HL and HIV positive patients were excluded. Patients signed written informed consent in accordance with local Institutional Review Board requirements and the Declaration of Helsinki. Clinical and follow-up data were obtained from clinical records, and a centralized database was collected at the University of Siena. All patients completed staging with whole body CT scan, FDG-PET and bone marrow biopsy; at pretreatment evaluation a complete blood cell count and a biochemistry panel including erythrocyte sedimentation rate and lactate dehydrogenase were performed. Induction treatment plan consisted, according to staging and EORTC criteria, of two to six courses of ABVD and, if indicated, IF-RT [4,6]. For relapsed/refractory patients, salvage therapy was consistent with local guidelines.

Assessment of response

Response to treatment was assessed with CT scan and FDG-PET after two cycles and after the completion of

DOI: 10.1002/hon

therapy, according to revised response criteria for malignant lymphoma [18]. Patients not achieving CR were considered as treatment failures. FDG-PET evaluation was made according to Deauville criteria; bone marrow biopsy was performed in patients who had bone marrow involvement at diagnosis. Duration of remission was clinically evaluated every 3 months and by CT scan every 6 months for 2 years; standard radiology with chest X-ray and abdominal ultrasound was used thereafter. FDG-PET assessment was not performed during follow-up period.

Immunohistochemistry

Formalin-fixed, paraffin embedded diagnostic specimens including tumour cells from each patient were selected for immunohistochemical analysis. All diagnosis was centrally revised at the University of Siena.

TAM were determined as CD68+ cells using a monoclonal antibody (anti-CD68 KP1, Dako®), as previously published [20]. Briefly, 1.5-mm duplicate cores were obtained, and cells were examined in three representative high-power fields. The relative percentage of CD68+ cells was established in relation to the overall cellularity and reported as an average of the scores of both duplicate cores examined. Based on this percentage, patients were divided into three groups: $<5\%$ (low expression, score 1), 5–25% (intermediate expression, score 2) and $>25\%$ (high expression, score 3).

Imaging technique of PET-CT

PET studies have been performed in all the Institutions with the same scanner during years and elaborated with unchanged protocols of acquisition and reconstruction throughout the interval of enrollment. For each patient basal, interim and final PET scans were performed at the same centre and with the same instrument. Early FDG-PET was performed 12–14 days after second part of second cycle. End of treatment FDG-PET was assessed at least 30 days after last ABVD course and never earlier 90 days after treatment in patients receiving IF-RT. Scans were performed with a Discovery 690 ST scanner (GE Medical Systems) in three centres and with a Gemini CT-PET scanner (Philips Electronics) in the other three centres. Patients were injected with 18F-FDG after a 6-h fast; serum glucose level in all patients at the time of injection was below 160 mg/dl, and all scans were executed from the vertex to the upper thigh. Emission data were acquired 60 min post 18F-FDG administration; acquisition time was 25 min per PET scan. Patients were injected with 3.7 MBq/kg body weight of FDG; the same dose was employed at the centres that used a GE scanner and at those that used a Philips scanner. After completion of CT, PET emission data (4 min per bed position) was obtained; CT data was used for attenuation correction.

The imaging procedures were standardized and harmonized in all the six involved centres as a part of the qualification procedures requested by Fondazione Italiana Linfomi (FIL), of which the centres are members. All the interim and end-of-therapy PET studies were evaluated qualitatively, by visual analysis, and the Deauville five-point scale was employed to describe and report the PET findings.

A complete metabolic response was defined as absence of any pathological 18F-FDG uptake or presence in the lesions of minimal residual uptake less than/equal to the liver activity. Therefore a Deauville score 1 to 3 identified a negative result while score 4–5 was considered as positive result.

The PET images were analysed and reported on site, and a central review was not performed: this is a limitation of the study even if the application of the Deauville scale may have contributed to obtain an acceptable level of inter-centre reproducibility.

Statistics

Categorical variables were analysed using Fisher's exact test. When we compared the three groups of CD68 expression we performed the Freeman–Halton extension of the Fisher's exact probability test with 2×3 contingency tables. For survival analysis, primary endpoint was PFS, defined as the time from the first day of treatment until disease progression, relapse, death for any cause or last follow-up (censored); patients that did not achieve CR after induction therapy were censored at that point for the progression analysis [18]. OS was defined as the time from the first day of treatment until death for any cause. Survival curves were estimated using the method of Kaplan and Meier and log rank test for significant associations; a p value < 0.05 was considered statistically significant. All statistical analyses were done with software MedCalc, version 2.0.

Results

Characteristics of patients

Clinical characteristics of enrolled patients are illustrated in Table 1. Out of the 200 patients, median age was 33.5 years (range 18–80), with a male/female distribution of 105:95. At presentation, the Ann Arbor stage was I in 11 patients (5.5%), II in 92 patients (46%), III in 55 patients (27.5%) and IV in 42 patients (21%); B symptoms were present in 102 patients (51%). Overall, 73/200 (36.5%) patients had early-stage disease (I–IIA) while 127/200 (63.5%) patients had advanced-stage disease (IIB–IV). According to the WHO classification, histological subtypes were nodular sclerosis in 162/200 (81%), mixed cellularity in 28/200

DOI: 10.1002/hon

Table 1. Characteristics of patients

Characteristic	Number of patients (%)
Age: median [range]	33.5 years [18–80]
Sex	
Men	105/200 (52.5%)
Women	95/200 (47.5%)
Histology	
Nodular sclerosis	162/200 (81%)
Mixed cellularity	28/200 (14%)
Lymphocyte rich	7/200 (3.5%)
Lymphocyte depleted	3/200 (1.5%)
Ann Arbor stage	
I	11/200 (5.5%)
II	92/200 (46%)
III	55/200 (27.5%)
IV	42/200 (21%)
Early stage (I–IIA)	73/200 (36.5%)
Advanced stage (IIB–IV)	127/200 (63.5%)
B symptoms	102/200 (51%)
CD 68 expression	
Low (<5%, TAM score 1)	26/200 (13%)
Intermediate (5–25%, TAM score 2)	100/200 (50%)
High (>25%, TAM score 3)	74/200 (37%)

Abbreviations: TAM, tumour associated macrophages

(14%), lymphocyte rich in 7/200 (3.5%) and lymphocyte depleted in 3/200 (1.5%) patients. All patients were treated with curative intent and received ABVD therapy; IF-RT was administered to early-stage disease and, in advanced-stage, at investigator discretion, for residual or bulky disease.

CD 68 expression was low in 26/200 patients (13%, TAM score 1), intermediate in 100/200 (50%, TAM score 2) and high in 74/200 (37%, TAM score 3). There was a higher percentage of stage II patients ($p = 0.02$) and a trend toward a higher representation of advanced-stage disease in the group with TAM score 3 (Table 2).

Response to treatment

Response to treatment is summarized in Table 3 and Table 4. Overall, as reported in Table 3, 168 patients (84%) achieved CR at the end of treatment, while the remaining 32 patients (16%) were considered as treatment failure; in this group 10 patients (5%) obtained a PR and 22 patients were non-responders (NR). Of 168 patients achieving CR, 20 relapsed and all but one received salvage therapy at investigator discretion followed by autologous bone marrow transplantation (ABMT). Because treatment outcome in HL is often much different for early and advanced-stage disease, we have reported separately the results (Table 4). Out of 73 patients with early-stage disease, 65/73 achieved CR (89.0%), while the remaining 8 patients (11.0%) were considered as treatment failure; out of 127 patients with advanced-stage disease, 103 obtained CR (81.1%), while the remaining 24 patients (18.9%) did not. Response to treatment was not different between early and advanced-stage patients (p = n.s.).

Early FDG-PET was negative in 163/200 patients (81.5%) and positive in 37/200 patients (18.5%), with a high negative predictive value of 93% (95%CI 88–96%). Of responders, 152/168 (90.5%) were early FDG-PET negative and 16/168 (9.5%) were early FDG-PET positive; of non responders, 11/32 (34.4%) were early FDG-PET negative and 21/32 (65.6%) were early FDG-PET positive ($p < 0.0001$). Out of 20 relapsed patients, 14 were early FDG-PET negative, and six were early FDG-PET positive.

The correlation between negative result of early FDG-PET and achievement of CR remained significant even if patients with early-stage and advanced-stage disease were analysed separately. In early-stage disease 57/65 CRs (87.6%) were early FDG-PET negative, while 8/65 (12.4%) were early FDG-PET positive; out of non-responders, 6/8 (75%) were

Table 2. Characteristics of patients according to the CD68 expression

Characteristic	CD68 low–intermediate N (%)	CD68 high N (%)	*P*
Age: median [range]	33.5 years [18–80]	33.5 years [18–79]	n.s.
Sex			
Men	69/126 (54.7%)	36/74 (48.6%)	n.s.
Women	57/126 (45.3%)	38/74 (51.4%)	
Histology			
Nodular sclerosis	104/126 (82.5%)	58/74 (78.4%)	n.s.
Mixed cellularity	17/126 (13.5%)	11/74 (14.8%)	n.s.
Lymphocyte rich	4/126 (3.2%)	3/74 (4.1%)	n.s.
Lymphocyte depleted	1/126 (0.8%)	2/74 (2.7%)	n.s.
Ann Arbor stage			
I	6/126 (4.7%)	5/74 (6.7%)	n.s
II	66/126 (52.4%)	26/74 (35.2%)	0.02
III	32/126 (25.4%)	23/74 (31.1%)	n.s.
IV	22/126 (17.5%)	20/74 (27.0%)	n.s.
Early stage (I–IIA)	52/126 (41.3%)	21/74 (28.4%)	0.07
Advanced stage (IIB–IV)	74/126 (58.7%)	53/74 (71.6%)	
B symptoms	61/126 (48.4%)	41/74 (55.4%)	n.s.

DOI: 10.1002/hon

Table 3. Response to treatment

	Number of patients	%	*p*
CR (responders)	168/200	84%	
Treatment failure	32/200	16%	
PR	10/200	5%	
NR	22/200	11%	
Early FDG-PET POS	37/200	18.5%	
Early FDG-PET NEG	163/200	81.5%	
Responders			<0.0001
Early FDG-PET POS	16/168	9.5%	
Early FDG-PET NEG	152/168	90.5%	
Treatment failure			
Early FDG-PET POS	21/32	65.6%	
Early FDG-PET NEG	11/32	34.4%	
Responders			n.s.
TAM score 1	23/168	13.7%	
TAM score 2	85/168	50.6%	
TAM score 3	60/168	35.7%	
Treatment failure			
TAM score 1	3/32	9.4%	
TAM score 2	15/32	46.9%	
TAM score 3	14/32	43.7%	
Early FDG-PET POS			n.s.
TAM score 1	4/37	10.8%	
TAM score 2	22/37	59.5%	
TAM score 3	11/37	29.7%	
Early FDG-PET NEG			
TAM score 1	22/163	13.5%	
TAM score 2	78/163	47.9%	
TAM score 3	63/163	38.6%	

Abbreviations: CR, complete response; PR, partial response; NR, non-responders; TAM, tumour associated macrophages.

early FDG-PET positive, while only 2/8 (25.0%) were early FDG-PET negative (p =0.0004); in advanced-stage disease 95/103 CRs (92.2%) were early FDG-PET negative, while 8/103 (7.8%) were early FDG-PET positive; of non responder patients, 15/24 (62.5%) were early FDG-PET positive, while only 9/24 (37.5%) were early FDG-PET negative ($p < 0.0001$).

Correlation between TAM score, attainment of CR and early FDG-PET result

In responders, TAM score was 1 in 23/168 patients (13.7%), 2 in 85/168 (50.6 %) and 3 in 60/168 (35.7%); in patients who had treatment failure TAM score was 1 in 3/32 patients (9.4%), 2 in 15/32 (46.9%) and 3 in 14/32 (43.7%) (p = n.s., Table 3 and Figure 1). The lack of association between TAM score and attainment of CR was confirmed even if we separately analysed patients with early-stage and advanced-stage disease (Table 4). Of 65 early-stage responders, 10/65 patients (15.3%) had score 1, 36/65 (55.3%) had score 2 and 19/65 (29.4%) had score 3; of 8 non responders, 1/8 (12.5%) patients had score 1, 5/8 (62.5%) had score 2 and 2/8 patients (25%) had score 3 (p = n.s.). Of 103 advanced-stage responders, 13/103 patients (12.6%) had score 1, 49/103 (47.6%) had score 2 and 41/103 (39.8%) had score 3; of 24 non responders, 2/24 patients (8.3%) had score 1, 10/24 (41.7%) had score 2 and 12/24 (50%) had score 3 (p = n.s.).

With regard to the possible correlation between TAM score and early-FDG-PET result, of 163 PET negative patients, 22/163 (13.5%) had score 1, 78/163 had score 2 (47.9%) and 63/163 had score 3 (38.6%); of 37 PET positive patients, 4/37 (10.8%) had score 1, 22/37 (59.5%) had score 2 and 11/37 (29.7%) had score 3 (p = n.s., Table 3). The lack of association between TAM score and early FDG-PET result remains even if we separately analysed patients with early-stage and advanced-stage disease (Table 4). Out of 59 PET negative patients with early-stage disease, 9/59 (15.2%) had score 1, 32/59 (54.2%) had score 2 and 18/59 (30.6%) had score 3; while out of 14 PET positive patients, 2/14 (14.3%) had score 1, 9/14 (64.3%) had score 2 and 3/14 (21.4%) had score 3 (p = n.s.). Out of 104 PET negative patients with advanced-stage disease, 13/104 (12.5%) had score 1, 46/104 (44.2%) had score 2 and 45/104 (43.3%) had score 3; while out of 23 PET positive patients, 2/23 (8.7%) had score 1, 13/23 (56.5%) had score 2 and 8/23 (34.8%) had score 3 (p = n.s.).

Survival analysis

After a median follow-up of 40 months 194/200 patients (97%) were alive, and six patients died, five of progressive disease and one because of a second neoplasm (gastric cancer). There was a trend to a lower PFS for advanced-stage patients compared to early-stage, but it did not reach statistical significance (p = 0.07, 2-year PFS of 84% and 73%, respectively, as shown in Figure S1). Patients with negative early FDG-PET had longer PFS compared to patients with a positive result; median PFS was not reached vs 7 months ($p < 0.0001$, Figure 2); while TAM score was not associated with PFS (p = n.s., Figure 3). The HR was 1.833 for TAM score 3 vs 1 (95% CI 0.766–4.383) and of 1.534 for TAM score 3 vs 2 (95% CI 0.842–2.794), while the estimated 3-year PFS was 80%, 77% and 65% for the three different groups, respectively. The same results were obtained when we separately analysed patients with early-stage and advanced-stage disease (Figures S2–S5). Median PFS and OS in the whole cohort were not reached.

Discussion and conclusion

In this study we observed that (i) early FDG-PET is a strong prognostic factor in cHL, (ii) TAM score is not, at least in our hands, a significant factor for outcome

DOI: 10.1002/hon

Table 4. Response to treatment in early and advanced-stage disease

	Early-stage disease			Advanced-stage disease		
	Number of patients	**%**	***P***	**Number of patients**	**%**	***p***
CR (responders)	65/73	89.0%		103/127	81.1%	
Treatment failure	8/73	11.0%		24/127	18.9%	
Early FDG-PET POS	14/73	19.2%		23/127	18.1%	
Early FDG-PET NEG	59/73	80.8%		104/127	81.9%	
Responders			0.0004			<0.0001
Early FDG-PET POS	8/65	12.4%		8/103	7.8%	
Early FDG-PET NEG	57/65	87.6%		95/103	92.2%	
Treatment failure						
Early FDG-PET POS	6/8	75%		15/24	62.5%	
Early FDG-PET NEG	2/8	25%		9/24	37.5%	
Responders			n.s.			n.s.
TAM score 1	10/65	15.3%		13/103	12.6%	
TAM score 2	36/65	55.3%		49/103	47.6%	
TAM score 3	19/65	29,4%		41/103	39.8%	
Treatment failure						
TAM score 1	1/8	12.5%		2/24	8.3%	
TAM score 2	5/8	62.5%		10/24	41.7%	
TAM score 3	2/8	25%		12/24	50%	
Early FDG-PET POS			n.s.			n.s.
TAM score 1	2/14	14.3%		2/23	8.7%	
TAM score 2	9/14	64.3%		13/23	56.5%	
TAM score 3	3/14	21.4%		8/23	34.8%	
Early FDG-PET NEG						
TAM score 1	9/59	15.2%		13/104	12.5%	
TAM score 2	32/59	54.2%		46/104	44.2%	
TAM score 3	18/59	30.6%		45/104	43.3%	

prediction in cHL, and (iii) there is not any correlation between TAM score and early FDG-PET result.

HL is a highly curable neoplasm but about 30% of patients with advanced-stage disease and up to 15% with early-stage disease will be refractory or finally relapse; moreover prognostic factors in cHL are still a matter of debate and to date treatment strategy is principally determined by the assessment of tumour burden [4–6]. Malignant RS cells are surrounded by non-neoplastic cells, including TAM, that might influence tumour survival through a deep cross-talk [12,22]. TAM could reflect inflammation

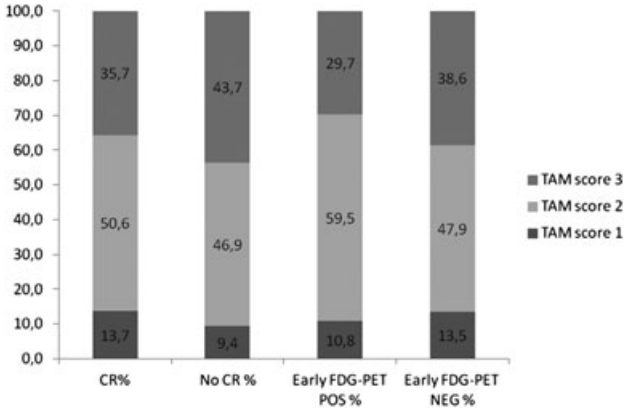

Figure 1. Correlation between TAM score, achievement of CR and early FDG-PET result

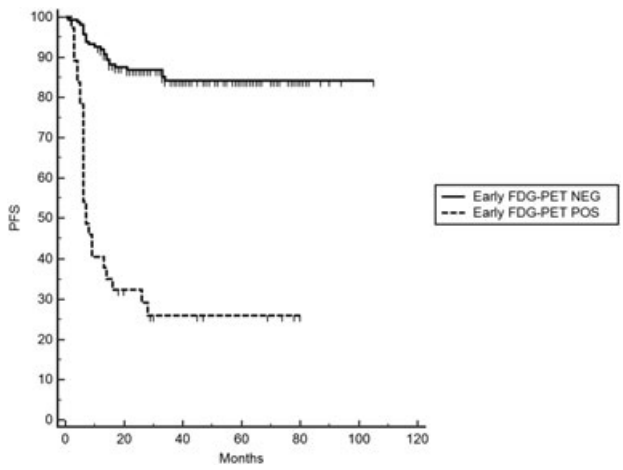

Figure 2. Correlation between early FDG-PET and PFS. Patients with negative early FDG-PET had longer PFS compared to patients with positive result; median PFS was not reached vs 7 months ($p < 0.001$)

DOI: 10.1002/hon

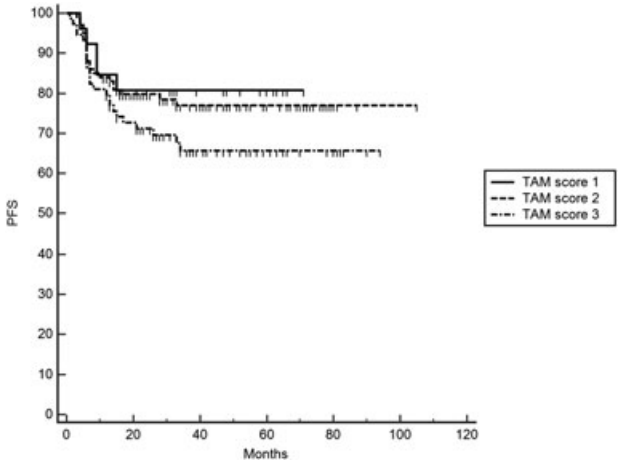

Figure 3. Correlation between TAM score and PFS. High TAM score did not show any correlation with PFS

and the non-malignant component and might be detected by PET [34].

The prognostic value of early FDG-PET is obtaining a growing consent, especially in advanced-stage disease where it is outperforming the IPS [15]. A recent multicentre FIL study evaluated 246 newly diagnosed HL patients with early-stage disease (I–IIA) treated with 4 cycles of ABVD followed by IF-RT to assess the predictive role of early FDG-PET on 2-year FFS. The positive and negative early FDG-PET predictive value was 53% and 95%, while the sensitivity and specificity were 65.5% and 92%, respectively. After a mean follow-up of 46 months, a positive early FDG-PET result was the only one factor that negatively affected FFS in a multivariate analysis, suggesting that interim PET could be considered a prognostic test also in early-stage HL [38].

Although many phase II or III trials are ongoing, to date an interim PET-driven approach is not yet suggested in clinical practice [5,19]. About imaging evaluation, the Deauville score has been accepted as the standard because it is simple and allows reproducible results [16,17].

Since 2010, when Steidl and colleagues analysed TAM prognostic value in GEP and IHC, we were very interested to this observation, because they realized an easy score with a strong association with PFS and DSS, useful both at diagnosis and at relapse [20]. In this study was observed a relationship between high TAM score and treatment outcome while in our study we did not identify any association; these different results are not surprising because of the different cohort of patients. First, in our study we analysed 200 consecutive HL patients, and our cohort is characterized by a relatively good prognosis because 51% of patients had stage I–II disease; conversely the Canadian cohort was enriched for all available cases of treatment-failure identified through the database of their Institution. In our study ORR was 84% with only 16% of chemo-resistant patients, while in the original report 79 out of 166 patients (47%) had a treatment-failure. Another reason explaining the different findings obtained could be identified in the choice of monoclonal antibodies addressing different epitopes, because we used KP-1 while for Steidl and colleagues the assessed epitope was PGM-1.

Despite promising results of this and other subsequent studies, other authors found no correlation between TAM (determined by CD68 expression) and outcome, as showed in Table 5 [23–33;35,36]. These discordant results could be because of the choice of TAM markers, the antibodies used for IHC and the choice of the cut-offs [12,21,27,29]. Azambuja stated an overall lack of reproducibility for IHC analysis, while Klein and colleagues suggested CD163 ensures greater reproducibility and could be superior to CD68 in predicting OS in cHL patients [27,29]. A possible explanation is that CD68 could stain both myelomonocytic cells and some connective tissue cells such as fibroblasts and endothelial cells. Steidl and colleagues used the KP1 antibody but, as declared in subsequent papers, KP1 clone for CD68 could react with myeloid and fibroblast cells, too; however the other available antibody PGM1 showed discordant association with prognosis, too [27,36]. Although the reproducibility of the Steidl score appears very low, when other cut-off values were adopted (e.g. moving from the ROC curves analysis) a relationship could be found. Yoon and colleagues performed separated analyses at different cut-off points and they determined the best cut-off of 20% to divide the patients in two groups employing the log-rank test, finding an association with EFS, DSS and OS [25]. Greaves defined three prognostic groups with low, intermediate and high CD68 expression ($<5\%$, 5% to 15% and $>15\%$), high TAM score was associated with inferior FFTF and OS [26]. Kamper and colleagues found a correlation between CD68$\geq$7.8% and lower EFS and OS, while Tan and colleagues found an optimum threshold of 12.7% to identify the group with reduced FFS and OS [23,31].

Only two recent studies investigated a possible correlation between TAM score and early FDG-PET result. Touati and colleagues analysed a cohort of 158 patients enrolled between 1995 and 2011 and found a correlation between CD68 $\geq$25% and inferior PFS and OS [35]. Early FDG-PET showed an association with high CD68 expression, but there are some biases because it was done in only 68 patients, treatment strategy was not uniform and data since 2005 (performed by a dual head coincidence gamma camera) were analysed retrospectively using the Deauville score. Agur and colleagues enrolled 98 patients with a median follow-up of 45 months and did not observe any correlation between TAM score, early FDG-PET and PFS [36].

In our study, 200 consecutive cHL patients were prospectively enrolled, and all of them underwent early FDG-PET scan. Characteristics of patients and response to treatment are consistent with previously published data, with a CR rate of 84%, PFS at 2 and 5 years of 76% and 73%,

DOI: 10.1002/hon

Table 5. Prognostic role of CD68+ tumour associated macrophages (TAM) in classical Hodgkin lymphoma

Study	Method	Patients number	Cut-off	Outcome correlation
Kayal and colleagues [33]	IHC	100	• Steidl score (<5%, 5–25%, >25)	No
			• <25%, ≥25% (lowest quartile vs others)	No
			• <12.9%, ≥12.9% (lowest quartile vs others)	No
			• <18.2, ≥18.2 (median)	No
Casulo and colleagues [28]	IHC	81 (relapsed/ refractory)	• Steidl score	No
			• <30%, ≥30%	Inferior OS in univariate, not in multivariate
Azambuja and colleagues [27]	IHC	265	Steidl score	No
Klein and colleagues [29]	IHC	88	<25%, ≥25%	No
Agur and colleagues [36]	IHC	98	Steidl score	No
Steidl and colleagues [20]	GEP, IHC	166	Steidl score	Inferior PFS, DSS
Yoon and colleagues [25]	IHC	144	<20%, ≥20%	Inferior EFS, DSS, OS
Tan and colleagues [23]	IHC	287	<12.7%, ≥12.7%	Inferior FFS, OS
Greaves and colleagues [26]	IHC	90	<5%, 5% to 15%, >15%	Inferior FFTF, OS
Jakovic and colleagues [30]	IHC	52	<25%, ≥25%	Trend to inferior EFS, inferior OS
Touati and colleagues [35]	IHC	158	<25%, ≥25%	Inferior PFS, OS
Kamper and colleagues [31]	IHC	288	<7.8%, ≥7.8%	Inferior EFS, OS
Deau and colleagues [32]	IHC	59	Steidl score	Refractoriness or early relapse
Tzankov and colleagues [24]	IHC	105	>0.82%	Inferior OS

Abbreviations: IHC, Immunohistochemistry; OS, overall survival; PFS, progression free survival; DSS, disease specific survival; EFS, event free survival; FFS, failure free survival; FFTF, freedom from treatment failure.

respectively, with a small rate of late relapses. According to the clinical risk, the number of ABVD cycles ranged from 2 to 6; in the subgroup receiving two cycles the PET scan performed after the second cycle was an End of Treatment scan and not a 'standard' interim PET, but it was the case of only 2 stage IA patients, not impacting on the overall study results. All the other 9 stage IA patients received 4 ABVD cycles because of unfavourable characteristics.

In our cohort the proportion of patients achieving CR was similar for early and advanced-stage disease (89% vs 81.1% respectively), while patients with advanced-stage disease had a trend to a lower PFS ($p = 0.07$). In the original Canadian cohort the proportion of patients with advanced-stage disease was higher than in our cohort (75.3% vs 63.5%) and advanced-stage was associated with shorter PFS both in univariate and in multivariate analysis ($p = 0.002$ and $p = 0.001$, respectively).

Negative early FDG-PET results were associated with attainment of CRs and an improved PFS in the whole cohort. This correlation was further confirmed in both the early and advanced-stage subsets, further highlighting the strong prognostic value for early FDG-PET result in all stages of cHL.

Because TAM score did not show any correlation with treatment response and with PFS, we believe that it is not a useful prognostic tool for clinical application, at least until a standardized method of determination will be available. In contrast to our original hypothesis, TAM score does not correlate with FDG-PET assessment.

In conclusion, we suggest to investigate in future multicentre studies the role of macrophage subpopulations and a possible correlation between early FDG-PET and the presence of new prognostic markers (such as FOXP3, CD20 and bcl-2), with the goal of providing a better stratification of cHL patients at diagnosis and proposing to each patient the most adequate therapeutic strategy.

Conflict of interest

The authors report no declaration of interests.

Author contributions

E.C. and A.F. designed the study, analysed and interpreted data, performed statistical analysis and drafted the manuscript; S.L., G.G., A.D.N., S.F., S.D.L. and L.L. performed immunohistochemistry and interpreted data; S.L. and L.L. revised diagnostic specimens; S.K. and R.B. collected clinical data; L.R., A.B., G.G., .M.C.C., S.M., E.A. and M.B. analysed and interpreted clinical data; M.B. finally revised and approved the manuscript.

References

1. Re D, Küppers R, Diehl V. Molecular pathogenesis of Hodgkin's lymphoma. *J Clin Oncol* 2005; **23**: 6379–6386.

DOI: 10.1002/hon

2. Thomas RK, Re D, Wolf J, Diehl V. Part I: Hodgkin's lymphoma—molecular biology of Hodgkin and Reed-Sternberg cells. *Lancet Oncol* 2004; **5**: 11–18.
3. Gobbi PG, Ferreri AJ, Ponzoni M, Levis A. Hodgkin lymphoma. *Crit Rev Oncol Hematol* 2013; **85**: 216–237.
4. Brusamolino E, Bacigalupo A, Barosi G, *et al.* Classical Hodgkin's lymphoma in adults: guidelines of the Italian Society of Hematology, the Italian Society of Experimental Hematology, and the Italian Group for Bone Marrow Transplantation on initial work-up, management, and follow-up. *Haematologica* 2009; **94**: 550–565.
5. Eichenauer DA, Engert A, André M, *et al*. ESMO Guidelines Working Group. Hodgkin's lymphoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* 2014; **25**(Suppl 3): 70–75.
6. Noordijk EM, Carde P, Dupouy N, *et al*. Combined-modality therapy for clinical stage I or II Hodgkin's lymphoma: long-term results of the European Organisation for Research and Treatment of Cancer H7 randomized controlled trials. *J Clin Oncol* 2006; **24**: 3128–3135.
7. Diehl V, Fuchs M. Early, intermediate and advanced Hodgkin's lymphoma: modern treatment strategies. *Ann Oncol* 2007; **18**: 71–79.
8. Björkholm M, Axdorph U, Grimfors G, *et al*. Fixed versus response-adapted MOPP/ABVD chemotherapy in Hodgkin's disease. A prospective randomized trial. *Ann Oncol* 1995; **6**: 895–899.
9. Cellai E, Magrini SM, Masala G, *et al*. The risk of second malignant tumors and its consequences for the overall survival of Hodgkin's disease patients and for the choice of their treatment at presentation: analysis of a series of 1524 cases consecutively treated at the Florence University Hospital. *Int J Radiat Oncol Biol Phys* 2001; **49**: 1327–1337.
10. Hasenclever D, Diehl V. A prognostic score for advanced Hodgkin's disease. International Prognostic Factors Project on Advanced Hodgkin's Disease. *N Engl J Med* 1998; **339**: 1506–1514.
11. Moccia AA, Donaldson J, Chhanabhai M, *et al*. International Prognostic Score in advanced-stage Hodgkin's lymphoma: altered utility in the modern era. *J Clin Oncol* 2012; **30**: 3383–3388.
12. Diehl V. Hematology. Are macrophages the bad guys in Hodgkin lymphoma? *Nat Rev Clin Oncol* 2010; **7**: 301–302.
13. 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**: 52–59.
14. Hutchings M, Mikhaeel NG, Fields PA, Nunan T, Timothy AR. Prognostic value of interim FDG-PET after two or three cycles of chemotherapy in Hodgkin lymphoma. *Ann Oncol* 2005; **16**: 1160–1168.
15. 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**: 3746–3752.
16. Gallamini A, Barrington SF, Biggi A, *et al*. 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. *Haematologica* 2014; **99**: 1107–1113.
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**: 683–690.
18. Cheson BD, Pfistner B, Juweid ME, *et al*. International Harmonization Project on Lymphoma. Revised response criteria for malignant lymphoma. *J Clin Oncol* 2007; **25**: 579–586.
19. Cheson BD, Fisher RI, Barrington SF, *et al*. Recommendations for initial evaluation, staging, and response assessment of Hodgkin and non-Hodgkin lymphoma: the Lugano classification. *J Clin Oncol* 2014; **32**: 3059–3068.
20. Steidl C, Lee T, Shah SP, *et al*. Tumor-associated macrophages and survival in classic Hodgkin's lymphoma. *N Engl J Med* 2010; **362**: 875–885.
21. Steidl C, Farinha P, Gascoyne RD. Macrophages predict treatment outcome in Hodgkin's lymphoma. *Haematologica* 2011; **96**: 186–189.
22. DeVita VT Jr. A breakthrough in Hodgkin's disease. *Nat Rev Clin Oncol* 2010; **7**: 179.
23. Tan KL, Scott DW, Hong F, *et al*. Tumor-associated macrophages predict inferior outcomes in classic Hodgkin lymphoma: a correlative study from the E2496 Intergroup trial. *Blood* 2012; **120**: 3280–3287.
24. Tzankov A, Matter MS, Dirnhofer S. Refined prognostic role of CD68-positive tumor macrophages in the context of the cellular micromilieu of classical Hodgkin lymphoma. *Pathobiology* 2010; **77**: 301–308.
25. Yoon DH, Koh YW, Kang HJ, *et al*. CD68 and CD163 as prognostic factors for Korean patients with Hodgkin lymphoma. *Eur J Haematol* 2012; **88**: 292–305.
26. Greaves P, Clear A, Coutinho R, *et al*. Expression of FOXP3, CD68, and CD20 at diagnosis in the microenvironment of classical Hodgkin lymphoma is predictive of outcome. *J Clin Oncol* 2013; **31**: 256–262.
27. Azambuja D, Natkunam Y, Biasoli I, *et al*. Lack of association of tumor-associated macrophages with clinical outcome in patients with classical Hodgkin's lymphoma. *Ann Oncol* 2012; **23**: 736–742.
28. Casulo C, Arcila M, Bohn OL, Teruya-Feldstein J, Maragulia J, Moskowitz CH. Tumor associated macrophages in relapsed and refractory Hodgkin lymphoma. *Leuk Res* 2013; **37**: 1178–1183.
29. Klein JL, Nguyen TT, Bien-Willner GA, *et al*. CD163 immunohistochemistry is superior to CD68 in predicting outcome in classical Hodgkin lymphoma. *Am J Clin Pathol* 2014; **141**: 381–387.
30. Jakovic LR, Mihaljevic BS, Perunicic Jovanovic MD, Bogdanovic AD, Andjelic BM, Bumbasirevic VZ. The prognostic relevance of tumor associated macrophages in advanced stage classical Hodgkin lymphoma. *Leuk Lymphoma* 2011; **52**: 1913–1919.
31. Kamper P, Bendix K, Hamilton-Dutoit S, Honoré B, Nyengaard JR, d'Amore F. Tumor infiltrating macrophages correlate with adverse prognosis and Epstein–Barr virus status in classical Hodgkin's lymphoma. *Haematologica* 2011; **96**: 269–276.
32. Deau B, Bachy E, Ribrag V, *et al*. Macrophage, mast cell and T lymphocyte infiltrations are independent predictive biomarkers of primary refractoriness or early relapse in classical Hodgkin lymphoma. *Leuk Lymphoma* 2013; **54**: 41–45.
33. Kayal S, Mathur S, Karak AK, *et al*. CD68 tumor-associated macrophage marker is not prognostic of clinical outcome in classical Hodgkin lymphoma. *Leuk Lymphoma* 2014; **55**: 1031–1037.
34. Fabbri A, Gozzetti A, Forconi F, Lauria F. Predicting the clinical course of Hodgkin lymphoma. *Nat Rev Clin Oncol* 2010; **7**.
35. Touati M, Delage-Corre M, Monteil J, *et al*. CD68-positive tumor-associated macrophages predict unfavorable treatment outcomes in classical Hodgkin lymphoma in correlation with interim fluorodeoxyglucose-positron emission tomography assessment. *Leuk Lymphoma* 2015; **56**: 332–341.

DOI: 10.1002/hon

36. Agur A, Amir G, Paltiel O, *et al*. CD68 staining correlates with the size of residual mass but not with survival in classical Hodgkin lymphoma. *Leuk Lymphoma* 2015; **56**: 1315–1319.
37. Swerdlow S, Campo E, Harris N, *et al*. *WHO Classification of Tumors of Haematopoietic and Lymphoid Tissues*. IARC Press: Lyon, 2008.
38. Rigacci L, Puccini B, Zinzani PL, *et al*. The prognostic value of positron emission tomography performed after two courses (INTERIM-PET) of standard therapy on treatment outcome in early stage Hodgkin lymphoma: a multicentric study by the Fondazione Italiana Linfomi (FIL). *Am J Hematol* 2015; **90**: 499–503.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site.

DOI: 10.1002/hon