Impact of risk factors on outcomes in early-stage Hodgkin’s lymphoma: an analysis of international staging definitions

B. Klimm1*, H. Goergen1, M. Fuchs1, B. von Tresckow1, B. Böll1, J. Meissner2, A. Glunz3, V. Diehl1, H. T Eich4, A. Engert1 & P. Borchmann1

1Department I of Internal Medicine, German Hodgkin Study Group (GHSG), University Hospital Cologne; 2Department V of Internal Medicine, University Hospital Heidelberg; 3Center of Internal Medicine, Department of Hematology, University Hospital Essen; 4Department of Radiotherapy, University Hospital Münster, Germany

Received 12 June 2013; revised 25 August 2013; accepted 28 August 2013

Background: In early-stage Hodgkin’s lymphoma (HL), treatment according to the early favorable or unfavorable subgroup is guided by staging definitions, which differ between various study groups worldwide. We analyzed risk factors used in different international staging systems and their impact on the outcome of early-stage HL patients.

Patients and methods: In 1173 early-stage HL patients treated homogenously within the German Hodgkin Study Group (GHSG) trials HD10 and HD11, the impact of three staging systems developed and used by the GHSG, the European Organization for Research and Treatment of Cancer (EORTC), and the National Comprehensive Cancer Network (NCCN) in discriminating risk groups for progression-free survival (PFS) and overall survival (OS) was assessed and the relevance of their single risk factors was investigated.

Results: All the three staging systems defined an unfavorable risk group out of early-stage patients of comparable size (56%, 55%, and 57%), having a significantly poorer PFS and OS as compared with the corresponding favorable group; 5-year differences between early favorable and early unfavorable in terms of PFS were 9.4% (HR 2.61, 95% CI 1.74–3.91), 6.7% (HR 2.10, 95% CI 1.41–3.13), and 8.6% (HR 2.14, 95% CI 1.45–3.16) with the GHSG, EORTC, and NCCN definition, respectively. Sensitivity was high for all systems (84%, 79%, and 83%); however, there was a low specificity with high rates of false-positive results (1-specificity 54%, 53%, and 55%, respectively). Models of high sensitivity included risk factors associated with large tumor burden and high tumor activity. Most risk factors for tumor-specific end points were also predictive of OS.

Conclusions: Differentiating between a favorable and an unfavorable risk group has significant impact on PFS and OS in early-stage HL patients in the modern treatment era. Risk-adapted treatment strategies using new risk factors with higher specificity are needed.

Key words: chemotherapy, combined modality treatment, early-stage Hodgkin’s lymphoma, outcome, risk factor

introduction

Patients with early-stage Hodgkin’s lymphoma (HL) are often subdivided into early favorable and early unfavorable stages and assigned to a corresponding treatment strategy. The differentiation has led to refined risk-adapted treatment protocols with high cure rates for HL patients in recent years. However, risk factor definitions for discrimination of early-stage patients differ between various study groups worldwide and the need for the definition of an early unfavorable group has repeatedly been questioned [1–5].

The definitions of risk factors have been developed over several generations of clinical trials. Most of these factors were originally defined in the radiotherapy era and proof of their impact and individual weight in the modern treatment era is missing. Furthermore, current combined modality treatment protocols that achieve better overall tumor control might have blunted the influence of certain risk factors on treatment outcomes. In addition, treatment today is often directly adapted to staging and risk factor existence and consequently differs between patients in the early favorable and unfavorable group. Thus, assessing the relevance of single risk factors without treatment bias is difficult and no larger cohort of uniformly treated early-stage patients has been analyzed thus far.

A large number of early-stage favorable and unfavorable patients were treated homogenously within our GHSG HD10 and HD11 trials with four cycles of ABVD (doxorubicin, bleomycin, vinblastine, and dacarbazine)
and consolidating involved-field radiotherapy (IFRT) [1, 2]. To shed more light on the ongoing question whether we really need separate early-stage subgroups, we used this dataset to analyze different international staging definitions and the significance of their risk factors on treatment outcomes in early-stage HL.

patient and methods

patient selection

Between May 1998 and January 2003, patients with newly diagnosed early-stage HL [clinical stages (CS) I and II] were included in the GHSG trials HD10 for early favorable stage HL \( (N = 1570) \) and HD11 for early unfavorable stage HL \( (N = 1370) \). In both studies, four cycles of ABVD was the standard chemotherapy, and patients were randomly assigned to receive either 20 or 30 Gy of consolidating IFRT [1, 2]. Except for the presence of risk factors, inclusion criteria did not differ between these studies. The details of inclusion criteria and staging procedures were described elsewhere [1, 2]. Both studies were registered at ClinicalTrials.gov \( (NCT00265018 \) and \( NCT00264953) \) and recently published [1, 2].

One thousand one hundred and seventy-three early-stage HL patients aged between 16 and 59 years from HD10 \( (525 \) patients) and HD11 \( (648 \) patients) qualified for the current analysis, including only patients treated with four cycles of ABVD followed by IFRT of 20 or 30 Gy. This cohort was chosen in order to assess the impact of risk factors on the outcome, without having a therapy bias by direct treatment adaptation according to the risk profile. Patients ≥60 years were not included, given that they are a risk group themselves that might not benefit from an intensified treatment for unfavorable stages [6, 7].

staging definitions

Staging definitions from the German Hodgkin Study Group (GHSG), the European Organization for Research and Treatment of Cancer (EORTC), and the National Comprehensive Cancer Network (NCCN) were evaluated in this analysis. According to these definitions, early-stage patients with CS I-II are classified as unfavorable if at least one of the risk factors shown in Table 1 is present [1–3, 5]. The staging definition of the National Cancer Institute of Canada Clinical Trials Group (NCIC-CTG) was not analyzed, since the corresponding early-stage patients and their inclusion criteria were not applicable to the available patient population due to exclusion of all patients with B-symptoms and larger tumor masses [4].

Table 1. Risk factor definitions in early-stage Hodgkin’s lymphoma (HL)

<table>
<thead>
<tr>
<th>GHSG</th>
<th>EORTC</th>
<th>NCCN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large mediastinal mass (ratio ≥ 1/3)*</td>
<td>Large mediastinal mass (ratio ≥ 0.35)</td>
<td>Large mediastinal mass (ratio &gt; 1/3)</td>
</tr>
<tr>
<td>≥1 extranodal lesion*</td>
<td>Age ≥ 50 years</td>
<td>Bulk &gt; 10 cm</td>
</tr>
<tr>
<td>ESR ≥ 50 (A) or ≥ 30 (B)</td>
<td>ESR ≥ 50 (A) or ≥ 30 (B)</td>
<td>ESR ≥ 50 (A)</td>
</tr>
<tr>
<td>≥3 nodal areas (out of 11 GHSG areas)</td>
<td>≥4 nodal areas (out of 5 supra-diaphragmatic. EORTC areas)</td>
<td>≥4 nodal regions (out of 17 Ann Arbor regions)</td>
</tr>
</tbody>
</table>

Patients in CS I-II are staged unfavorable, if at least one of the listed risk factors is present.

*Within the GHSG system, patients with CSIIb and a large mediastinal mass or ≥1 extranodal lesion are considered as advanced stage.

HL, Hodgkin’s Lymphoma; ESR, erythrocyte sedimentation rate; A, without B-symptoms; B, with B-symptoms; GHSG, German Hodgkin Study Group; EORTC, European Organization for Research and Treatment of Cancer; NCCN, National Comprehensive Cancer Network.

statistical analysis

The primary end point was HL-related failure, defined as salvage therapy (after no change or partial response) or relapsed or progressive HL, within 2.5 years after start of therapy. Patients observed shorter than 2.5 years without HL-related failure were excluded from analyses regarding this end point (2.5 year analysis set). Secondary end points included death within 5 years, PFS and OS. The detailed description of endpoints and statistical methods is given in supplementary Appendix 1, available at Annals of Oncology online.

results

patient characteristics

The flow charts of patients with the different analysis sets for primary and secondary end points are depicted in supplementary Figure S1, available at Annals of Oncology online. The median age of the 1173 qualified patients was 33 years (16–59 years). In total, 53% of patients were males and 47% females; 51% received 30 Gy of IFRT and 49% received 20 Gy. The median observation time of all qualified patients was 80 and 91 months for calculation of PFS and OS, respectively. Ninety patients (8% of the 2.5-year analysis set) had a HL-related failure within 2.5 years; 30 patients (3% of the 5-year analysis set) died within 5 years.

significance of an early unfavorable risk group

All three staging definitions led to a comparable proportion of early unfavorable patients; 55%, 54%, and 57% of patients presented with at least one risk factor at initial diagnosis when applying the GHSG, EORTC, and NCCN system, respectively. To assess whether discrimination between a favorable and an unfavorable risk group in early-stage patients according to their initial risk profile is reasonable, Kaplan–Meier analyses on PFS were carried out using the GHSG, EORTC, or NCCN staging definitions. As demonstrated in Figure 1, all three staging definitions truly selected an unfavorable risk group out of the early-stage patients, having a significantly poorer outcome \( (P < 0.001) \) when compared with the corresponding favorable subgroup. The 5-year difference in PFS was 9.4% for the GHSG definition, 6.7% for the EORTC definition, and 8.6% for the NCCN definition.
Figure 1. Kaplan–Meier analyses on progression-free survival (PFS) using staging definitions of the German Hodgkin Study Group (GHSG), the European Organization for Research and Treatment of Cancer (EORTC), or National Comprehensive Cancer Network (NCCN).
comparison of staging definitions and risk factor analysis

Sensitivity with respect to the primary end point was high for all the three systems (84%, 79%, and 83% with GHSG, EORTC, and NCCN system, respectively), suggesting that the high proportion of patients experiencing HL-related failure within 2.5 years was correctly classified in the unfavorable risk group. However, there was also a high rate of false-positive results (1-specificity); 54%, 53%, and 55% of patients not having an event later on were classified unfavorable.

In univariate analyses, most factors from these three definitions as well as some other candidate risk factors such as presence of bulky disease ≥5 cm or factors from the International Prognostic Score (IPS) for advanced stage HL [8] had a relevant impact on the event rate (supplementary Table S1, available at Annals of Oncology online). Exceptions included the commonly used risk factors B-symptoms and age; reasons for non-significance in this setting are discussed below. In multivariate analyses, the GHSG staging definition had 4/4, the EORTC definition 2/4, and the NCCN definition 3/5 risk factors with significant impact ($P < 0.05$) on the event rate (Table 2).

Regarding the corresponding ROC curves of the three systems, the specificity of the existing definitions could not substantially be improved by weighting factors or changing cut-off points without diminishing sensitivity <80% (supplementary Figure S2, available at Annals of Oncology online). Furthermore, when including all univariately relevant candidate risk factors in a multivariate logistic regression model, no relevant improvement of specificity was achievable without considerable loss of sensitivity (data not shown). However, all models of high sensitivity included especially those risk factors indicating large tumor burden, such as large mediastinal mass or involvement of a certain number of lymph node areas, and those pointing at a systemic inflammation, such as an elevated erythrocyte sedimentation rate (ESR).

impact of risk factors and risk groups on overall survival

To assess whether risk factors for tumor-specific end points were also predictive of overall survival (OS), univariate analyses were also carried out in the 5-year analysis set; no multivariate analyses were conducted due to the small number of events. In line with the results on tumor-specific endpoints, sensitivity regarding death within 5 years was high for all three systems (80%, 89%, and 83% with GHSG, EORTC, and NCCN system, respectively), while there was also a high rate of false-negative results (1-specificity 55%, 54%, and 56%, respectively). Factors indicating a large tumor size and high tumor activity were similarly relevant for OS (supplementary Table S1, available at Annals of Oncology online), in particular a large mediastinal mass ($P < 0.001$), bulky disease >10 cm ($P < 0.001$), and high ESR ($P = 0.011$). However, involvement of a certain number of nodal areas had a lower impact (all $P > 0.05$). In addition, age was a crucial prognostic factor for OS ($P = 0.006$ for age ≥50 years).

Kaplan–Meier analyses on OS of the risk groups identified by the three staging definitions are presented in Figure 2. Small but significant differences in OS were detected between patients classified as early favorable or unfavorable ($P = 0.006, P < 0.001$, and $P = 0.002$, respectively) with 5-year differences of 2.6%, 3.3%, and 2.9% for the GHSG, EORTC, and NCCN definition, respectively.

discussion

Although treatment decisions in early-stage HL are guided by the presence of risk factors, no larger analysis has evaluated the impact of these risk factors in the modern combined modality strategy.

Table 2. Multivariate analyses of staging systems

<table>
<thead>
<tr>
<th></th>
<th>2.5-year analysis set, $N = 1107$</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds ratio for HL-failure within 2.5 years</td>
<td>95% Confidence limits</td>
<td>$P$ value</td>
<td></td>
</tr>
<tr>
<td>GHSG system, $N = 1107$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large mediastinal mass (ratio &gt; 1/3)</td>
<td>3.3</td>
<td>2.0–5.5</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Extranodal disease</td>
<td>2.3</td>
<td>1.1–4.8</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>ESR ≥ 50 mm/h (A) or ≥ 30 mm/h (B)</td>
<td>1.6</td>
<td>1.0–2.5</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>≥3 nodal areas (out of 11 GHSG areas)</td>
<td>2.6</td>
<td>1.6–4.1</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>EORTC system, $N = 1018$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large mediastinal mass (ratio &gt; 1/3)</td>
<td>3.9</td>
<td>2.4–6.4</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Age ≥ 50 years</td>
<td>0.8</td>
<td>0.3–2.0</td>
<td>0.6</td>
<td></td>
</tr>
<tr>
<td>ESR ≥ 50 mm/h (A) or ≥ 30 mm/h (B)</td>
<td>1.5</td>
<td>0.9–2.4</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>≥4 nodal areas (out of 5 supradiaphragmatic EORTC areas)</td>
<td>2.1</td>
<td>1.3–3.4</td>
<td>0.003</td>
<td></td>
</tr>
<tr>
<td>NCCN system, $N = 1077$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large mediastinal mass (ratio &gt; 1/3)</td>
<td>2.2</td>
<td>1.1–4.3</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Bulky disease &gt; 10 cm</td>
<td>2.0</td>
<td>1.0–4.0</td>
<td>0.046</td>
<td></td>
</tr>
<tr>
<td>ESR ≥ 50 mm/h</td>
<td>1.6</td>
<td>1.0–2.5</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>B-Symptoms</td>
<td>1.0</td>
<td>0.5–1.9</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>≥4 nodal regions (out of 17 Ann Arbor regions)</td>
<td>2.4</td>
<td>1.5–3.8</td>
<td>&lt;0.001</td>
<td></td>
</tr>
</tbody>
</table>

HL, Hodgkin’s Lymphoma; ESR, erythrocyte sedimentation rate; GHSG, German Hodgkin Study Group; EORTC, European Organization for Research and Treatment of Cancer; NCCN, National Comprehensive Cancer Network.
The aim of this retrospective study thus was to answer the question if the risk factors still have an impact on treatment outcome and if the different staging systems using these risk factors reliably define a cohort of poor prognosis early-stage HL patients. The following major findings emerge from our analysis:

Figure 2. Kaplan–Meier analyses on overall survival (OS) using staging definitions of the GHSG, EORTC, or NCCN.
The staging definitions used by the GHSG, EORTC, and NCCN define an unfavorable risk group of comparable size out of the early-stage patients, having a significantly poorer PFS and OS compared with the corresponding favorable risk group. All three staging definitions demonstrate high sensitivity but low specificity for tumor-specific as well as survival end points. Multivariate models of high sensitivity for the tumor-specific end point include risk factors indicating large tumor burden and systemic inflammation. Most risk factors for tumor-specific end points are also predictive of OS.

Risk factors defining treatment groups in HL were empirically defined decades ago when treatment often was confined to radiotherapy alone [9–11]. However, knowledge on their current relevance and impact is missing. Validation of these old findings is needed, since risk factors might have lost their significance over time with better treatment strategies [12]. In 1173 early-stage HL patients homogenously treated with four cycles of ABVD followed by IFRT, particularly large tumor burden and an elevated ESR were shown as high impact risk factors. Surprisingly, the commonly used risk factors B-symptoms and age (with the EORTC cut-off of 50 years as well as the IPS cut-off of 45 years) had no significant impact on the tumor-specific event rate. This may be caused by the pre-selection of patients in our analysis: patients presenting with CS IIB with risk factor large mediastinal mass and/or extranodal lesions are considered advanced stage HL within the GHSG staging system and therefore not enrolled into the trials for early-stage HL. Thereby, the prognostic relevance of the potential risk factor B-symptoms might have been impaired. Furthermore, we excluded patients ≥60 years, assuming that the poorer prognosis of an unfavorable subgroup might be overcome only by an intensified treatment strategy which, however, might not be applicable to elderly patients [6, 7].

In contrast to our results, a recently reported Canadian analysis did not find any influence on PFS or OS when retrospectively stratifying early-stage patients according to GHSG or NCIC-CTG criteria. However, the number of patients was smaller, treatment was heterogeneous, and intensity was adapted to risk groups, making the assessment of risk factors difficult [13]. Another analysis included early-stage patients treated with abbreviated Stanford V chemotherapy followed by RT. Although only 101 patients were analyzed, differences in freedom from progression rates between favorable and unfavorable patients were significant when applying GHSG criteria to this cohort. However, other staging systems failed to discriminate risk groups in this smaller analysis [14].

What are the clinical implications of our study? Discrimination of risk groups still has significant impact on PFS and OS in early-stage HL in the modern treatment era. Risk-adapted treatment strategies are being pursued, trying to overcome the role of initial staging and prognostic factors. In the GHSG HD10 trial for early favorable stages, treatment has been reduced to two cycles of ABVD followed by 20 Gy IFRT without loss of efficacy, [1] leading to 5-year PFS of 92.2% (95% CI 87.8% to 95.0%) for patients <60 years (personal communication GHSG) [1]. For the early unfavorable risk group, the intensified 2 + 2 regimen used in the GHSG HD14 trial (two cycles of escalated BEACOPP followed by two cycles of ABVD) resulted in a 5-year PFS of 95.4% (95% CI 93.7% to 97.1%) with an advantage of 6.2% when compared with four cycles of ABVD [15]. Further analyses of low-risk and high-risk subgroups within HD14 suggested that patients with a large mediastinal mass or elevated ESR particularly benefit from the more intensive chemotherapy. Thus, these tailored approaches seem to overcome the older prognostic factors with respect to PFS. Another analysis of non-randomized registry data even suggested to treat early unfavorable stage patients with strategies for advanced stage HL, when presenting with B-symptoms or a large mediastinal mass [16]. However, treatment intensification leading to better tumor control does not always translate into better OS [15]. Furthermore, all risk factors are lacking specificity, which may lead to over treatment of some patients when trying to achieve higher overall cure rates for all. Given that 5-year OS rates are already >95% in early-stage patients, intensification strategies are not generally accepted and used worldwide.

Unfortunately, a new prognostic model with equal sensitivity and higher specificity could not be built from the routinely assessed risk factors including clinical patient- and tumor-related factors. In addition, the three staging systems pose clinical difficulties, associated, e.g. with the differing definition of the risk factor ‘number of nodal sites’. Real improvement may only be achieved by identification and inclusion of different factors (e.g. new biomarkers) with much higher specificity and by construction of a simpler universal model to prevent the current confusion and to avoid potential errors in applying trial results.

Another approach to keep the burden of intensive therapy to those in need is response-adapted therapy based on positron emission tomography, which is currently being investigated in several large randomized clinical trials for early-stage HL (EORTC H10, GHSG HD16, GHSG HD17, UK NCRI RAPID). However, these studies are not designed to challenge the established upfront risk stratification and will therefore not affect the results presented in this analysis.

In conclusion, differentiating between a favorable and an unfavorable risk group has significant impact on PFS and OS in early-stage HL patients in the modern treatment era. Further advances in risk- and response-adapted treatment strategies as well as identification of new factors with higher specificity may help to tailor treatment intensity to the individual needs of our patients.

acknowledgements

We are indebted to the patients who participated in the GHSG trials HD10 and HD11, providing the basis for this analysis. We thank all colleagues from participating centers and all members of the GHSG trial center and the Bundesministerium für Bildung und Forschung for their administrative and technical support. The work was presented at the 2012 Annual Meeting of the American Society of Hematology, December 8–11, 2012, Atlanta, GA, USA.

funding

Financial Support from the Deutsche Krebshilfe including funding of the HD10 and HD11 trial (grant number 70-2368-Di 8).
Despite improvement with intensive multi-agent chemotherapy, 2-year progression-free survival (PFS) rates for adults with high-risk Burkitt’s lymphoma (BL) remains <55%.

---

**A multicenter phase II study incorporating high-dose rituximab and liposomal doxorubicin into the CODOX-M/IVAC regimen for untreated Burkitt’s lymphoma**


1Division of Hematology/Oncology, Tufts Medical Center, Boston; 2Division of Hematology/Oncology, Washington University School of Medicine, St. Louis; 3School of Pharmacy, University of Wisconsin Carbone Cancer Center, Madison; 4Medicine, Division of Hematology and Oncology, Advocate Lutheran General Hospital, Park Ridge; 5Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago; 6Division of Hematology/Oncology, University of Rochester Medical Center, Rochester; 7Hematology and Oncology, University Hospitals Case Medical Center, Cleveland; 8Division of Hematology and Medical Oncology, Rush University Medical Center, Chicago; 9Division of Hematology/Oncology and the Robert H. Lurie Comprehensive Cancer Center, Northwestern University Feinberg School of Medicine, Chicago, USA

Received 8 July 2013; revised 28 August 2013; accepted 29 August 2013

**Background:** Despite improvement with intensive multi-agent chemotherapy, 2-year progression-free survival (PFS) rates for adults with high-risk Burkitt’s lymphoma (BL) remains <55%.

---

*Correspondence to: Dr. Andrew Evens, Division of Hematology/Oncology, Tufts Medical Center, 800 Washington Street, Boston, MA 02111, USA. Tel: +1-617-636-8077; Fax 1-617-636-7060; E-mail: aevens@tuftsmedicalcenter.org

© The Author 2013. Published by Oxford University Press on behalf of the European Society for Medical Oncology. All rights reserved. For permissions, please email: journals.permissions@oup.com.