Bone Marrow Transplantation (2020) 55:2224–2233
https://doi.org/10.1038/s41409-020-0934-1

ARTICLE

Multidimensional geriatric assessment for elderly hematological patients (≥60 years) submitted to allogeneic stem cell transplantation. A French–Italian 10-year experience on 228 patients

Nicola Polverelli[1] · Paolo Tura[1,2] · Giorgia Battipaglia[2] · Michele Malagola[1] · Simona Bernardi[1,3] · Lisa Gandolfi[1] · Tatiana Zollner[1] · Camilla Zanaglio[1,3] · Mirko Farina[1] · Enrico Morello[1] · Alessandro Turra[1] · Mohamad Mohty[2] · Domenico Russo[1]

Received: 7 February 2020 / Revised: 27 April 2020 / Accepted: 28 April 2020 / Published online: 12 May 2020

Abstract

Nowadays, the evaluation of elderly patients' eligibility for allogeneic stem cell transplantation (allo-SCT) is crucial. We evaluated the feasibility and efficacy of a multidimensional geriatric assessment, the Fondazione Italiana Linfomi (FIL) score, on a cohort of 228 patients older than 60 years submitted to allo-SCT in Italy and France from 2008 to 2018. Based on FIL score, available in 215 patients, 125 (58%) patients were classified as "fit" and 90 as "unfit/frail." The hematopoietic cell transplantation-specific comorbidity index (HCT-CI) was measured in 222 patients (97%); 71 (32%) patients had HCT-CI 0, 75 (34%) patients scored 1–2, and 76 (34%) ≥3. A total of 121 (53%) patients died after a median follow-up of 36 months. FIL score was found to highly predict survival, due to an excess of NRM in unfit/frail group, and confirmed its independent prognostic role on OS (HR: 0.37; 95% CI: 0.25–0.55; $p < 0.0001$). On the contrary, the HCI-CI failed in allo-SCT outcome prediction (HR: 1.06; 95% CI: 0.96–1.16; $p = 0.27$). In summary, a comprehensive geriatric assessment with FIL score seems to add significant prognostic information in elderly patients submitted to allo-SCT. The pretransplant adoption of this easy-to-use tool could help the patients' selection and management.

Introduction

Thanks to the introduction of reduced intensity conditioning (RIC) regimens, the improvement of HLA-typing, donor selection, and supportive care (anti-infective and graft-vs-host disease (GVHD) prophylaxis and therapy), advanced age does not represent an absolute limit for allogeneic stem cell transplantation (allo-SCT) [1, 2]. However, non-relapse mortality (NRM) is still a major issue and limits the use of this curative approach to only a minority of fit elderly patients. Selection of eligible patients older than 60 years is therefore crucial to maximize the benefit of allo-SCT.

The easier way to measure patient fitness is to evaluate patients' performance. Karnofsky Performance Status (KPS) was first introduced in the late 1940s and it was found to be an independent predictor of outcome in allo-SCT [3–5]. However, the majority of elderly allo-SCT patients present a preserved KPS, thus limiting the efficacy of this score in this setting. Later, the index of independence in activity of daily living (ADL) [6] and instrumental activities of daily living (IADL) [7] were adopted. While ADL impairments were rare, IADL compromission did impact allo-SCT outcome in elderly patients [8]. Thereafter, many other functionality assessments were studied (e.g., falls, exhaustion, timed up and go, 6-min walk test, etc.), confirming a

These authors contributed equally: Mohamad Mohty, Domenico Russo

Supplementary information The online version of this article (https://doi.org/10.1038/s41409-020-0934-1) contains supplementary material, which is available to authorized users.

✉ Nicola Polverelli
nicola.polverelli@unibs.it

[1] Unit of Blood Diseases and Stem Cells Transplantation, Department of Clinical and Experimental Sciences, University of Brescia, ASST Spedali Civili of Brescia, Brescia, Italy

[2] Department of Hematology and Cellular Therapy, "Saint Antoine Hospital" AP-HP, Paris, France

[3] Laboratorio CREA (Centro di Ricerca Emato-oncologica AIL), ASST Spedali Civili of Brescia, Brescia, Italy

high prevalence of limitations in elderly allo-SCT population [9].

Beside performance scores, comorbidities scores were largely applied. The hematopoietic cell transplantation-specific comorbidity index (HCT-CI) or Sorror score is currently the most used tool for transplant-risk assessment. Taking into account the number and severity of 15 medical conditions, HCT-CI stratifies patients into three risk categories (patients score 0, 1–2, or ≥3), with significantly different overall survival (OS) and risk of NRM [10]. In geriatrics, Cumulative Illness Rating Score for Geriatricians (CIRS-G) was shown to be a reliable measure of health state and physical impairment in elderly population [11]. This scale format provides for 14 independent categories grouped under body systems. Each domain is scored from 0 to 4 points, according to the degree of severity of organ involvement. The cumulative final score varies between 0 and 56. This score was also evaluated in allo-SCT elderly population, showing high prevalence of comorbidity, but no information on outcome is yet available [12].

By combining functionality and comorbidity scores, many comprehensive geriatric assessments (CGA) have been proposed [13]. Such scores demonstrated their efficacy in AML, myelodysplastic syndromes (MDS), multiple myeloma, and lymphomas prognostication [14].

In particular, a combined comorbidity and functionality score was developed on behalf of Fondazione Italiana Linfomi (FIL), the so-called FIL score [15]. This tool combines four parameters: (1) the CIRS-G [11], a geriatric-specific comorbidity score, (2) the ADL [6], (3) IADL [7], and (4) age > 80 years. The sum of these parameters stratifies patients into three categories: fit, unfit, and frail.

The four components of FIL score were chosen with the aim to evaluate three different area of elderly frailty: functionality (ADL and IADL), geriatric-specific comorbidity (CIRS-G), and biologic age (age > 80), making FIL score a quite comprehensive instrument.

This score was retrospectively evaluated in a monocentric study showing a good ability in identifying elderly diffuse large B-cell lymphoma (DLBCL) patients who could benefit from a curative approach [16]. Thereafter, it was prospectively validated on a cohort of 173 (>69 years) DLBCL patients, confirming its ability to predict OS [15]. Up-to-now, scant data are available on the use of combined functional and comorbidity scores in allo-SCT, and no study has evaluated FIL score in this setting. We therefore conducted this study with the aim of evaluating the role of the multidimensional FIL score in predicting allo-SCT outcome, in terms of NRM, relapse incidence (RI), and OS in a large cohort of elderly (>60 years) patients who received an allo-SCT in two European Transplant Centers.

Patients and methods

We analyzed 228 patients with hematological diseases, aged 60 years and older, consecutively submitted for allo-SCT at Spedali Civili of Brescia and Saint Antoine Hospital in Paris, from March 2008 to April 2018. The median follow-up of the entire cohort was 36 months (range: 1–118).

Clinical and laboratory data of patients submitted for allo-SCT were retrospectively collected from clinical charts or electronic databases available in Transplant Centers. Patients' characteristics are detailed in Table 1.

The median age of the cohort was 64 years (range: 60–76). Male sex was predominant ($n = 155$, 68%). AML was the most common indication for transplant (53%). Disease risk index (DRI) was calculated according to the CIBMTR assignment tool [17] on all the 227 patients with hematological neoplasms and was distributed as follows: low risk in 6 (3%), intermediate in 126 (56%), high in 76 (33%), and very high in 19 (8%) patients.

A matched unrelated donor (MUD) was used in 41%, sibling in 30%, alternative donors (haploidentical donor or cord blood) in 29% of cases. Peripheral blood-derived stem cells were most commonly employed (82%). Myeloablative conditioning (MAC) regimens were reserved for only a minority (20%) of patients. Busulfan-based conditioning regimens were the most employed (156/228, 68.4%), total body irradiation was given in 34 out of 228 cases (14.9%). In RIC setting, busulfan–fludarabine conditioning was predominant (65.4%), followed by TBI-based (17.0%), cyclophosphamide-based (14.8%), others (2.8%). The median dose of CD34+ infused cells was 5 (range: 0.01–12.92) $\times 10^6$/kg recipient body weight; CD3+ cell dose was 18 (range: 0.2–47.8) $\times 10^7$/kg.

The local transplant physicians prospectively evaluated KPS [3] in 217 (95%) and Sorror HCT-CI score in 222 (97%) patients; median time from HCT-CI evaluation to transplant was 19 days (range: 7–85). The FIL score was calculated for 215 (94%) patients, on the basis of the ADL and IADL measured within 3 months from transplantation in all patients undergoing allo-SCT at our Institutions after 2007 and CIRS-G retrospectively calculated from information recorded in hospital stored clinical charts, with the exclusion of hematology items (Supplementary Material).

Acute GVHD (aGVHD) was diagnosed and graded according to the Mount Sinai Acute GVHD International Consortium criteria [18] and chronic GVHD (cGVHD) according to the 2015 National Institutes of Health (NIH) criteria [19]. GVHD prophylaxis was given to all patients and consisted of: anti-thymocyte globulin (ATG) plus methotrexate (MTX) and calcineurin inhibitors (CNI) in 67 (29.4%), CNI plus MTX in 69 (30.3%), ATG plus CNI

Table 1 Clinical characteristics of patients submitted to allo-SCT.

	$N = 228$
Age in years, median (range)	64 (60–76)
Sex, male (%)	155 (67.9)
Diagnosis, *n* (%)	
AML	121 (53.0)
MDS	33 (14.5)
MPNs	18 (7.9)
Lymphoproliferative neoplasms	52 (22.8)
Other	4 (1.8)
DRI for 227 evaluable pts, *n* (%)	
Low	6 (2.6)
Intermediate	126 (55.5)
High	76 (33.5)
Very high	19 (8.4)
Donor gender, *n* (%)	
Female	83 (36.4)
Sex mismatch female to male	53 (23.2)
Donor, *n* (%)	
Sibling	69 (30.3)
MUD	94 (41.2)
Haploidentical	43 (18.9)
Cord blood	22 (9.6)
Stem cells source, *n* (%)	
Peripheral blood	188 (82.5)
Bone marrow	18 (7.9)
UCB	22 (9.6)
CD34+ $\times 10^6$/kg, median (range)	5 (0.01–12.92)
CD3+ $\times 10^7$/kg, median (range)	18 (0.20–47.80)
Conditioning regimen, *n* (%)	
MAC	46 (20.2)
RIC	182 (79.8)
FIL score (available for 215 patients), *n* (%)	
Fit	125 (58.1)
Unfit/frail	90 (41.9)
Sorror HCT-CI, median (range) (available for 222 patients)	2 (0–9)
Karnofsky PS, median (range) (available for 217 patients)	90 (40–100)
Follow-up in months, median (range)	36 (1–118)

plus mycophenolate mofetil (MMF) in 62 (27.2%), and CNI, MMF plus posttransplant cyclophosphamide in 30 (13.2%) cases.

More patients submitted for allo-SCT in Italy belonged to DRI class high or very high (55 vs. 36%, $p = 0.007$), received transplant with conventional donors (84.1 vs. 66.0%, $p = 0.004$), conditioned by RIC regimen (92.8 vs. 74.2%, $p = 0.001$). Less FIL unfit/frail patients were encountered in the Italian cohort (17.5 vs. 52.0%, $p < 0.001$); on the contrary, Sorror classes distribution was comparable between the two centers ($p = 0.81$).

All patients were on standard anti-infection prophylaxis with acyclovir, fluconazole, and sulfamethoxazole–trimethoprim.

The collection of data was approved by the local ethics committee for each Center and all patients had previously given informed consent. The study was conducted according to the Declaration of Helsinki guidelines.

Statistical analysis

Categorical variables were compared using the chi-square test or Fisher's exact test, and continuous variables were compared using the Student's *t* test. OS was calculated using the Kaplan–Meier estimator; the log-rank test was used to detect subgroup differences in survival distributions and to compare factors impacting OS.

A Fine–Gray regression model for competing risks was used for aGVHD and cGVHD, NRM and RI calculations. Death without the event of interest was considered as a competing event. NRM comprised death from toxicity, infection, aGVHD, and cGVHD.

Univariate and multivariate analyses of survival were carried out using the Cox proportional-hazards regression model for OS and with Fine–Gray proportional-hazards regression for competing events for NRM. The following variables were included in the univariate analysis: number of CD34+ and CD3+ infused cells, HLA-mismatch (MM) number, stem cell source (PB, BM, or CB), conditioning regimen (MAC vs. RIC), use of ATG, donor sex, KPS (≥90), FIL score (fit vs. unfit/frail), Sorror HCT-CI (total sum, ≥1 vs. <1, ≥2 vs. <2, ≥3 vs. <3), DRI category, diagnosis of AL/MDS, age at allo-SCT, male sex.

All variables associated with OS ($p < 0.05$) in the univariate analysis were included in the multivariate analysis. KPS was found to be highly associated with FIL fitness, by the chi-square test, and was therefore excluded in the multivariate analysis.

The likelihood ratio was computed for FIL score (and its components), KPS and Sorror HCT-CI on the basis of proportional-hazards regression; larger likelihood ratio values (>10) indicate a large increase in model prediction, while values approaching 1 indicate no ability in predicting outcome [20]. Also, c-(concordance)-statistics were calculated, whereby results range from 0.5 (meaning no ability to discriminate) to 1 (indicating perfect discrimination) [21].

Statistical analyses were performed with EZR software (v1.40) [22].

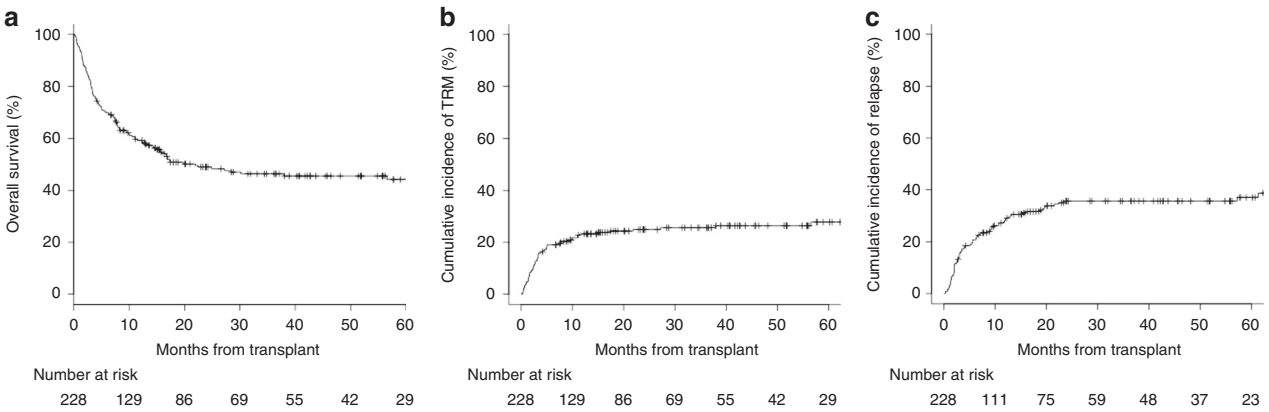

Fig. 1 Outcomes after allo-SCT. Overall survival (**a**), cumulative incidence of non-relapse mortality (NRM) (**b**), and relapse incidence (**c**) of study cohort (228 patients).

Results

Outcomes after allo-SCT

The median survival time of the entire cohort was 21.9 months (95% CI: 14.3–63.7), with an estimated OS of 69.9%, 59.2%, and 49.0% at 6, 12, and 24 months, respectively (Fig. 1a).

At last follow-up, 121 (53%) patients had died. The cause of death was consistent with NRM in 60 (49%) cases (31 infection, 13 toxicity, 9 aGvHD, and 7 cGvHD). Cumulative incidence of NRM was 19.0%, 23.3%, and 24.9% at 6, 12, and 24 months from transplant, respectively (Fig. 1b). Overall, 81 (36%) patients experienced disease relapse during follow-up, with an RI of 20.7%, 28.2%, and 35.7% at 6, 12, and 24 months, respectively (Fig. 1c). Disease relapse was the final cause of death in 61 (51%) cases.

aGVHD occurred in 118 (51.7%) patients, after a median time from allo-SCT of 26 days (range: 6–324), with a cumulative incidence of 44% at 3 months. aGVHD grading was distributed as follows: grade I in 66/228 (28.9%) patients, grade II in 33 (14.5%), grades III–IV in 19 (8.3%).

cGVHD was recorded in 72 of the 185 (38.9%) patients surviving more than 100 days after allo-SCT. Cumulative incidence of cGVHD was 14.1%, 32.7%, 37.9% at 6, 12, and 24 months, respectively, for an incidence rate of 24.3 per 100 patients/year. According to NIH criteria, 19/185 (10.3%) patients had mild, 35 (18.9%) moderate, and 18 (9.7%) severe cGVHD.

Pretransplant fitness assessment

Overall, HCT-CI evaluation before transplant was available for 222/228 (97%) patients. HCT-CI was 0 in 71/222 (31.9%) patients, while the majority of patients had HCT-CI impairment, score 1–2 in 75 (33.8%), and ≥3 in 76 (34.3%) cases. Patients who scored HCT-CI ≥ 3 were less likely to have received MAC (11.8 vs. 25.3%; $p = 0.023$), while other features did not differ between the HCT-CI groups. Among the comorbidities, pulmonary conditions were the most frequently encountered (46.4%), followed by heart valve disease (39.7%), solid cancer (23.2%), diabetes (19.9%), arrhythmia (16.6%), cardiac (7.3%), cerebrovascular disease (7.3%), rheumatologic (7.3%), hepatic impairment (7.3%), peptic ulcer (7.3%), active infection (4.6%), and psychiatric disorders (4.6%).

The FIL score was measured in 215/228 (94%) patients and classified patients as "fit" in 125/215 (58.1%), "unfit" in 19 (8.8%), and "frail" in 71 (33.1%). Causes of FIL unfitness or frailty were related to: IADL impairment in 63 (70.0%), CIRS-G score in 20 (22.2%), ADL in 4 (4.5%), and IADL plus ADL deficits in 3 (3.3%) patients, respectively. No patient was aged > 80 years at transplant. According to the FIL score, fit patients were more frequently diagnosed with AML (64.0 vs. 37.8%, $p < 0.001$), belonged to a lower DRI class (67.2 vs. 46.1%, $p < 0.001$). More "unfit/frail" individuals had HCT-CI impairment ($p = 0.022$) and KPS ≤ 80 ($p < 0.001$) compared with "fit" subjects (Table 2). Figure 2 represents the distribution of KPS according to FIL categories.

Predictive factors for overall survival and non-relapse mortality

Interestingly, HCT-CI was not able to discriminate between higher and lower risk patients in our cohort of elderly patients. In fact, patients with HCT-CI ≥ 3 had a 6-, 12-, and 24-month OS of 71.6%, 61.1%, and 51.4% compared with 70.7%, 59.1%, and 46.5% of patients with a lower score, respectively ($p = 0.539$) (Fig. 3a). NRM and RI were also comparable in patients with HCT-CI ≥ 3 and <3; NRM was 17.3%, 23.0%, and 24.7% vs. 19.9%, 21.4%, and

Table 2 Characteristics of patients according to FIL score; study cohort was divided into two categories: fit and unfit/frail patients.

	Fit ($N = 125$)	Unfit/frail ($N = 90$)	p value
Age in years, median (range)	64 (60–72)	64 (60–76)	0.509
Sex, male (%)	85 (68.0)	60 (66.7)	0.883
Diagnosis, *n* (%)			
AML	80 (64.0)	34 (37.8)	<0.001
MDS	16 (12.8)	16 (17.8)	
MPNs	4 (3.2)	14 (15.6)	
Lymphoproliferative neoplasms	23 (18.4)	24 (26.7)	
Other	2 (1.6)	2 (2.2)	
DRI, number on 214 evaluable (%)			
Low	2 (1.6)	4 (4.5)	0.002
Intermediate	82 (65.6)	37 (41.6)	
High	35 (28.0)	36 (40.4)	
Very high	6 (4.8)	12 (13.5)	
Donor gender, *n* (%)			
Female	43 (34.4)	37 (41.1)	0.315
Sex mismatch female to male	27 (31.8)	24 (40.0)	0.306
Donor, *n* (%)			
Sibling	38 (30.4)	25 (27.8)	0.102
MUD	58 (46.4)	32 (35.6)	
Cord blood	11 (8.8)	8 (8.9)	
Haploidentical	18 (14.4)	25 (27.8)	
Stem cells source, *n* (%)			
Peripheral blood	101 (80.8)	77 (85.5)	0.447
Bone marrow	13 (10.4)	5 (5.5)	
UCB	11 (8.8)	8 (9.0)	
CD3+ $\times 10^7$/kg, median (range)	5.0 (0.01–12.9)	5.1 (0.1–10.2)	0.225
CD34+ $\times 10^6$/kg, median (range)	17.9 (0.2–47.8)	19.2 (1.7–41.0)	0.767
Conditioning regimen, *n* (%)			
MAC	30 (24.0)	16 (17.8)	0.314
RIC	95 (76.0)	74 (82.2)	
Sorror HCT-CI, median (range)	1 (0–9)	2 (0–9)	0.046
HCT 0, *n* (%)	49 (39.2)	20 (22.2)	0.022
HCT 1–2, *n* (%)	34 (27.2)	36 (40.0)	
HCT ≥ 3, *n* (%)	42 (33.6)	34 (37.8)	
Karnofsky PS, median (range)	100 (60–100)	90 (40–100)	<0.001
KPS > 80, *n* on 205 evaluable (%)	100 (82.0)	42 (50.6)	<0.001
Follow-up months, median (range)	37.1 (1.0–118.0)	29.9 (1.0–107.7)	0.160

23.2% ($p = 0.905$), respectively, and RI was 18.6%, 27.2%, and 35.0% vs. 25.1%, 30.9%, and 36.5% ($p = 0.709$) in the HCT-CI ≥ 3 and <3 groups, respectively (Fig. 3b, c).

Unlike HCT-CI, comorbidity evaluation by CIRS-G score significantly stratified patients with different survival, despite a low number of CIRS-G impaired patients ($n = 20$). CIRS-G fit patients were projected to a 24-month OS of 54.5 vs. 20.0% of unfit/frail patients ($p < 0.0001$). Twenty-four-month estimated NRM and RI were 21.9% vs. 40.0% ($p = 0.06$) and 32.7% vs. 50.0% ($p = 0.09$) in fit and unfit/frail CIRS-G patients, respectively (Fig. 3d, e). Also, ADL/IADL-impaired patients experienced worse survival. Twenty-four-month OS, NRM and RI were 58.8% vs. 35.4% ($p = 0.0002$), 16.7% vs. 37.6% ($p = 0.002$), and 35.5% vs. 32.2% ($p = 0.95$) in ADL/IADL fit and unfit/frail patients, respectively.

Combining CIRS-G, IADL and ADL, FIL score performed even better. FIL "fit" patients experienced an OS of 86.2%, 74.9%, and 65.6% compared with 54.4%, 44.3%, and 31.6% in FIL "unfit/frail" patients at 6, 12, and

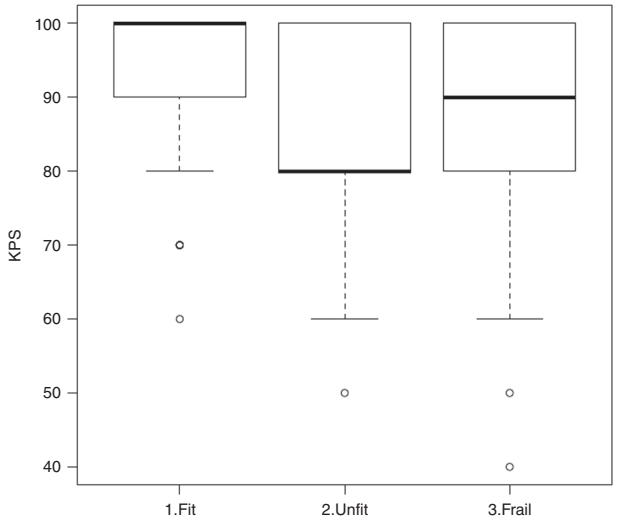

Fig. 2 Karnofsky Performance Status distribution among different FIL categories (overall *p* value < 0.001). Fit vs. unfit ($p = 0.005$), fit vs. frail ($p < 0.001$), unfit vs. frail ($p = 0.83$). In box and whiskers plot, center line represents the median values; the boxes' limits represent the lower and the higher quartile; whiskers define the 10–90 percentiles; and points represent the outliers.

24 months, respectively ($p < 0.0001$) (Fig. 4a), while no differences in transplant outcome were observed between unfit and frail patients ($p = 0.81$). This finding was related to an excess of NRM; in fact, transplant-related deaths were significantly higher in the "unfit/frail" in comparison with the "fit" group: NRM at 6, 12, and 24 months was 7.3%, 10.8%, and 12.9% in the "fit" group and 31.1%, 36.8%, and 38.1% in the "unfit/frail" group ($p < 0.0001$), respectively. Interestingly, despite the higher percentage of DRI high/very high patients among unfit/frail patients, RI was similar between "fit" and "unfit/frail" groups, with a 6-, 12-, and 24-month RI of 16.2%, 23.2%, and 32.7% vs. 24.4%, 31.2%, and 32.3%, respectively ($p = 0.39$) (Fig. 4b, c).

By univariate analysis, higher DRI category (HR: 1.51; 95% CI: 1.05–2.16; $p = 0.02$), KPS ≥ 90 (HR: 0.53; 95% CI: 0.36–0.78; $p = 0.0013$), and FIL "fit" status (HR: 0.35; 95% CI: 0.24–0.52; $p < 0.0001$) were associated with survival. On the contrary, when evaluated as either a continuous variable (HR: 1.06; 95% CI: 0.96–1.16; $p = 0.27$) or as a categorical variable, HCT-CI was not associated with OS: HCT-CI ≥ 1 or <1 (HR: 1.12; 95% CI: 0.75–1.66;

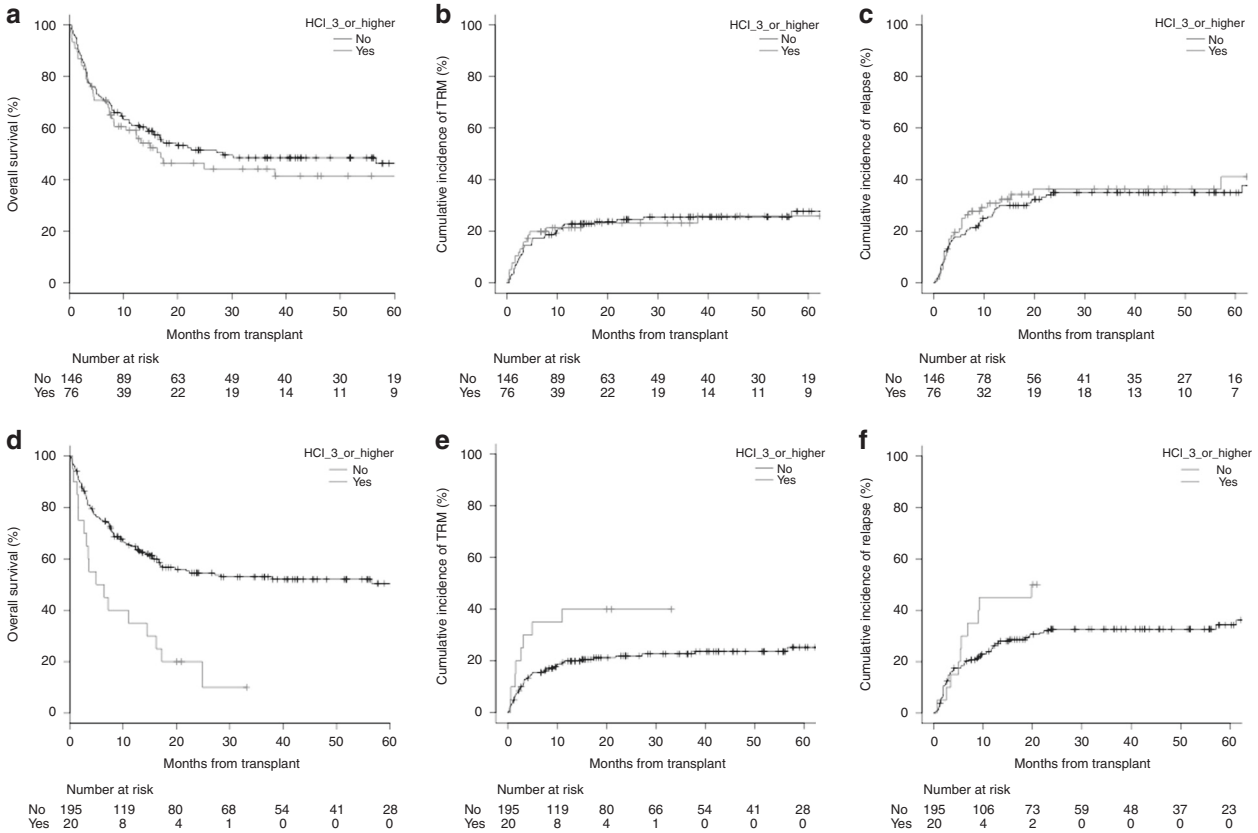

Fig. 3 Outcomes after allo-SCT according to HCT-CI and CIRS-G score. Overall survival (**a**), cumulative incidence of non-relapse mortality (NRM) (**b**), and relapse incidence (**c**) according to Sorror HCT-CI. The black line represents patients with score ≥ 3, red line patients with score < 3. Overall survival (**d**), cumulative incidence of non-relapse mortality (NRM) (**e**), and relapse incidence (**f**) according to CIRS-G unfitness/frailty as assessed by FIL score. The black line represents fit patients, red line unfit/frail patients.

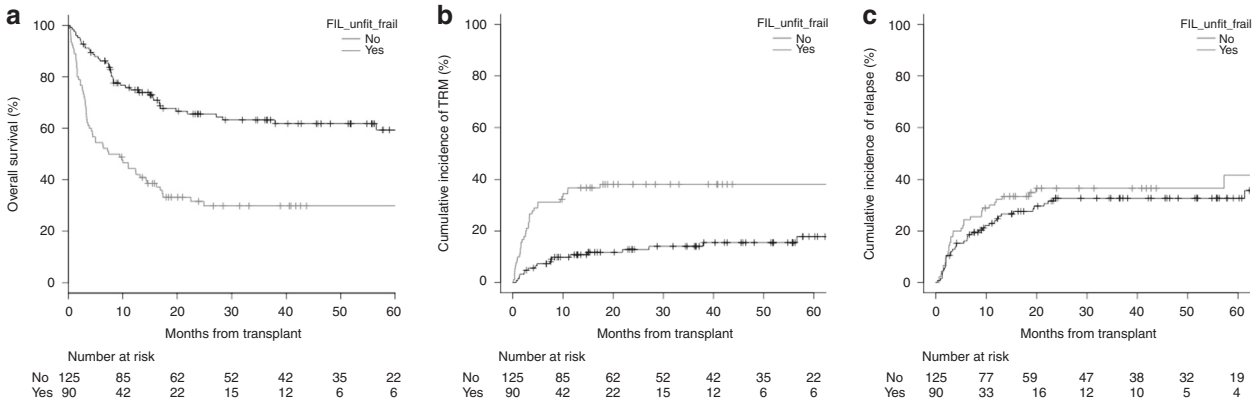

Fig. 4 Outcomes after allo-SCT according to FIL score. Overall survival (**a**), cumulative incidence of non-relapse mortality (NRM) (**b**), and relapse incidence (**c**) according to FIL score. The black line represents "fit" patients, red line "unfit/frail" patients.

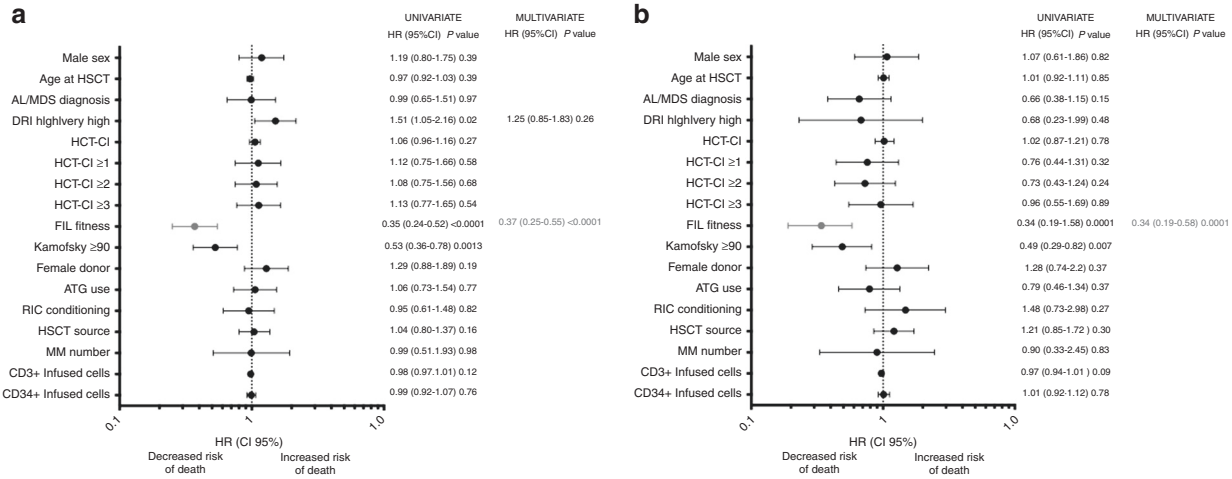

Fig. 5 Factors associated with OS and NRM. Uni and multivariable analysis of pretransplant factors associated with OS (**a**) and NRM (**b**); HSCT hematopoietic stem cell transplantation, AL acute leukemias, MDS myelodysplastic syndrome, DRI disease risk index, HCT-CI hematopoietic cell transplantation-comorbidity index, FIL Federazione Italiana Linfomi, ATG anti-thymocyte globulin, RIC reduced intensity conditioning, MM mismatch.

$p = 0.579$), HCT-CI ≥ 2 or <2 (HR: 1.08; 95% CI: 0.75–1.56; $p = 0.679$), HCI ≥ 3 or <3 (HR: 1.13; 95% CI: 0.77–1.65; $p = 0.540$).

Due to collinearity between FIL score and KPS, only FIL score and DRI underwent multivariate analysis, that confirmed the independent predictive role of the FIL score (fit vs. unfit/frail, HR: 0.37; 95% CI: 0.25–0.55; $p < 0.0001$) on OS (Fig. 5a).

Also on NRM, FIL score affirmed its prognostic role (HR: 0.34; 95% CI: 0.19–0.58; $p = 0.0001$) (Fig. 5b).

Comparison between FIL score, HCT-CI and KPS

Then, we performed a head-to-head comparison of performances of the FIL score and its components with HCT-CI and KPS.

Based on likelihood-ratio test and c-statistics, CIRS-G and ADL/IADL impairment had weak predictive role on OS (11.20 and 0.547 [standard error 0.016], 12.40 and 0.591 [standard error 0.024]) and NRM (4.10 and 0.546 [standard error 0.026], 10.90 and 0.622 [standard error 0.035]), respectively. By combining CIRS-G and IADL/ADL into FIL score, we observed a stronger association between unfitness/frailty and poorer survival due to an increased NRM, as evaluated by the likelihood-ratio test (28.45 and 19.37, respectively). C-statistics also confirmed the predictive role of FIL score on OS and NRM (0.637 [standard error 0.023] and 0.668 [standard error 0.032]). Conversely, Sorror HCT-CI had a significantly lower likelihood-ratio test value: OS and NRM (0.46 and 0.20, respectively); also, C-statistics confirmed the poor predictive role on OS and NRM (0.520 [standard error 0.020] and 0.511 [standard

error 0.040]). Also for KPS, likelihood-ratio test and c-statistics for OS and NRM were significantly lower (9.64 and 0.573 [standard error 0.023], 8.53 and 0.586 [standard error 0.034]) compared to FIL score.

Discussion

This retrospective study on a fairly large cohort of elderly patients (≥60 years) submitted to allo-SCT in two European Transplant Centers showed the potential ability of comprehensive FIL score in predicting allo-SCT NRM and OS. This score was, indeed, initially conceived for elderly patients diagnosed with DLBCL and this is the first study exploring its use in a cohort of elderly transplanted patients. Our experience supports the design of prospective trials to further validate the use of FIL score before allo-SCT for improving the evaluation of patients' eligibility to and management after allo-SCT.

In fact, taking into account the dramatic increase in the number of elderly hematological patients who are candidates to allo-SCT in the last few decades, the evaluation of patient fitness is becoming more important [23]. Also, the availability of novel and effective medical therapies raises the need for an effective transplant-risk assessment, thus tailoring the therapy according to patients' fitness and limiting the risk of NRM.

Up-to-now, HCT-CI is the most employed tool used for patient selection for allo-SCT [10]. This score was developed by Sorror et al. in a retrospective cohort of 1055 patients, submitted for allo-SCT mainly for malignant disease in Seattle [10]. Accordingly, HCT-CI assessment has been integrated in clinical practice and it is strongly recommended before transplantation by the Foundation for the Accreditation of Cellular Therapy and the European Joint Accreditation Committee (Accreditation Manual 7th edition). Nevertheless, it should be considered that the Sorror study cohort was transplanted in 1997–2003, the median age was 44.8 years, no patient received transplant from alternative donor and chronic myeloid leukemia was the primary indication for allo-SCT, making this population not representative of a contemporary cohort of transplant patients [24].

Later, the effectiveness of HCT-CI score was prospectively validated by two Italian and American studies [25, 26]. However, in CIBMTR experience, elderly population was a minority (22.8%) and only one-third of patients received RIC conditioning. Although HCT-CI confirmed good prediction of allo-SCT outcome, the authors suggested to further improve the ability of the comorbidity score with some components of geriatric assessments in subjects older than 60 [26]. In GITMO trail, HCT-CI showed a c-concordance statistics equal to 0.59 in RIC group, indicating a faint prediction in this setting.

Beside this solid literature, some research groups recently questioned the effectiveness of HCT-CI. Nakaya et al. reported no ability of the original HCT-CI to predict survival and NRM, especially in cord blood recipients [27]. An European cooperative work documented a weak predictive value of HCT-CI in a cohort of 812 patients receiving an RIC transplant for AML in first CR. HCT-CI also showed poor prediction on allo-SCT outcome in an American cohort of 203 allo-SCT patients older than 50 years [8].

Moreover, several studies have highlighted the complexity of evaluation of fitness in older cancer patients. In order to specifically capture peculiar problems of a senior population, a geriatric-specific comorbidity score was developed [11]. The CIRS-G to date represents the gold standard for rating the total burden of disease in the elderly. Also, many different multidisciplinary approaches by CGA have been reported in onco–hematology [13]. These experiences led the American Society of Clinical Oncology to strongly recommend the use of CGA in patients ≥ 65 years receiving chemotherapy [28].

In the allo-SCT setting, frailty, as evaluated by Fried's criteria, was shown to be associated with an eightfold higher risk of severe or life-threatening nervous system complications, 1.9 times higher risk of grades III–IV toxicities, and 3.1 times higher risk of overall mortality [29]. However, only few studies have evaluated the role of CGA in elderly allo-SCT patients in depth, with conflicting results. Muffly et al. first published a pilot study reporting a high prevalence of undocumented CGA impairments in older allo-SCT recipients, then they demonstrated an independent prognostic utility for IADL, slow walk speed, HCT-CI, and low mental health by short-form-36 mental component summary and elevated C-reactive protein in a prospective cohort of 203 allo-SCT patients ≥ 50 years old enrolled at the University of Chicago [8, 12]; IADL impairment and mild renal dysfunction were found to correlate with NRM and OS in a cohort of 457 elderly (≥60 years) patients undergoing allo-SCT at Memorial Sloan Kettering Cancer Center in New York [30]; in contrast a German study on 106 patients showed no differences in survival according to pretransplant CGA assessment [31].

By combining the geriatric-specific comorbidity assessment (CIRS-G), ADL, IADL, the FIL scores appeared to us to be a complete tool capable of capturing the complexity of the elderly allo-SCT population.

The main characteristics of our cohort were very similar to other recent elderly patient groups in terms of disease, conditioning regimen, GVHD prophylaxis, and outcomes [23, 30, 31].

The majority of patients presented with comorbidities, as demonstrated by HCT-CI impairment in 68% of cases, with around 50% of them being of pulmonary origin; nevertheless, comorbidities did not represent a contraindication to allo-SCT and resulted in a decrease in FIL score in only in a minority (22.2%) of cases. In our cohort, HCT-CI was not able to identify higher-risk patients; conversely, the FIL score was better able to stratify patients with different risks of NRM and OS, with unfit/frail patients identified as having a dismal outcome (32% compared with 66% at 2 years in fit patients). FIL score impairment was predominantly determined by IADL scale results (73.3% of cases), while basic functions as classified by the ADL were preserved in almost all patients.

Our study results highlight two main aspects: (1) the comorbidities themselves, as evaluated by HCT-CI, but not with CIRS-G score, could be less relevant for elderly patient selection for allo-SCT; (2) on the contrary, a comprehensive evaluation of patient functionality seems to be critical for allo-SCT success in the elderly.

In such a setting, a higher independence and a subject's ability to carry out complex actions may overcome physical limits. For instance, the ability to take care of oneself as well as compliance and adherence to daily home therapy and the early recognition of medical problems (infections, GVHD occurrence), leading to prompt access to specialist care, could be much more important than concurrent stable illness such as pulmonary function decrease, heart valve incontinence, controlled diabetes, etc. [32, 33].

We recognize the limits of our experience: the retrospective nature of the study, calculation of the CIRS-G score from medical records, and inter-rater variability could have affected the final results and the reliability of the data in terms of incorrect categorization, lack of information, and physician opinion. Also, higher-risk patients (by HCT-CI) may have received more careful medical surveillance, counterbalancing the unfavorable role of concomitant diseases. Nevertheless, the large cohort of patients, the use of easy to calculate scales (ADL, IADL, and CIRS-G), and survival endpoints are in the study's favor.

In the future, the evaluation of more precise indexes of patient functionality (e.g., muscle strength, etc.) could furtherly improve the prognostic value of FIL score [34].

In conclusion, our study suggests that FIL score is a potential model of interest for pretransplant evaluation of elderly candidates for allo-SCT, allowing an effective prognostication and helping transplant physicians in the selection and management of patients who can actually benefit from transplantation. Prospective larger trials and validation of our results in further cohorts are warranted to confirm these preliminary results before the adoption in clinical practice.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Passweg JR, Baldomero H, Basak GW, Chabannon C, Corbacioglu S, Duarte R, et al. The EBMT activity survey report 2017: a focus on allogeneic HCT for nonmalignant indications and on the use of non-HCT cell therapies. Bone Marrow Transplant. 2019;54:1575–85. https://doi.org/10.1038/s41409-019-0465-9.
2. Ringden O, Boumendil A, Labopin M, Canaani J, Beelen D, Ehninger G, et al. Outcome of allogeneic hematopoietic stem cell transplantation in patients age >69 years with acute myelogenous leukemia: on behalf of the acute leukemia working party of the european society for blood and marrow transplantation. Biol Blood Marrow Transplant. 2019;25:1975–83. https://doi.org/10.1016/j.bbmt.2019.05.037.
3. Yates JW, Chalmer B, McKegney FP. Evaluation of patients with advanced cancer using the Karnofsky Performance Status. Cancer. 1980;45:2220–4. 10.1002/1097-0142(19800415)45:8<2220::aid-cncr2820450835>3.0.co;2-q.
4. Sorror M, Storer B, Sandmaier BM, Maloney DG, Chauncey TR, Langston A, et al. Hematopoietic cell transplantation-comorbidity index and Karnofsky Performance Status are independent predictors of morbidity and mortality after allogeneic nonmyeloablative hematopoietic cell transplantation. Cancer. 2008;112:1992–2001. https://doi.org/10.1002/cncr.23375.
5. Artz AS, Pollyea DA, Kocherginsky M, Stock W, Rich E, Odenike O, et al. Performance status and comorbidity predict transplant-related mortality after allogeneic hematopoietic cell transplantation. Biol Blood Marrow Transplant. 2006;12:954–64. https://doi.org/10.1016/j.bbmt.2006.05.015.
6. Katz S, Akpom CA. A measure of primary sociobiological functions. Int J Health Serv. 1976;6:493–508. https://doi.org/10.2190/UURL-2RYU-WRYD-EY3K.
7. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist. 1969;9:179–86.
8. Muffly LS, Kocherginsky M, Stock W, Chu Q, Bishop MR, Godley LA, et al. Geriatric assessment to predict survival in older allogeneic hematopoietic cell transplantation recipients. Haematologica. 2014;99:1373–9. https://doi.org/10.3324/haematol.2014.103655.
9. Artz AS. Biologic vs physiologic age in the transplant candidate. Hematol Am Soc Hematol Educ Program. 2016;2016:99–105. https://doi.org/10.1182/asheducation-2016.1.99.
10. Sorror ML, Maris MB, Storb R, Baron F, Sandmaier BM, Maloney DG, et al. Hematopoietic cell transplantation (HCT)-specific comorbidity index: a new tool for risk assessment before allogeneic HCT. Blood. 2005;106:2912–9. https://doi.org/10.1182/blood-2005-05-2004.
11. Miller MD, Paradis CF, Houck PR, Mazumdar S, Stack JA, Rifai AH, et al. Rating chronic medical illness burden in geropsychiatric practice and research: application of the Cumulative Illness Rating Scale. Psychiatry Res. 1992;41:237–48. https://doi.org/10.1016/0165-1781(92)90005-n.
12. Muffly LS, Boulukos M, Swanson K, Kocherginsky M, Cerro PD, Schroeder L, et al. Pilot study of comprehensive geriatric assessment (CGA) in allogeneic transplant: CGA captures a high

prevalence of vulnerabilities in older transplant recipients. Biol Blood Marrow Transplant. 2013;19:429–34. https://doi.org/10.1016/j.bbmt.2012.11.006.
13. Liu JJ, Extermann M. Comprehensive geriatric assessment and its clinical impact in oncology. Clin Geriatr Med. 2012;28:19–31. https://doi.org/10.1016/j.cger.2011.10.001.
14. Abel GA, Klepin HD. Frailty and the management of hematologic malignancies. Blood. 2018;131:515–24. https://doi.org/10.1182/blood-2017-09-746420.
15. Tucci A, Martelli M, Rigacci L, Riccomagno P, Cabras MG, Salvi F, et al. Comprehensive geriatric assessment is an essential tool to support treatment decisions in elderly patients with diffuse large B-cell lymphoma: a prospective multicenter evaluation in 173 patients by the Lymphoma Italian Foundation (FIL). Leuk Lymphoma. 2015;56:921–6. https://doi.org/10.3109/10428194.2014.953142.
16. Tucci A, Ferrari S, Bottelli C, Borlenghi E, Drera M, Rossi G. A comprehensive geriatric assessment is more effective than clinical judgment to identify elderly diffuse large cell lymphoma patients who benefit from aggressive therapy. Cancer. 2009;115:4547–53. https://doi.org/10.1002/cncr.24490.
17. Armand P, Kim HT, Logan BR, Wang Z, Alyea EP, Kalaycio ME, et al. Validation and refinement of the disease risk index for allogeneic stem cell transplantation. Blood. 2014;123:3664–71. https://doi.org/10.1182/blood-2014-01-552984.
18. Harris AC, Young R, Devine S, Hogan WJ, Ayuk F, Bunworasate U, et al. International, multicenter standardization of acute graft-versus-host disease clinical data collection: a report from the Mount Sinai Acute GVHD International Consortium. Biol Blood Marrow Transplant. 2016;22:4–10. https://doi.org/10.1016/j.bbmt.2015.09.001.
19. Jagasia MH, Greinix HT, Arora M, Williams KM, Wolff D, Cowen EW, et al. National Institutes of Health Consensus Development Project on criteria for clinical trials in chronic graft-versus-host disease: I. The 2014 Diagnosis and Staging Working Group report. Biol Blood Marrow Transplant. 2015;21:389–401 e381. https://doi.org/10.1016/j.bbmt.2014.12.001.
20. McGee S. Simplifying likelihood ratios. J Gen Intern Med. 2002;17:646–9. https://doi.org/10.1046/j.1525-1497.2002.10750.x.
21. Harrell FE Jr, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. Stat Med. 1984;3:143–52. https://doi.org/10.1002/sim.4780030207.
22. Kanda Y. Investigation of the freely available easy-to-use software 'EZR' for medical statistics. Bone Marrow Transplant. 2013;48:452–8. https://doi.org/10.1038/bmt.2012.244.
23. Muffly L, Pasquini MC, Martens M, Brazauskas R, Zhu X, Adekola K, et al. Increasing use of allogeneic hematopoietic cell transplantation in patients aged 70 years and older in the United States. Blood. 2017;130:1156–64. https://doi.org/10.1182/blood-2017-03-772368.
24. Passweg JR, Baldomero H, Bader P, Basak GW, Bonini C, Duarte R, et al. Is the use of unrelated donor transplantation leveling off in Europe? The 2016 European Society for Blood and Marrow Transplant activity survey report. Bone Marrow Transplant. 2018;53:1139–48. https://doi.org/10.1038/s41409-018-0153-1.
25. Raimondi R, Tosetto A, Oneto R, Cavazzina R, Rodeghiero F, Bacigalupo A, et al. Validation of the hematopoietic cell transplantation-specific comorbidity index: a prospective, multicenter GITMO study. Blood. 2012;120:1327–33. https://doi.org/10.1182/blood-2012-03-414573.
26. Sorror ML, Logan BR, Zhu X, Rizzo JD, Cooke KR, McCarthy PL, et al. Prospective validation of the predictive power of the hematopoietic cell transplantation comorbidity index: a Center for International Blood and Marrow Transplant Research Study. Biol Blood Marrow Transplant. 2015;21:1479–87. https://doi.org/10.1016/j.bbmt.2015.04.004.
27. Nakaya A, Mori T, Tanaka M, Tomita N, Nakaseko C, Yano S, et al. Does the hematopoietic cell transplantation specific comorbidity index (HCT-CI) predict transplantation outcomes? A prospective multicenter validation study of the Kanto Study Group for Cell Therapy. Biol Blood Marrow Transplant. 2014;20:1553–9. https://doi.org/10.1016/j.bbmt.2014.06.005.
28. Mohile SG, Dale W, Somerfield MR, Hurria A. Practical assessment and management of vulnerabilities in older patients receiving chemotherapy: ASCO Guideline for Geriatric Oncology Summary. J Oncol Pract. 2018;14:442–6. https://doi.org/10.1200/JOP.18.00180.
29. Pamukcuoglu M, Bhatia S, Weisdorf DJ, DeFor TE, Ustun C, Nayar M, et al. Hematopoietic cell transplant-related toxicities and mortality in frail recipients. Biol Blood Marrow Transplant. 2019;25:2454–60. https://doi.org/10.1016/j.bbmt.2019.07.030.
30. Lin RJ, Elko TA, Devlin SM, Shahrokni A, Jakubowski AA, Dahi PB, et al. Impact of geriatric vulnerabilities on allogeneic hematopoietic cell transplantation outcomes in older patients with hematologic malignancies. Bone Marrow Transplant. 2019. https://doi.org/10.1038/s41409-019-0654-6.
31. Deschler B, Ihorst G, Schnitzler S, Bertz H, Finke J. Geriatric assessment and quality of life in older patients considered for allogeneic hematopoietic cell transplantation: a prospective risk factor and serial assessment analysis. Bone Marrow Transplant. 2018;53:565–75. https://doi.org/10.1038/s41409-017-0021-4.
32. Turner A, Hochschild A, Burnett J, Zulfiqar A, Dyer CB. High prevalence of medication non-adherence in a sample of community-dwelling older adults with adult protective services-validated self-neglect. Drugs Aging. 2012;29:741–9. https://doi.org/10.1007/s40266-012-0007-2.
33. Mizokami F, Mase H, Kinoshita T, Kumagai T, Furuta K, Ito K. Adherence to medication regimens is an effective indicator of cognitive dysfunction in elderly individuals. Am J Alzheimers Dis Other Demen. 2016;31:132–6. https://doi.org/10.1177/1533317515598859.
34. Beaudart C, Zaaria M, Pasleau F, Reginster JY, Bruyere O. Health outcomes of sarcopenia: a systematic review and meta-analysis. PLoS ONE. 2017;12:e0169548. https://doi.org/10.1371/journal.pone.0169548.