Biol Blood Marrow Transplant 22 (2016) 2270–2275

Biology of Blood and Marrow Transplantation

journal homepage: www.bbmt.org

ASBMT American Society for Blood and Marrow Transplantation

Myeloablative versus Reduced-Intensity Conditioning in Patients with Myeloid Malignancies: A Propensity Score-Matched Analysis

Hassan Sibai, Umberto Falcone, Uday Deotare, Fotios V. Michelis, Jieun Uhm, Vikas Gupta, John Kuruvilla, Jeffrey H. Lipton, Matthew D. Seftel, Hans A. Messner, Dennis (Dong Hwan) Kim *

Allogeneic Blood and Marrow Transplant Program, Princess Margaret Cancer Centre, University Health Network, University of Toronto, Toronto, Ontario, Canada

Article history:
Received 7 July 2016
Accepted 31 August 2016

Key Words:
Propensity score matching
Allogeneic hematopoietic stem cell transplantation
Conditioning regimen
Acute myeloid leukemia
Myelodysplasia

A B S T R A C T

Reduced-intensity conditioning (RIC) has been shown to have similar overall survival (OS) but higher relapse rates compared with myeloablative (MAC) regimens in patients with myeloid malignancies undergoing allogeneic hematopoietic stem cell transplantation (allo-HSCT). Using propensity score matching (PSM) analysis, well-balanced pairs of different variables can be compared effectively. We retrospectively compared allo-HSCT recipients with acute myeloid leukemia or myelodysplasia receiving a RIC regimen (FBT200; fludarabine 30 mg/m^2/day for 4 days, busulfan 3.2 mg/kg/day for 2 days, and total body irradiation [TBI] 200 cGy) or MAC regimen (FBT400; fludarabine 50 mg/m^2/day for 4 days, busulfan 3.2 mg/kg/day for 4 days, and TBI 400 cGy). A total of 248 patients (121 in the RIC group and 127 in the MAC group) were included in the analysis. No statistically significant difference was observed in 2-year OS (RIC group, 45.2 ± 5.0%; MAC group, 51.7 ± 5.2%; $P = .541$), nonrelapse mortality (NRM; RIC group, 28.7 ± 2.8% MAC group, 34.7 ± 4.6%; $P = .368$), and acute graft-versus-host disease (GVHD) ($P = .171$) or chronic GVHD ($P = .605$) at 1 year. The cumulative incidence of relapse (CIR) at 2 years was statistically significantly different between the 2 groups, however (RIC, 26.1 ± 2.6%; MAC, 14.2 ± 3.5%; $P = .033$). When PSM was applied to the study population, 42 case-control pairs were evenly matched. PSM analysis confirmed no statistically significant difference in 2-year OS (RIC, 49.0 ± 9.1%; MAC, 54.9 ± 7.7%; $P = .718$), NRM (RIC, 22.2 ± 2.3%; MAC, 33.3 ± 2.8%; $P = .238$), or CIR (RIC, 25.7 ± 2.6%; MAC, 9.5 ± 1.1%; $P = .315$) in the PSM pairs. Our findings demonstrate that after applying PSM, FBT 200 RIC conditioning has comparable OS, NRM, and CIR to FBT 400 MAC conditioning before allo-HSCT.

INTRODUCTION

Allogeneic hematopoietic stem cell transplantation (allo-HSCT) is a curative treatment modality for patients with acute myeloid leukemia (AML) and myelodysplastic syndrome (MDS), as well as other hematologic disorders [1-4]. Although myeloablative conditioning (MAC) regimens, such as high doses of total body irradiation (TBI) plus chemotherapy [2] and combination chemotherapy regimens including busulfan and cyclophosphamide [1], are still widely used, the use of reduced-intensity conditioning (RIC) regimens has been increasing over the past 2 decades [5-7].

RIC regimens include reduced doses of chemotherapeutic agents and/or low-dose TBI. Adding low-dose TBI to RIC regimens has been shown to be a well-tolerated option, with antileukemic activity comparable to that of MAC regimens [8]. Whether RIC results in similar post-transplantation outcomes as MAC remains a matter of debate, however. Although RIC is suspected to increase the risk of relapse owing to the lower intensity of chemotherapy, it is associated with less conditioning-related toxicity and nonrelapse mortality (NRM) compared with MAC [5,9-12].

In addition to recipient age, other factors, including disease status at transplantation, donor type, cytogenetics, and hematopoietic cell transplantation comorbidity index (HCT-CI) [13], contribute to the outcome variables in both MAC and RIC allo-HSCT. In most transplantation studies, these pretransplantation factors are not entirely matched between

Financial disclosure: See Acknowledgments on page 2275.

H.S. and U.F. contributed equally to this work.

Previously presented in part at the American Society of Hematology's 56th Annual Meeting, San Francisco, California, December 6-9, 2014.

* Correspondence and reprint requests: Dennis (Dong Hwan) Kim, MD, PhD, Department of Medical Oncology and Hematology, Princess Margaret Cancer Center, Rm 5-126, 610 University Ave, Toronto, ON, Canada M5G 2M9.

E-mail address: dr.dennis.kim@uhn.ca (D. Kim).

http://dx.doi.org/10.1016/j.bbmt.2016.08.030
1083-8791/

comparable groups. The propensity score is the probability of treatment assignment conditional on observed baseline characteristics. Propensity score matching (PSM) analysis is performed to balance the variables affecting the choice of treatment among different treatment options [14,15].

We conducted a retrospective study including patients undergoing allo-HSCT for AML or MDS, with the aim of comparing MAC and RIC regimens, using PSM analysis to obtain well-matched pairs of patients. Our hypothesis is that by this methodology, the OS and relapse rates obtained with RIC regimans are comparable to those obtained with MAC regimens.

PATIENTS AND METHODS

Study Design and Eligibility Criteria

We retrospectively compared MAC and RIC regimens in patients who underwent allo-HSCT for AML or MDS at Princess Margaret Cancer Centre (PMCC) between January 2009 and December 2013. This study was approved by the Research Ethics Board of the University Health Network/PMCC. Only patients with AML in first complete remission (CR1) or second complete remission (CR2) and patients with MDS with <10% bone marrow blasts at the time of transplantation were included in the study. For AML, cytogenetic risk at diagnosis was characterized as favorable/intermediate versus adverse, as described by the Southwest Oncology Group/Eastern Cooperative Oncology Group [16]. In patients considered at good cytogenetic risk, allo-HSCT was offered for AML refractory to first-line treatment or in relapse after achieving CR2. For MDS cytogenetics, 2008 World Health Organization minimal MDS cytogenetic criteria were used [17].

Inclusion was restricted to patients receiving fludarabine/busulfan (FB) plus TBI with either RIC (FBT200: fludarabine 30 mg/m^2/day for 4 days, busulfan 3.2 mg/kg/day for 2 days, and TBI 200 cGy) or MAC (FBT400: fludarabine 50 mg/m^2/day for 4 days, busulfan 3.2 mg/kg/day for 4 days, and TBI 400 cGy). The decision to offer RIC was based primarily on patient age (≥60 years) and/or the presence of comorbidities [9,18-20]. Stem cell sources included matched-related donors (MRDs), matched-unrelated donors (MUDs), or mismatched unrelated donors (MMUDs). Donor peripheral blood stem cells were mobilized with granulocyte colony-stimulating factor. Haploidentical or cord blood allograft recipients were excluded from this study. GVHD prophylaxis in MRD allo-HSCT consisted of cyclosporine A combined with either mycophenolate mofetil (15 mg/kg by mouth or i.v. twice daily; dose rounded to the nearest multiple of 250) from day 0 for 30 days and then stopped without taper or methotrexate (15 mg/m^2 on allo-HSCT day +1 and 10 mg/kg on allo-HSCT days +3, +6, and +11). In patients receiving MUD allo-HSCT, T-cell depletion was performed using low-dose alemtuzumab or rabbit antithymocyte globulin (ATG) in combination with cyclosporine A.

Pretransplantation variables included in the PSM analysis were age, HCT-CI, remission status (CR1 vs CR2, and MDS with <10% blasts) at the time of transplantation, diagnosis (AML vs MDS), cytogenetic risk group (adverse risk versus favorable [relapsed] and intermediate risk), donor type (MRD vs MUD and MMUD), and period effect (year of transplantation). We adopted PSM analysis to adjust the risk factors affecting the choice of conditioning regimen by creating well-balanced pairs of RIC and MIC patients.

Definitions of Statistical Endpoints

OS duration was measured from the date of allo-HSCT until death from any cause. Alive patients were censored on the date of their last follow-up. The cumulative incidence of relapse (CIR) was calculated from the date of allo-HSCT until relapse. Relapse was defined as ≥5% blasts in a bone marrow aspirate or peripheral blood, or the development of extramedullary leukemia after allo-HSCT. NRM was calculated as death without evidence of disease relapse. Acute and chronic GVHD were diagnosed and graded using established criteria [21,22].

Statistical Analysis

Patient and disease characteristics were reported using descriptive statistics. Data were updated as of December 2013. The main outcome variables of interest included OS, CIR, and cumulative incidence of NRM. Cumulative incidence functions were used to estimate CIR and NRM in a competing risks setting. For relapse incidence, death from a nonrelapse cause was counted as a competing event, whereas relapse was taken as a competing event for NRM. The cumulative incidences of grade II-IV acute GVHD and chronic GVHD at 1 year were estimated taking into account death as a competing event. To adjust for any potential biases derived from imbalanced pretranplantation factors between the RIC and MAC groups, we adopted PSM analysis. In an initial step, the propensity score was calculated using a binary logistic regression model. The following 7 independent pretransplantation factors were included in the binary logistic regression model for calculation of propensity score: age, remission status, diagnosis, cytogenetic risk group (adverse-risk versus favorable/intermediate), donor type, HCT-CI, and period effect. Then each patient from the FBT200 group was matched and paired to another patient from the FBT400 group with the smallest differences in propensity score using an in-house script. For risk factor analysis, we included the foregoing 7 covariates with conditioning regimen (ie, MAC vs RIC) in the analysis for OS, CIR, and NRM. These analyses were performed in the entire cohort and separately in the PSM groups. After the univariable analysis, multivariable analysis was performed using the Cox proportional hazard regression model for OS and the Fine-Gray method for CIR and NRM. A stepwise selection procedure was applied using the criteria for variable selection, $P = .05$ for variable entry and $P = .1$ for variable removal. In addition to all of the variables remaining in the multivariable model, the variable of interest, conditioning regimen, was included in the final model.

Hazard ratios (HRs) and 95% confidence intervals (CIs) were estimated for the significant risk factors. Paired analysis was adopted throughout the PSM analysis for survival. For MVA in the PSM groups, only the statistically significant variables (plus conditioning regimen) on MVA for the whole cohort were considered. Statistical analyses were performed using EZR on R Commander version 1.11 [23].

RESULTS

Patients, Diseases, and Treatment Characteristics

Patient characteristics are summarized in Table 1. A total of 248 patients (134 males; 54%) underwent allo-HSCT using peripheral blood stem cells as a graft source. The median age at transplantation for the entire cohort was 54 years (range, 18-71 years). FBT400 was administered to 127 patients (51.2%), and FBT200 was administered to 121 patients (48.8%). Remission status was CR1 in 132 patients (53%) and CR2 in 56 patients (23%). Fifty-four patients (21.7%) had adverse risk cytogenetics, and 194 (78.3%) had good/intermediate risk cytogenetics. The stem cell source was MRD in 103 patients (41.5%), MUD in 115 patients (46.4%), and a MMUD in 30 patients (12.1%).

Demographic data and pretransplantion characteristics differed between the FBT200 and FBT400 groups. Older age ($P < .001$), higher HCT-CI score ($P = .001$), and more MRD ($P = .061$) were observed in the FBT200 group; however, no between-group differences were found in CR status at HSCT ($P = .316$), subtype of diagnosis (AML versus MDS; $P = .206$), or cytogenetic risk group ($P = .261$).

Using PSM analysis to overcome baseline imbalances, 42 case-control pairs (84 patients) were selected (Table 1). Pretransplantation variables became well balanced between the FBT200 and FBT400 groups after applying PSM. Median age at transplantation remained different between the 2 PSM groups (58 years in FBT200 versus 55 years in FBT400; $P = .009$).

Overall Outcomes

With a median follow-up of 18 months among survivors in the overall population (n = 248), the 2-year OS, NRM, and relapse incidence rate was 48.0 ± 3.6%, 34.6 ± 3.6% and 24.8 ± 3.5%, respectively. There was no difference in OS (45.2 ± 5.0% in RIC vs 51.7 ± 5.2% in MAC patients at 2 years; $P = .541$) or NRM (28.7 ± 2.8% in RIC vs 34.7 ± 4.6% in MAC patients at 2 years; $P = .368$) between FBT200 and FBT400 groups. The CIR at 2 years was higher in the RIC group (26.1 ± 2.6% in RIC vs 14.1 ± 1.4% in MAC; $P = .033$, Figure 1A). In the PSM subgroup of patients, no statistical difference was noted in 2-year OS (RIC 49.0 ± 9.1%, MAC 54.9 ± 7.7%; $P = .718$), NRM (RIC 22.2 ± 2.3%, MAC 33.3 ± 2.8%; $P = .238$), or CIR (RIC 25.7 ± 2.6%, MAC 9.5 ± 1.1%; $P = .315$, Figure 1B) between FBT200 and FBT400 groups.

Acute and Chronic GVHD According to Conditioning for Overall and PSM Pairs

The overall incidence of grade 3-4 acute GVHD was 19.4 ± 1.9%. When comparing between the two conditioning regimens, FBT400 and FBT200, there was no difference in rates