Statistical modelling of competing risks with incomplete data: with applications to allogeneic stem cell transplantation

Bonneville, E.F.

Chapter 4

Impact of comorbidities and body mass index on the outcomes of allogeneic hematopoietic cell transplantation in myelofibrosis: A study on behalf of the Chronic Malignancies Working Party of EBMT

Chapter based on: Polverelli, N., **Bonneville, E. F.**, de Wreede, L. C., et al. (2024) Impact of comorbidities and body mass index on the outcomes of allogeneic hematopoietic cell transplantation in myelofibrosis: A study on behalf of the Chronic Malignancies Working Party of EBMT. *American Journal of Hematology*, 99, 993—996. DOI: 10.1002/ajh.27262.

The process of selection of feasibility for transplant in myelofibrosis (MF) is determined by several factors such as age, disease stage, comorbidities, performance status, and donor availability (Acosta-Medina *et al.*, 2023). The Myelofibrosis Transplant Scoring System (MTSS) has emerged as a valuable tool for selecting suitable candidates for transplant in MF. By incorporating patient-, transplant-, and donor-specific variables, the MTSS has proven its effectiveness in stratifying patients at varying risks of non-relapse mortality (NRM) and overall survival (OS). More recently, a CIBMTR/EBMT score has been identified as an effective tool for MF transplant candidates' prognostication. However, it should be noted that the prognostic ability of both scores may be reduced by a lack of information on the presence of comorbidities and body mass index (BMI) prior to the transplant, which was not available in these analyses (Kröger *et al.*, 2024).

In order to assess the role of comorbidities and BMI in MF patients undergoing transplantation the Chronic Malignancies Working Party (CWMP) of the EBMT performed a retrospective study with the aim to provide more comprehensive and reliable data on the impact of these factors on transplant outcomes and to identify potential areas for improvement in current MF transplantation protocols. The policy of such study is consistent with that previously published (Polverelli *et al.*, 2021). Inclusion and exclusion criteria, definitions, and methodology are available in Appendix A.

Overall, 4086 patients were included in the final analysis. Patients' characteristics are available in the online supplementary materials. Out of 3157 patients with fully reported comorbidity data, 1701 patients (54%) had at least one comorbidity, with pulmonary conditions being the most prevalent (12.7% moderate and 6.8% severe), as documented also in other transplant scenarios (Polverelli *et al.*, 2020). Other comorbidities present in more than 5% of cases were cardiac disorders (8.6%), diabetes (5.7%) and prior-solid tumor (5.4%). An overview of all comorbidities is available in Figure 4.1.

Concerning the HCT-CI, 1701 (54%) patients had a low (0), 762 (24%) intermediate (1, 2) and 694 (22%) high-risk (=3) score, respectively. Table S2 reports the clinical characteristics stratified according to different HCT-CI classes (see also Table 5.1 from Chapter 5 for the unstratified characteristics). As expected, higher risk class did correlate with increased use of RIC regimens (high risk with 70% vs. 69% and 62% in intermediate, and low risk, respectively), decreased KPS (KPS <80 in 11% vs. 8.1 and 5%). Moreover, the proportion of the splenectomised patients was higher for the high-risk HCT-CI category (14% vs. 9% and 6% in high, intermediate, and low HCT-CI categories, respectively), leading to a lower prevalence of massive splenomegaly (=15 cm) (17% vs. 22% and 28%). Compared to previous cohorts in which the HCT-CI had been developed and subsequently validated (Sorror *et al.*, 2005), the prevalence of comorbidities was higher in our study. Overall, these differences underscore a significant shift in the characteristics of the transplant population over time, as transplantation is increasingly considered in older patients with comorbidities.

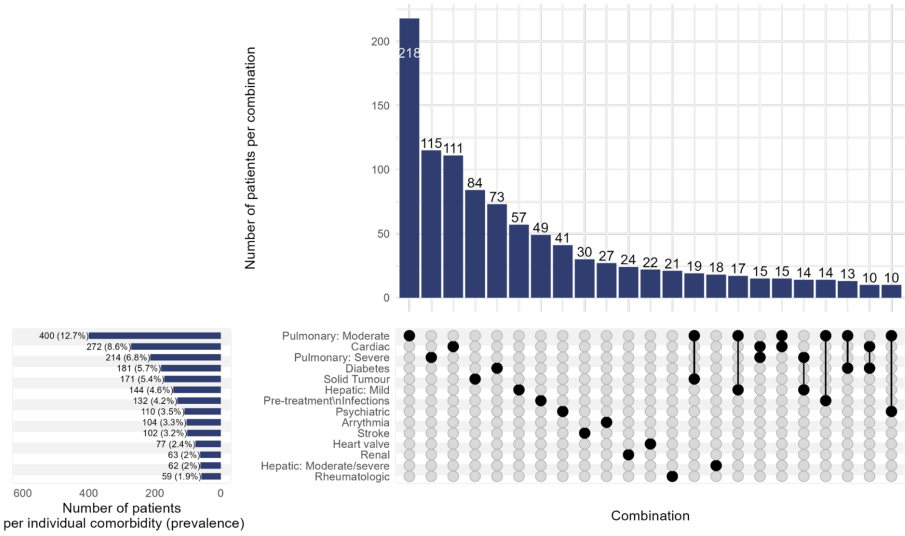

Figure 4.1: UpSet plot visualising the prevalences of individual comorbidities, together with most commonly occurring combinations of comorbidities in the cohort. The bar chart in the bottom-left hand corner of the plot shows the prevalences of individual comorbidities, while main bar chart shows the number of patients with a specific combination of comorbidities (defined by the black dots under it). For example, 400 (12.7%) patients had a moderate pulmonary comorbidity (either as only comorbidity, or in combination with other comorbidities), while 218 patients had a moderate pulmonary comorbidity as only comorbidity. 10 patients had both cardiac comorbidity and diabetes, but no other comorbidities. Combinations of comorbidities occurring in fewer than 10 patients are not shown. As a result, both IBD (n = 22, 0.7%) and Peptic Ulcer (n = 37, 1.17%) do not feature in this plot.

By univariable analysis, both NRM and OS were statistically associated with HCT-CI risk categories. The 5-year expected NRM was 27% (25%–30%), 33% (29%–36%), and 36% (32%–40%) in low, intermediate, and high-risk HCT-CI groups ($p < .001$), respectively. The 5-year estimated OS was 58% (55%–61%), 52% (47%–56%), and 46% (42%–51%) for the low, intermediate, and high HCT-CI scores, respectively ($p < .001$) (Figure S2). No statistical differences were observed in relapse incidence ($p = .22$), and incidence of grade 2-4 acute GvHD ($p = .056$) or chronic GvHD ($p = .46$) depending on the HCT-CI. Table S3 details the causes of NRM.

After adjusting for other variables well known to be associated with NRM and OS in MF, high-risk HCT-CI was strongly associated with both NRM (HR 1.32, 95% CI 1.12–1.55, $p < .001$) and OS (HR 1.27, 95% CI 1.11–1.46, $p < .001$), relative to patients with a low-risk HCT-CI (score of 0) (Figure S3). Also, splenectomy status did not appear to affect NRM in the context of high HCT-CI class ($p = .95$). Therefore, the presence of comorbidities continues to play a negative prognostic role on allo-HCT outcomes and should be integrated into the selection process for MF patients undergoing transplantation along with the existing MTSS and CIBMTR/EBMT tools.

A total of 2679 patients had information on BMI at time of transplant: 50 patients were classified as underweight (1.9%), 1318 as normal weight (49.2%), 964 as overweight (36%), and 347 as grade 1 to 3 obese (13%). Median BMI was 24.9 (range 12.1–46.1). The high prevalence of overweight and obese individuals suggested that patients with robust nutritional reserves were more often considered suitable for transplantation, while cachectic or sarcopenic patients may have had their transplant deferred due to a general tendency among physicians to avoid transplantation in such conditions, generally associated with worse transplant course. As compared to under-normal weight patients (1368, 51.1%), overweight/obese patients were more frequently males (69% vs. 57%), and had been more frequently exposed to ruxolitinib (40% vs. 34%). Continuous BMI was weakly correlated with all other comorbidities. Aside from a correlation of 0.11 with diabetes, correlations with any other comorbidity did not exceed ±0.07 (Table S4).

Despite differences in comorbidities and patient characteristics between different BMI classes, on univariable analysis, no significant differences were found across the BMI groups in terms of NRM ($p = .5$), OS ($p = .3$), grade II-IV acute GvHD ($p = .73$), or chronic GvHD ($p = .6$). By contrast, a modest difference was found regarding relapse incidence ($p = .031$). Furthermore, within a multivariable model that accounted for other variables known to correlate with NRM and OS in MF, including weight loss before allo-HCT, BMI was determined to have no significant impact on either NRM ($p = .59$) or OS ($p = .41$). Figure 4.2 shows the hazard ratio plots of BMI (relative to a reference BMI of 21.75) on OS and NRM respectively, and highlight the very limited impact of BMI on OS and NRM. Likelihood ratio tests also confirm the lack of non-linear effects on both OS ($p = .25$) and NRM ($p = .33$). These findings contrast with the original HCT-CI data, which identified a BMI >35 as a risk factor in both NRM and

OS after allo-HCT (Sorror *et al.*, 2005). In this context, it seems evident that overweight and obese MF patients should not be excluded from a potentially curative life-saving procedure. In MF, overweight or obesity can be associated with milder disease activity, resulting in better nutritional status and suggesting a greater likelihood of improved survival. On the other hand, even patients with lower BMIs can derive benefits from a transplant procedure.

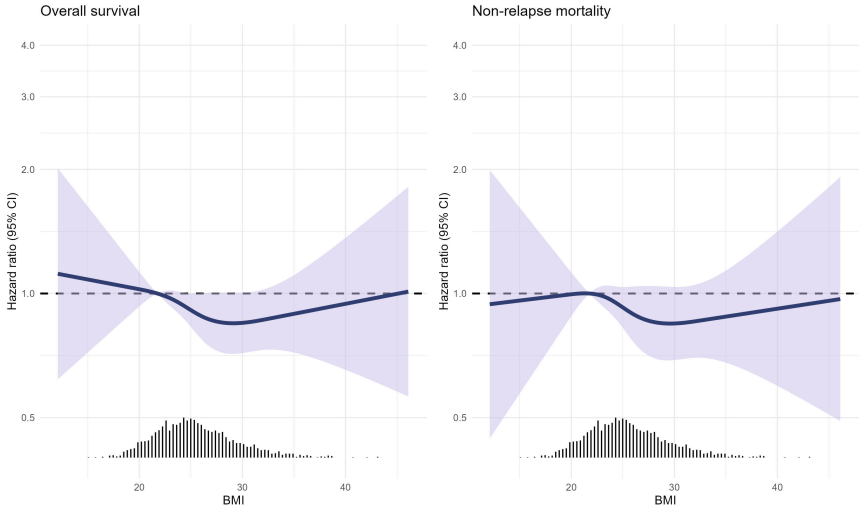

Figure 4.2: Non-linear effect of BMI on overall survival (left) and non-relapse mortality (right) as part of multivariable model adjusting for variables known to be associated with both outcomes in patients with MF. Models are based on n = 3982 patients with complete outcome information on OS and NRM, where covariates were multiply imputed using MICE (100 imputed datasets). Displayed are therefore the pooled coefficients. Hazard ratios should be interpreted relative to a reference BMI of 21.75 (mid-point of the 'normal' BMI category). The plots also show the (marginal) distribution of observed BMI values.

Importantly, evidence of weight loss >10% within 6 months prior to allo-HCT was significantly associated with higher risk of NRM (HR 1.19, 95% CI 1.01–1.39, p = .042), and a trend toward shortened OS (HR 1.19, 95% CI 0.96–1.46, p = 0.108). Therefore, it seems to be vital to consider transplantation not when the disease is already symptomatic with ongoing weight loss. In this case, the optimisation of nutritional status where possible should be considered.

The major strengths of this study rely on the novelty of this information in the largest sample of MF transplant patients, with comprehensive assessment of comorbidities

and BMI and significant follow-up. These numbers permit a comprehensive identification of factors associated with allo-HCT outcomes. Nevertheless, it's important to acknowledge some limitations in this study. First, despite adjusting for weight-loss prior to allo-HCT, the conclusions regarding the effect of BMI on mortality were likely also affected by selection effects that could not be modeled. In particular, patients with BMI >35 (i.e., for whom one would expect a clear negative impact on mortality) in this study may have been selected for allo-HCT based on other more favorable disease characteristics or lack of other comorbidities. Second, the study lacks detailed information regarding specific treatments administered for the management of comorbidities and weight loss before transplantation, which could potentially impact allo-HCT results; additionally, the definition of comorbidities lacks an understanding of their functional impact, and there is a current suggestion to include concomitant frailty assessment in cancer patients (Polverelli *et al.*, 2020). Third, the amount of missing data for both main variables and important adjustment factors was substantial. We chose not to exclude patients on the basis of unavailable comorbidity and BMI information (which may have rendered the cohort less representative), and instead made use of the observed information in the data in order to multiply impute the missing values, thereby potentially enhancing the robustness of the study results. However, the authors believe that despite these limitations, the significance of the topic, which pertains to an ever-growing number of MF patients over the years, outweighs these concerns.

In conclusion, this study, for the first time in a robust fashion, highlights the prognostic significance of HCT-CI in MF patients undergoing allo-HCT. Additionally, it suggests that BMI at the time of transplantation has a limited impact on transplant outcomes in this patient population. These findings enhance our understanding of risk factors and can guide clinical decision-making for MF patients considering allo-HCT. Nevertheless, future research should aim to validate these findings and explore the possibility of integrating comorbidity assessment alongside existing scoring systems and splenomegaly evaluation (Polverelli *et al.*, 2021, 2023) to develop effective tools for selecting MF patients as candidates for transplantation.

Supplementary materials

Supplementary tables and figures are available online at https://doi.org/10.1002/ajh.27262.

Appendix A: Methods

Inclusion criteria:

- All primary and secondary MF patients submitted to allo-HCT in chronic phase from any donor (except syngeneic or matched-related), conditioning regimen and stem cell source.
- Age greater or equal to 18 years-old.
- First transplant.
- Transplant performed between 2009 and 2019.

Exclusion criteria:

- Accelerated or Blastic phase MF.

Definitions

The comorbidities encountered at allo-HCT and their severity cut-offs were categorised according to the original study by Sorror *et al.* (2005). BMI at allo-HCT was measured by dividing a person's weight in kilograms by the square of his or her height in meters. According to WHO criteria, the following categories were defined: underweight (BMI <18.5), normal weight (18.5–24.9), overweight (25–29.9), class I obesity (30–34.9), class 2 obesity (35–39.9), and class 3 obesity ($\geq$ 40).

Statistical analysis

Pre-transplant patient characteristics were reported using the median and interquartile range (IQR) for continuous variables, and frequencies and proportions for categorical variables. Baseline was defined as time of first allo-HCT. Primary endpoints were overall survival (OS) and cumulative incidence of non-relapse mortality (NRM). OS was defined as time from allo-HCT to death, and NRM was defined as death without evidence of relapse or progression. Secondary endpoints included the cumulative incidence of relapse (including progression), as well as the cumulative incidences of acute (grades II-IV) and chronic Graft-versus-host disease (aGvHD and cGvHD, respectively). Median follow-up was determined using the reverse Kaplan–Meier method. OS was estimated using the Kaplan–Meier product limit estimation method, and differences in subgroups until 60 months were assessed by means of the Log-Rank test. Cumulative incidences of relapse and NRM were analysed together in a competing risks framework. Competing risks analyses were also separately applied to estimate the cumulative incidences of aGvHD and cGvHD, where death and second allo-HCT were considered as competing events. Subgroup differences in competing risks analyses were assessed using Gray's test.

HCT-CI was calculated using information for individual comorbidities reported by centers. The impact of HCT-CI (as low 0 risk, intermediate 1–2 risk or high ≥ 3 risk) on OS was assessed using a multivariable Cox proportional hazards model, and its impact of NRM was assessed using a cause-specific Cox model. Both models adjusted for the following variables: patient sex, patient age (per decade), donor type (identical sibling donor vs. other), Karnofsky performance score (KPS; ≥ 90, 80 and < 80), patient/donor cytomegalovirus match (patient/donor negative or other), ruxolitinib treatment prior to allo-HCT (given or not), conditioning (standard or reduced), and the individual components of the Dynamic international prognostic scoring system (Passamonti *et al.*, 2010) risk score other than patient age. These DIPSS components included continuous white blood cell count $x10^9$/L, hemoglobin g/L, peripheral blood blasts (per 5%), as well as weight loss (more than 10% in 6 months prior to allo-HCT) and presence of night sweats. Continuous components of the DIPSS were modelled linearly, after testing for non-linear effects using likelihood ratio tests (comparing the model with restricted cubic splines with one using linear effects).

A separate multivariable model was also fit (only for outcome NRM), evaluating the impact of individual comorbidities comprising the HCT-CI, instead of the HCT-CI itself. This was done in order to compare the implied weights to those assigned by the original HCT-CI. Furthermore, a subset analysis was considered to explore HCT-CI's prognostic value alongside MTSS-specific parameters. However, due to a small sample (around 100 patients) with complete MTSS data, this investigation was not feasible.

The impact of BMI on OS and NRM was also investigated by means of multivariable Cox models. We allowed for a potential flexible non-linear effect of BMI by using restricted cubic splines with two internal knots, placed at the 33rd and 67th percentile of BMI values. The reference BMI value was set to 21.75, the center point of the normal BMI category. This model adjusted for the same variables as those described above, together with a (continuous) partial HCT-CI score which omits the contribution of BMI. Importantly, we adjust for characteristics pertaining to disease risk, among which weight loss prior to allo-HCT (which is related to both BMI at allo-HCT and the risk of death). This is in view of trying to obtain a 'direct' effect of BMI on NRM, while mitigating potential bias from not adjusting for illness-related weight loss (which could make being overweight/obese appear protective against mortality—see Lennon *et al.*, 2016).

Missing values in covariates, which were assumed to be missing at random, were multiply imputed using multivariate imputations by chained equations (MICE) (van Buuren *et al.*, 1999), with 100 imputed datasets and 20 iterations. This approach efficiently makes use of the observed information in the data, and can offer estimates that are more precise and potentially less biased (Bonneville *et al.*, 2023). The imputation model for each partially observed covariate included remaining covariates in the multivariable models, the competing event indicator as a categorical variable, and the estimated marginal cause-specific cumulative hazards of both relapse and NRM

at an individual's event or censoring time (Bonneville *et al.*, 2022). Each comorbidity variable was imputed separately (rather than imputing the HCT-CI itself), and BMI was imputed as a continuous variable. The year of allo-HCT, and the interval between diagnosis and allo-HCT were used as continuous auxiliary variables in the imputation procedure. Default imputation methods were used: predictive mean matching for continuous variables, logistic regression for binary variables, proportional odds for ordered categorical, and multinomial regression for underordered multi-level categorical variables. Prior to imputing the missing covariates, missing relapse and aGvHD times (for patients with known event, but missing event time) were singly imputed by sampling the observed times from patients for whom the times were observed.

Outcomes in the analysis for aGvHD and death before aGvHD were artificially censored at 4 months after allo-HCT, while remaining outcomes were censored at 60 months after allo-HCT. All statistical tests were two-sided, at a significance level of 0.05. All analyses were performed in R version 4.2.1, using 'survival', 'cmprsk', 'prodlim', 'rms', and 'mice' packages.