CIBMTR®
CENTER FOR INTERNATIONAL BLOOD & MARROW TRANSPLANT RESEARCH

Methodology Employed for Annual Report on Hematopoietic Cell Transplant Center-Specific Survival Rates

Last Updated February 5, 2016

Introduction

The purpose of the annual report on transplant center-specific survival rates is to provide potential hematopoietic cell transplant (HCT) recipients, their families, and the general public with a comparison of survival rates among the centers in the C.W. Bill Young Cell Transplantation Program (the Program) network. Transplant centers may use these reports for quality improvement initiatives. Reporting center-specific survival rates is a requirement of the Stem Cell Therapeutic and Research Act of 2005 (re-authorized in 2010), and prior to that, the 1990 Transplant Amendments Act. Because centers vary considerably in the risk level of cases treated, a statistical model was developed to adjust for several risk factors known or suspected to influence outcome. The outcome reported is one-year overall survival, for recipients of allogeneic HCT in the United States only. No attempts are made to incorporate other outcomes, such as relapse or disease-free survival.

The first center-specific risk-adjusted comparisons were published in 1994[1] and yearly since then. The current iteration of the report prepared by the Center for International Blood and Marrow Transplant Research (CIBMTR) includes recipients of both unrelated and related donor transplants facilitated by the Program for a three-year time window. The methodology for this analysis has undergone various transformations over the years. The methodology in current use has been employed since 2005, thus allowing direct comparisons over the most recent nine reports. This method adjusts for risk using a censored data logistic regression model[2,3,4] that allows inclusion of recipients with incomplete one-year follow-up. Note that although the method has remained the same, the types of patients studied changed with the inclusion of related-donor transplants in the 2010 report, which may affect comparisons over time. A risk-adjusted one-year survival rate is calculated for each center, based on results of the censored data logistic regression.

Results are published on the Program website (http://bloodcell.transplant.hrsa.gov), and a version of this report, as approved by HRSA, is distributed to HCT centers. Previously, this report was

published in a printed version of the *National Marrow Donor Program* (*NMDP) Transplant Center Directory*. The print version of this Directory was eliminated in 2009, and this information now is available online only at www.bethematch.org/access. Raw numbers of transplants and surviving recipients are published for each center, stratified by diagnosis and age. Each center included in the report performed at least one unrelated or related donor transplant over the three-year window of time for analysis.

Methods

Recipients and data

The current analysis includes first unrelated or related donor transplants performed in a three-year time interval, with follow-up through one year after the last recipient was transplanted. The rolling three-year window of transplants for inclusion was adopted with the 2011 report, replacing a rolling five-year window used previously. This change was based on the recommendation of the 2010 Center-Specific Outcomes Analysis Forum[5], in order to represent more current transplant center outcomes. A minimum of one-year follow-up is required for all eligible cases. All U.S. transplant centers that performed at least one HCT in the time interval are considered for inclusion in the report, provided they had sufficient data with at least one year of follow-up available. Typically, about 170 U.S. transplant centers are included in the analysis, with about 22,000 first allogeneic transplants performed by domestic transplant centers in the Program network during this time.

Demographics of the included cases are provided in tables for recipients of unrelated donor transplants and recipients of related donor transplants, broken down by conditioning regimen according to myeloablative vs. reduced-intensity or non-myeloablative. Baseline and follow-up data used for the analysis are provided to the CIBMTR by the transplant centers at the time of transplant (baseline), and at 100 days, six months and annually post-transplant, using standardized forms. Race was self-reported by recipients or by the staff at the center. Occasionally, data is not available for significant characteristics for subjects as reported by the centers. If there were sufficient numbers of such subjects, they were included in the multivariate modeling as a distinct category of the covariate. However, when a category of a significant variable had too few subjects (generally less than 20 subjects) to fit the multivariate model, those subjects were imputed to the relevant highest frequency category within the variable.

Statistical Analysis

Rationale for a fixed effects censored data logistic regression model

One of the CIBMTR's goals for the transplant center-specific outcomes analysis is to calculate a fair and accurate predicted survival rate given a center's recipient case mix. To do this, a fixed effects censored regression model is used. The fixed-effect logistic regression model provides information about how the recipients actually treated in a particular center would have fared had they been transplanted at a "generic" transplant center within the Program. This model assumes *no center effect*. In other words, it assumes that recipients are dying at the same uniform rate across all Program transplant centers, after adjusting for covariates. The model also adequately accounts for recipients with incomplete follow-up at one year.

Every effort is made to update follow-up information on each recipient. Some recipients are indeed lost to follow-up, and their final survival status at one year is unknown. To address this

problem, the analysis only includes centers that demonstrated 90% completeness of follow-up, meaning that the one-year status was known for at least 90% of their transplanted recipients. However, there are still some recipients for whom survival status at one year is incomplete, although many recipients had follow-up done just prior to one year. If these recipients are excluded from the center-specific analysis, it may bias the survival estimates. A censored data version of logistic regression based on pseudo-values proposed by Andersen et al.,[2] Klein and Andersen,[3] and Klein et al.[4] addresses this issue. This method is a generalization of logistic regression that simplifies to logistic regression (on the one-year survival probabilities) when there is no censoring present. This regression technique is used to estimate the fixed effects and predict the recipients' survival probabilities based on their patient characteristics alone. These predicted survival probabilities are then used to construct confidence limits for a center's survival probability according to the characteristics of the patients transplanted at that center. The actual survival observed at that center can be compared to these intervals to assess the performance of the center. This method is described in more detail below.

Details of fixed effects censored data logistic regression and confidence limits

Modeling for the center specific outcomes analysis can be broken down into four steps, as outlined below.

I. Definition of pseudo-values

To compute the pseudo-value for recipient *i*, first compute the pooled sample Kaplan-Meier estimate of survival at one year based on the entire sample, $\hat{S}_p(1)$. Next compute the Kaplan-Meier estimate of survival at one year based on the entire dataset with observation *i* removed $\hat{S}_p^{(i)}(1)$. The *i*th pseudo-value is defined by $\hat{\theta}_i = n\hat{S}_p(1) - (n-1)\hat{S}_p^{(i)}(1)$.

If there is no censoring, then the *i*th pseudo-value is simply the indicator that the *i*th recipient was alive at one year. These pseudo-values will then be used in a regression model using a logit link, similar to a standard logistic regression model, as described in the next section. The parameters of the regression model can be estimated using generalized estimating equations (GEE), which are implemented in PROC GENMOD in SAS.

II. Model building

Let $(Z_{i1},...,Z_{ip})$ denote the set of covariates in the final model for recipient i. First fit a fixed effects censored data logistic regression model with no center effect,

$$\varphi_i = \ln\frac{\theta_i}{1-\theta_i} = \beta_0 + \sum_{l=1}^{p}\beta_l Z_{il} .$$

III. Predicted and observed survival

From the fitted logistic regression model, each recipient has an estimated survival rate

$$\hat{p}_i = \frac{\exp(\hat{\varphi}_i)}{1+\exp(\hat{\varphi}_i)}$$

based on his or her risk characteristics. The predicted survival rate at center *j* based on recipient characteristics $E(S_j)$ is the average of the estimated survival rates for all recipients at center *j*,

$$E(S_j) = \left(\sum_{i\in C_j}\hat{p}_i\right) * \frac{1}{n_j} .$$

The observed one-year survival rate at center *j* can be computed using the Kaplan-Meier estimate of survival using the recipients at center *j*. This simplifies to the sample proportion of recipients alive when there is no censoring prior to one year present.

IV. Confidence Limits

Confidence limits are generated using a bootstrapping methodology. However, the bootstrap technique was modified slightly from previous years' reports to improve the coverage probabilities of the intervals, as described in Logan et al.[6] Previously, binary outcomes were generated for each individual to simulate the confidence limits; however, a more accurate prediction interval that controls the type I error rate can be obtained by re-sampling the residuals from the general linear model instead. Define the scaled Pearson residual for patient *i* by

$$r_i = \frac{\hat{\theta}_i - \hat{p}_i}{\sqrt{\hat{p}_i(1-\hat{p}_i)}},$$

then the bootstrap re-sampling algorithm to generate a prediction interval for center *j* is as follows. For *b*=1 to 10,000:

1. Generate r_i^{*b} for patient *i* by sampling with replacement from the set of residuals $\{r_i, i = 1,...,n\}$
2. Compute the bootstrap predicted value for patient *i* as $Y_i^{*b} = \hat{p}_i + r_i^{*b}\sqrt{\hat{p}_i(1-\hat{p}_i)}$
3. Compute the predicted center outcome for center *j* as $S_j^{*b} = \frac{1}{n_j}\sum_{i \in C_j} Y_i^{*b}$.

Then the 95% predicted confidence bounds for survival at center *j* are obtained by taking the 2.5th and 97.5th percentile of S_j^{*b} across the 10,000 bootstrap samples.

This confidence interval refers to the survival rate that might be observed at that center if there were no center effect and those recipients had been transplanted at any center in the network. The observed survival rate can be compared with this confidence interval to see if there is evidence of the center over-performing or under-performing the overall network.

Results

Risk factors

Based on the recommendation of the 2010 Center-Specific Outcomes Analysis Forum[5], variables recognized as clinically important are forced into the model regardless of whether they are statistically significant. After careful discussion with clinical and statistical transplant experts, the following essential risk factors are included in the model:

- Diagnosis (and disease status/stage)
- Donor type: matched sibling donor vs. other related vs. unrelated donor
- Coexisting disease (HCT-specific comorbidity index (HCT-CI), Sorror et al.[9])

- HLA matching[a]
- Recipient age
- Donor age (unrelated donor marrow or peripheral blood stem cells (PBSC) only)
- Recipient and donor gender
- Recipient cytomegalovirus (CMV) serology
- Recipient race/ethnicity (self-reported)
- Recipient Karnofsky/Lansky Performance Status score at transplant
- Prior autologous transplant
- Resistant disease (non-Hodgkin lymphoma (NHL) and Hodgkin lymphoma (HL) only)
- Time from diagnosis to transplant (acute leukemia not in first complete remission or Primary Induction failure (CR1/PIF))

In addition, the following variables are believed to be of uncertain clinical relevance, and so they are included in the model only if statistically significant ($p<0.05$).

- T-cell lineage in acute lymphoblastic leukemia, Philadelphia chromosome in acute lymphoblastic leukemia
- NHL subtype[b]
- Year of transplant
- Donor parity
- Donor race
- Donor CMV serology

For the 2015 report, preparative regimen intensity was removed as a distinct covariate in the multivariate adjustment model. In previous years, it had a statistically significant effect only for lymphomas, plasma cell disorders, chronic leukemia other than CML, and solid tumors. The general modeling approach for the center outcomes analysis has been to adjust primarily for patient and disease-related factors and minimize the adjustment for center-based factors that are at the discretion of the transplant centers. Additionally, for these diseases the choice of preparative regimen by the center is substantially correlated with other variables, including whether or not a previous autologous HCT has occurred, the patient's performance score, HCT-CI, and age. These patient- and disease-related variables remain in the model. Testing was performed to determine impact on the multivariate model, and removal of the preparative regimen intensity had little effect on the results of the modeling.

---

[a] For PBSC and marrow transplants, high-resolution typing at HLA-A, -B, -C, and –DRB1 is used for the cases where it is available. For the remaining patients with PBSC and bone marrow graft sources, the best available matching information at HLA-A, -B, -C, and -DRB1 is used.[7] For cord blood transplants, low/intermediate-resolution typing at HLA-A and -B, and high-resolution typing at -DRB1 only are considered.

[b] **NHL subtypes: Indolent B-cell lymphoma** includes the following sub-types: Lymphoplasmacytic; splenic marginal zone B-cell; extranodal marginal B-cell of mucosal-associated lymphoid tissue; nodal marginal zone B-cell; grades I, II and III follicle center cell; follicle center cell, grade unknown; Waldenstrom's macroglobulinemia. **Aggressive B-cell lymphoma** includes the following sub-types: Diffuse large B-cell including primary mediastinal B-cell lymphoma, Burkitt, and high grade Burkitt-like B-cell. **Mantle cell lymphoma** includes the following sub-type: Mantle cell lymphoma. **Nodal T-cell lymphoma** includes the following sub-types: Anaplastic large-cell T/null-cell, primary cutaneous type; peripheral T-cell NOS; angioimmunoblastic T-cell; anaplastic large cell T/null-cell, primary systemic type. **Extranodal T-cell lymphoma** includes the following sub-types: Extranodal NK/T-cell nasal type; enteropathy-type T-cell; hepatosplenic gamma-delta T-cell; subcutaneous panniculitis-like T-cell; mycosis fungoides; sezary syndrome; large T-cell granular lymphocytic leukemia; aggressive NK-cell leukemia; adult T-cell lymphoma/leukemia. **Other B-cell lymphoma. Other T-cell/NK-cell lymphoma.**

The results of the multivariate model are presented in a set of tables where each variable and its associated odds ratio are described, along with 95% confidence limits.

Center-specific results

Final center-specific results are presented, along with centers' historical performance in tables, and on the public website. Numbers of transplanted recipients at each center, actual (observed) survival at one year, predicted survival at one year, 95% confidence intervals for predicted survival, and performance status are displayed for each center. Centers whose actual survival is outside the 95% confidence limits for predicted survival have a "–1" in the performance status column if performing below the confidence limit, and a "1" in the performance status column if performing above the confidence limit. Centers with a "0" in the performance status column are performing as predicted. Since the censored data logistic regression model assumes no center effect, centers with smaller numbers of transplants (e.g. N = 1 or 2) will *not* have their predicted survival proportion regress toward the network average. Rather, the confidence limits around the predicted survival at that center will simply be much wider than those of larger centers.

Results are also displayed for centers via a visual box-plot graphic. Centers are arranged by center number, while reading from left to right across these figures. The actual survival at each center is superimposed with each box plot (using the symbol '•') to give the reader an instantaneous picture of how close the center is to under- or over-performing. A dashed line is included to denote the overall network survival average, using the Kaplan-Meier estimate of one-year survival from the entire cohort of patients who underwent first allogeneic HCT in the time interval included in the analysis.

Patients can find information about all U.S. transplant centers performing allogeneic transplants in the online U.S. Transplant Center Directory on http://bethematch.org. Listings are organized by state and can be found at bethematch.org/access. Along with center outcomes, each listing includes a description of that center's program, contact information, the number of transplants performed over a specified time period and survival statistics by patient's age, disease type and stage for both related and unrelated donor transplants. A link to the Transplant Center Directory can also be found on the Health Resources and Services Administration (HRSA) http://bloodcell.transplant.hrsa.gov website.

Because the outcome of interest is one-year survival, at least one year of follow-up time is required to be included in the analysis. Data are refreshed once a year. After the report on transplant center-specific survival rates is approved by HRSA, the Transplant Center Directory is repopulated with the new data.

Summary

A censored data logistic regression model is fitted to survival data for first unrelated and related donor hematopoietic cell transplants at U.S. centers. The model is adjusted for recipient age; recipient race/ethnicity; Karnofsky/Lansky score; coexisting disease (Sorror HCT-CI); recipient CMV status; disease/status; interval from diagnosis to transplant (ALL and AML only); NHL subtype; resistant disease (NHL and HL only); CLL and other chronic leukemia stage; prior autologous transplant; year of transplant; donor type/graft type/HLA matching; donor/recipient

gender match (bone marrow or PBSC only); and donor age (unrelated donor bone marrow or PBSC only). The report on transplant center-specific survival rates helps to identify centers that may have performed above or below confidence limits compared to the overall network of transplant centers during this specified time period.

References

1. National Marrow Donor Program Center-Specific Report. *Internal NMDP Report*, 1998.
2. Andersen PK, Klein JP and Rosthoj S. (2003). Generalized linear models or correlated pseudo-observations with applications to multi-state models. *Biometrika*, 90: 15-27.
3. Klein JP and Andersen PK. (2005) Regression modeling of competing risks data based on pseudo-values of the cumulative incidence function. *Biometrics*, 61: 223-229.
4. Klein JP, Logan BR, Harhoff M and Andersen PK. (2007) Analyzing survival curves at a fixed point in time. *Statistics in Medicine*, 26: 4505-4519.
5. Recommendations from the 2010 Center-Specific Outcomes Analysis Forum can be found at http://www.cibmtr.org/Meetings/Materials/CSOAForum/Documents/Center_Forum_Summary2010.pdf.
6. Logan BR, Nelson G and Klein JP. (2008). Analyzing center-specific outcomes in hematopoietic cell transplantation. *Lifetime Data Analysis*, 14: 389-404.
7. Weisdorf D, Spellman S, Haagenson M, et al. (2008). Classification of HLA-matching for retrospective analysis of unrelated donor transplantation: revised definitions to predict survival. *Biol Blood Marrow Transplant*, 14:748-758.
8. Bacigalupo A, Ballen K, Rizzo D, Giralt S, Lazarus H, Ho V, Apperley J, Slavin S, Pasquini M, Sandmaier BM, Barrett J, Blaise D, Lowski R and Horowitz M. (2009). Defining the intensity of conditioning regimens: working definitions. Biol Blood Marrow Transplant, 15:1628-1633.
9. Sorror ML, Maris MB, Storb R, Baron F, Sandmaier BM, Maloney DG, Storer B. (2005). Hematopoietic Cell Transplantation (HCT)-specific comorbidity index: a new tool for risk assessment before allogeneic HCT. *Blood*, 106(8):2912-2919.
10. Graf E, Schmoor C, Sauerbrei WF, and Schumacher M. (1999). Assessment and comparison of prognostic classification schemes for survival data. Statistics in Medicine, 18: 2529-2545.