Maternal Perspectives on Children’s Health-Related Quality of Life During the First Year After Pediatric Hematopoietic Stem Cell Transplant

Susan K. Parsons, MD, MRP, Mei-Chiung Shih, PhD, Katherine N. DuHamel, PhD, Jamie Ostroff, PhD, Deborah K. Mayer, RN, MSN, Jane Austin, PhD, D. Richard Martini, MD, Sharon E. Williams, PhD, Laura Mee, PhD, Sandra Sexson, MD, Sherrie H. Kaplan, PhD, William H. Redd, PhD, and Sharon Manne, PhD

1The Institute for Clinical Research and Health Policy Studies, Tufts-New England Medical Center, 2Tufts University School of Medicine, 3Clinical Research Program, Children’s Hospital, 4Harvard School of Public Health, 5Mount Sinai School of Medicine, 6Memorial Sloan-Kettering Cancer Center, 7Children’s Memorial Hospital, Northwestern University Medical School, 8Packard Children’s Hospital, Stanford University Medical School, 9Emory University School of Medicine, 10University of California, and 11Fox Chase Cancer Center

Objective To assess the longitudinal health-related quality of life (HRQL) of children receiving hematopoietic stem cell transplantation (HSCT). Methods Mothers (N = 160) of HSCT recipients aged 5–20 at six US transplant centers completed the Child Health Ratings Inventories (CHRIs), the Disease Impairment Inventory (DSII)–HSCT module, and the Short Form (SF)-36 at baseline, 3, 6, and 12 months. Results HRQL domain scores at baseline varied by recipient age and program site. Longitudinal data over the first year post-HSCT revealed lowest functioning at baseline and 3 months, with largest improvement in functioning between the 3 and 6-months assessments and continued improvement from 6 to 12 months. Recipients of unrelated donor transplants had steepest declines in functioning at 3 months and great HSCT-specific issues at 3 and 6 months. Among children who survived the first year, functioning at 12 months was similar across transplant types and surpassed baseline scores. Children who did not survive the first year exhibited deterioration in HRQL in the months before death and trajectories were strikingly different than for survivors. Conclusions This study offers the first glimpse of the 12-month trajectory of HRQL following pediatric HSCT from mothers’ perspectives. This study also highlights the importance of and approaches to addressing missing data in longitudinal research.

Key words children's self-assessment; health status; health-related quality of life; hematopoietic stem cell transplantation; longitudinal studies; parent report.

Since its inception more than 30 years ago, hematopoietic stem cell transplant (HSCT) has provided potentially life saving therapy to growing numbers of children and adults with malignancies as well as benign disorders of the bone marrow, the immune system, and metabolism. (Guinan, Krance, & Lehmann, 2002). There are two principal types of HSCT: allogeneic, in which the source of healthy hematopoietic progenitor cells are from a related or unrelated donor, and autologous, in which the progenitor cells come from the patient him or herself.
For further details about this treatment, the reader is referred to the cited chapter by Guinan et al. (2002) or http://www.marrow.org or http://cis.nci.nih.gov/fact/7_41.htm.] HSCT now represents the second most frequent major organ transplant in the United States; approximately 2,000 transplants are performed annually in patients less than 20 years old (Center for International Blood & Marrow Transplant Research, 2005). Improvements in supportive care and the control of transplant-related toxicity have led to a steadily increasing survivor cohort, now estimated at 90,000 (all ages) in the United States, representing 40% of survivors worldwide (Center for International Blood & Marrow Transplant Research, 2005).

The physiological and psychological demands of HSCT place recipients and their families uniquely at risk for altered health-related quality of life (HRQL).1 After a specific treatment regimen (transplant type, source of progenitor cells, and preparative regimen) is determined, based on the child’s underlying disease, the HSCT process begins with a prolonged hospitalization and/or intensive outpatient management at a specialized transplant center. All HSCT recipients experience physical isolation during this period, which may exacerbate feelings of emotional and social isolation. During the first several months following transplant, the recipient remains in protective isolation, removed from normal role activities although receiving frequent outpatient visits to monitor bone marrow function and to evaluate and manage complications (Cohen, Ley, & Tarzian, 2001). After 6–12 months, the recipient and family slowly begin to resume normal activities (Baker, Zabora, Polland, & Wingard, 1999; Parsons et al., in press). The child and parent attempt to reenter their social environments; for the child this means getting reacquainted with peers and returning to school, and for the parent, it means reestablishing family and work life. However, a return to the “old” normal may not be possible as they may face a “new” normal fraught with the physical, psychological, and social long-term effects of the HSCT. For some, the threat of disease recurrence continues to loom. The time trajectories for recovery may not necessarily be synchronous for the HSCT recipient and the parent and are likely to vary (Parsons et al., in press).

For more than a decade, researchers have grappled with issues of HRQL in adult HSCT survivors and have produced a literature from which many important lessons can be gleaned. Although many studies have revealed important intersubject differences, the general consensus is that overall HRQL for most adult survivors of HSCT is reasonably good and improves with time (Chao et al., 1992; Wingard, Curbow, Baker, & Piantadosi, 1991). A parabolic pattern of recovery, however, has been described (McQuellon et al., 1998) with worst functioning at the time of HSCT hospital discharge and improvement by day +100 post transplant. In studies by Chao (1992) and Syrjala (Syrjala, Chapko, Vitaliano, Cummings, & Sullivan, 1993) both physical and psychological functioning was lowest at day +90; of note, neither of these studies included an assessment at the time of discharge from the transplant. Patient characteristics associated with worse functioning include older age (Andrykowski et al., 1990; Andrykowski, Henslee, & Barnett, 1989; Schmidt et al., 1993), female gender (Wingard, 1994), and lower educational level (Andrykowski et al., 1990). Several study findings suggest that recovery following transplant is influenced by transplant type (autologous vs. allogeneic), the presence or absence of chronic graft versus host disease (GVHD) (Andrykowski et al., 1995; Schmidt et al., 1993; Syrjala et al., 1993; Wingard, 1994) and underlying disease (Bush, Donaldson, Haberman, Dacanay, & Sullivan, 2000).

In contrast to this growing body of research on adult HSCT survivors, research on pediatric HSCT survivors has lagged significantly behind. This is due in part to the lack of instruments designed to evaluate the multidimensionality of the HRQL construct in this population. The methodological challenges of child self-report, including its developmental relevance, accessibility, and the complex interpretation of proxy reporting pose another major barrier (Hinds & Haase, 1998; Parsons, Barlow, Levy, Supran, & Kaplan, 1999; Rebok et al., 2001). Two previous longitudinal studies in pediatric HSCT recipients focused on psychological adjustment and behavioral issues, utilizing the Child Behavior Checklist (CBCL) (Achenbach, 1983; Barrera, Boyd-Pringle, Sumbler, & Saunders, 2000; Phipps et al., 1995). In the Phipps study, changes in adjustment revealed decline in social competence and an improvement in problem behavior, both of which may reflect the decline in “opportunity,” because of forced isolation and intensity of treatment. In the Barrera study, no significant differences were found in behavioral summary

---

1Quality of life (QOL) is a subjective, multidimensional concept that encompasses a variety of domains, including physical health, psychological state, levels of functional independence, social relationships, environmental features, spiritual concerns, and personal beliefs (World Health Organization, 1947). The terms QOL and HRQL are often used interchangeably, although the later refers to the impact of the health states on overall QOL, principally one’s own health issues (Eisen, Donald, Ware, & Brook, 1990).
scores pre- and posttransplant. These two studies highlight the previously documented limitations of the CBCL in pediatric populations (Perrin, Stein, & Drotar, 1991): the CBCL, originally developed to identify pathological differences, may not be sensitive to the subtle differences within a group, particularly medically ill children and the instrument evaluates only one domain of HRQL, rather than a multidimensional construct.

Recent measurement development provides some insight into the HRQL issues for children following HSCT. Specifically, the Behavioral, Affective, and Somatic Experiences Scale (BASES), developed by Phipps and coworkers, measures the acute aspects of HRQL for children in five distinct areas: somatic distress, mood disturbance, quality of interactions, activity, and compliance (Phipps, Dunavant, Jayawardene, & Srivastiva, 1999). Parent, child, and nurse reports are available. Results demonstrate significant difference with time. Somatic distress (e.g., nausea and vomiting and mucositis) was greatest in the acute peri-transplant period, and as expected, had the highest cross-informant correlation. Correlations for nonobservable dimensions, such as mood, were lower (Phipps et al., 1999). Differences in patterns of scores were also noted by transplant type. As noted, the richness of this instrument is its ability to characterize HRQL in the acute peri-transplant period, particularly the impact of transplant on emotional distress and impaired functioning.

To enhance the measurement of HRQL of pediatric HSCT recipients both during and following the transplant, we developed an HSCT-specific module, the Disease Impairment Inventory–HSCT (DSII–HSCT) (Parsons et al., 1999; Parsons et al., in press), to be used with the Child Health Ratings Inventories (CHRIs), a generic HRQL measure for children with chronic disease (Kaplan et al., 1995). Both the CHRIs and the DSII–HSCT have two age-specified versions: 5–12 years and 13–21 years. In a cross-sectional application in 122 pediatric HSCT recipients and their parents, generic and disease-specific domain scores varied markedly by time posttransplant and by the clinical severity of the child’s condition as rated by the health care provider (Parsons et al., in press). Although the initial studies we conducted with the CHRIs instruments provided important “proof of concept” that this type of study could be conducted in this population, despite the morbidity and logistical aspects of follow-up care, the findings were limited by the cross-sectional design and a sample at a single institution.

In this report, we describe a multiinstitutional, longitudinal assessment over the first twelve months following HSCT, based on maternal reports of the child’s HRQL. Since this is one of the first studies of its kind in this population, the study aims were largely descriptive in nature, rather than testing hypotheses. We were particularly interested in quantifying the trajectory of the children’s recovery over the first year and to understand the link between the child’s HRQL and clinical status.

**Methods**

**Patients and Procedures**

This study reports on the sample of 160 mothers who participated in a larger study of maternal adaptation with pediatric HSCT at six transplant centers nationwide; the principal findings of the larger study are reported elsewhere (Manne et al., 2001; Manne et al., 2003; Manne et al., 2004; Rini et al., 2004). The sample for the current report was restricted to mothers of children aged 5–20 years at the time of HSCT, representing 56% of the total sample recruited for the larger study (n = 286). The remaining 44% of the original sample was mothers of children under five for which the CHRIs was not originally developed or validated. Eligible mothers were at least 18 years of age and the child’s primary or designated caregiver. They were required to possess a working knowledge of English and to be available for a baseline assessment before the child’s transplant. As previously reported (Rini et al., 2004), 31% of eligible mothers studywide declined participation. No differences were detected between consenters and refusers based on demographic characteristics, transplant type or disease status. The study was approved by the Human Subjects Committees at each of the participating institutions.

**Study Procedures**

After informed consent was obtained, participants completed a baseline assessment of study measures by self-report. These assessments were coordinated by trained study personnel at each of the participating sites. In addition to completion of the baseline assessment before the child’s transplant, follow-up assessments were obtained at 3, 6, and 12 months posttransplant. Participants were allowed to remain in the study if they missed only one of the scheduled assessments.

**Study Measures**

**Assessment of Child’s HRQL (CHRIs and DSII–HSCT)**

The principal outcome measure for the study was the maternal assessment of the child’s HRQL. At each of the designated time points, participants completed the 20-item, age-appropriate version of the CHRIs, a generic,
parent report measure. Although not collected at baseline (before the HSCT), participants also completed the 10-item HSCT-specific module, DSII-HSCT at each of the follow-up assessments. (Parsons et al., 1999; Parsons et al., in press). The generic and HSCT-specific questions are designed to assess children’s status in the week before the interview. All questions are introduced with the same stem: “During the past week, how much has not feeling well, or problems with health, gotten in the way when your child wants to do each of the following?” For the frequency items, a 5-point Likert scale ranged from “never” to “a whole lot of the time,” whereas for the intensity items, the 5-point scale ranged from “not at all” to “a whole lot.” For the HSCT items, the 5-point rating scale ranged from “1,” most to “5,” least. For most of the HSCT items, the respondent also was given the option of indicating that “this hasn’t happened to my child” or “my child doesn’t have to do this.” This choice was designed to distinguish between experiencing the problem and not being bothered by it versus not experiencing the problem within the reference period. The scores for generic and HSCT domains were then transformed to a 0–100-point scale using conventional psychological scaling methods (DeVellis, 2003). Similar to other HRQL instruments, higher scores for the generic items connoted higher level of functioning; for the HSCT items, higher scores connoted experiencing more problems (Aaronson et al., 1993).

The domains of HRQL for the CHRIs, derived from our cross sectional analyses of pediatric HSCT recipients and their families, included physical functioning, role functioning, emotional functioning, and energy (Parsons et al., in press). Of note, the energy domain, consisting of four items, included two items related to school attendance (“missed school” and “energy after school”). Given the extended hospitalization and required duration of protective isolation for 6–12 months following transplant, we were concerned that the two items would have little relevance to the population and may alter the structure of that domain. Therefore, we performed exploratory factor analysis with the two school-related items removed. The remaining two items, “energy to play” and “need to rest,” heavily loaded on the emotional functioning domain, supporting the inclusion of those items with emotional functioning. This grouping was also supported by the internal consistency reliability of the domain, as measured by Cronbach’s alpha (Cronbach, 1951) (α = 0.86–0.89 at all time points). In the analyses we report herein, we relied on a three-domain structure for the generic measure: physical functioning; role functioning; and emotional functioning. The HSCT-specific module also had three domains: hassles; worry; and body image, as originally hypothesized when the measure was developed.

Assessment of Maternal Quality of Life (SF-36)
To assess the mother’s own quality of life, participants completed the SF-36 at each scheduled assessment. This widely used measure, originally developed as part of the Medical Outcomes Study (Ware & Sherbourne, 1992), yields domain scores in multiple areas of general functioning (e.g., physical, emotional, role-physical, role-emotional, social) in addition to bodily pain, vitality, general health perception, and change in health (Ware, Snow, Kosinski, & Gandek, 1993). Mean summary scores were calculated for each domain of the SF-36 using established scoring algorithms (Ware et al., 1993). Composite scores (e.g., physical and mental) were not used, because comparisons were not directly available with the CHRIs.

Medical Assessment
After centralized training on data abstraction, including review of standard definitions and established toxicity scales, study personnel extracted medical information from the child’s medical record. At baseline, information was collected on the child’s diagnosis and disease stage, date of diagnosis, planned conditioning regimen, stem cell donor (autologous, related or unrelated allogeneic), and progenitor source (bone marrow, peripheral blood stem cells, cord blood). For allogeneic transplant recipients, information was also collected on the extent of the donor match and planned prophylaxis against acute GVHD, an immunologic reaction of the donor cells against the recipient. At each follow-up assessment, information was collected on the onset and severity of acute and chronic GVHD (Glucksberg et al., 1974; Shulman et al., 1980), regimen-related toxicity (Bearman et al., 1988), infection (National Cancer Institute, 1998), recurrent malignancies, and death.

Demographics
Demographic variables on age, gender, and race were collected on the child and mother; income and education were also collected from parent participants.

Statistical Methods
Participant and Patient Characteristics
Descriptive statistics were calculated to summarize the demographic and medical variables. Using established scoring algorithms, mean summary scores were calculated for each domain of the CHRIs and the DSII-HSCT.
Domain-specific Cronbach’s alpha (Cronbach, 1951) was also computed.

Baseline Evaluation
Multiple linear regressions were performed to evaluate whether baseline HRQL scores differed by treatment characteristics (indication for transplant, transplant type, relapse, and survival status at 12 months) or by recipient characteristics (gender and age) and to assess potential interactions between transplant type and survival on baseline HRQL scores. Age was evaluated both continuously and categorically, the latter defined by the groupings specified by the CHRIs (school-age, 5–12 years; adolescents, 13–21 years). These variables were considered first in a univariate and then in a multiple-covariate model. All models adjusted for program site.

Attrition
Given the expected mortality from HSCT over the first year (~30%), completion rates were computed for each period for the entire sample and by transplant type to inform how attrition should be addressed in the analysis. Over the 12-month period, sixty participants dropped out of the study for a completion rate of 63% at 12 months. Mortality was the major reason for attrition across all transplant types (43/60, 72%); voluntary drop out made up the remainder.

Assessing Missing Data
In the setting of this study, if missing HRQL data caused by lost follow-up of participating subjects were directly or likely related to the child’s condition and thus, also related to their unobserved HRQL measures, then the missingness would be nonignorable. It is well known that analyses ignoring this type of missing data can lead to biased results. To determine whether the attrition (which was principally caused by mortality) was ignorable, we used boxplots of HRQL scores by survival status to visually inspect and two-sample t-tests to formally test whether children who died before the next assessment had worse functioning at the current assessment than children who survived to the next assessment. Plots of mean HRQL scores were also used to explore the HRQL trajectories by survival status. These plots also informed the selection of the statistical method to deal with the missing data.

In addition, the study design permitted participants to skip one assessment and still remain in the study. Differences in clinical complications between respondents who skipped a scheduled assessment were compared with those who completed that assessment, using Fisher’s exact test, to evaluate whether this intermittent missingness was associated with clinical complications.

Longitudinal Data Assessment
We used a pattern mixture model (Fitzmaurice, Laird, &Shneyer, 2001) to handle the nonignorable missing data caused by mortality. The model included the main effects of program site, survival group, categorical time, transplant type, major HSCT complications (including infection, acute, and chronic GVHD; end-organ toxicity; and relapse), maternal emotional functioning, and recipient’s gender and age, as well as transplant type by time, survival status by time, and early infection by survival status interactions. The survival status by time interaction allowed the HRQL trajectory to be estimated separately for children who survived to at least 12 months (“survivors”) and those who died before 12 months (“nonsurvivors”). The transplant type by time interaction allowed the trajectory to differ between transplant types. Major HSCT complications were included in the model to examine whether the trajectory differences between transplant types could be explained by the anticipated variations in relevant clinical variables between transplant types. An unstructured covariance matrix was used to model the correlations between multiple HRQL scores on the same child. Likelihood ratio tests were used to evaluate main effects and interactions with two or more degrees of freedom; Wald tests were used to evaluate those with one degree of freedom.

More specifically, time was treated as a categorical variable with four different levels, allowing the mean HRQL to freely vary at the four assessment time points (baseline, 3, 6, and 12 months) without assuming a linear or quadratic parametric model; transplant type had three levels (autologous, related allogeneic, and unrelated allogeneic); and survival group had four levels (died before 3 months, died between 3 and 6 months, died between 6 and 12 months and survived beyond 12 months). For HSCT complications, adjacent ranks on each standardized scale were grouped together if they had statistically similar coefficients. Collapsed clinical variables that maintained the clinical integrity or meaningfulness of the scale were entered into subsequent models and their coefficients were allowed to vary by survivor status. In some cases, if the number of clinical events within the study sample was insufficient to estimate separate coefficients for survivors and nonsurvivors, we assumed a common coefficient. Maternal emotional functioning, as reported on the SF-36, was selected after examining Pearson’s correlations (Landis & Koch, 1977) between maternal reports of the child’s
functioning with reports of her own functioning. [For all domains of general and transplant-specific functioning, maternal emotional functioning was significantly correlated with reports of the child’s functioning (Pearson’s correlation coefficient $r = .18–.39$).] We also explored maternal emotional functioning by time and recipient’s age by time interactions, that is, whether the relation between maternal emotional functioning and HRQL score changed over time and whether the HRQL trajectory differed by recipient’s age. These interactions were not included in the final model because they were nonsignificant.

Of note, we did not consider the interaction between transplant type and survival or the three-way interaction between transplant type, survival, and time in the longitudinal data analysis. We assumed that the differences between survivors and nonsurvivors were the same for all three transplant types, after adjusting for other clinical variables.

All analyses were performed using SAS (SAS Institute, 2004). Statistical significance was determined by a $p$-value $\leq .05$.

Results

Participant and Patient Characteristics

The analysis included 160 mothers of children aged 5–20 years, receiving HSCT. The patient and disease characteristics of the transplant recipient are summarized in Table I. Most of the children (77%) had reported good grade of toxicity, although the remaining 23% had intermediate or poor grade toxicity, as previously defined. Among the 118 children who received an allogeneic HSCT, 27% developed intermediate or poor acute GVHD; 16% developed chronic GVHD. Systemic or disseminated infection occurred in 53% of children before day 100; late systemic or disseminated infection was much less common (18%). Among children with malignancies ($n = 136$), disease recurrence was reported in 15% of children by 12 months following transplant. Death from all causes was 27% ($n = 43$) by the end of the first year. These results highlight the variability in clinical outcomes for these children, as well as the relative infrequency of severe events.

Baseline Evaluation

In univariate analyses, significant differences in baseline HRQL scores were detected between program sites for all three generic domains. Recipient age, whether considered continuously or categorically, was significantly associated with physical and emotional functioning, but not with role functioning. No significant differences were detected between recipient gender or between transplant characteristics including transplant types, indications for transplant, 12-month relapse status, and 12-month survival status. Similar results were found in analyses with multiple covariates, taking all of these variables and age (as a continuous variable) together. In particular, for every one year increase in recipient age,
physical functioning and emotional functioning domain scores decreased by 1.1 and 1.3 points, respectively ($p = .005$; $p < .001$). Further, the interaction of transplant type by survival status was not significant for any generic domain.

**HRQL Trajectory by Domain**

Mean summary scores were computed for generic domains at baseline and for both generic and HSCT-specific domains at the three follow-up assessment periods, using all available data without consideration for the reasons for or timing of attrition (Table II). Data from parent reports of the CHRI and DSII–HSCT for pediatric HSCT survivors from our cross sectional data are included for comparison (Parsons et al., in press).

Among the three generic domains, lowest functioning was reported at baseline and 3 months. The largest improvement in functioning was detected between the 3 and 6-month assessments. Although scores continued to improve across all domains from 6 to 12 months, the greatest improvement was noted in physical functioning over that time interval. For the HSCT-specific domains, greatest hassles, worry, and body image issues were reported at the 3-month assessment, with steady improvement over the ensuing assessment time points. Of note, 12-month scores for generic and HSCT-specific domains were comparable to our previously reported cross sectional data of parent reports (Parsons et al., in press).

**HRQL by Transplant Type**

As noted above, one of the primary objectives of the study was to describe the HRQL trajectory by transplant type. Figure 1 shows the mean trajectories by transplant type for each domain based on all available data. Distinct paths emerged over the 12-month period by transplant type. While the paths for autologous and related allogeneic transplants were similar, the path for unrelated allogeneic transplants appeared different. Recipients of unrelated donor transplants had the steepest decline in generic functioning at 3 months. These recipients also had greater HSCT-specific issues at 3 and 6 months. Despite differences at 3 and 6 months among the three transplant types, generic and HSCT-specific functioning was similar by 12 months. For all generic domains, 12-month scores exceeded baseline scores. To formally evaluate whether these observed differences in HRQL trajectories between transplant types were statistically significant, we first needed to assess the nature of the missing data to determine the appropriate statistical method.

**Assessing Missing Data**

Figure 2 shows boxplots comparing the domain scores of children who survived to the next period with domain scores of children who died before the next period. For example, if a patient died before the 6-month assessment, his scores at 3 months were compared with 3-month scores for children who survived to at least the six month time period. These boxplots reveal that children who died before the next assessment had worse HRQL scores than those who survived to the next assessment, suggesting that the missing data due to death were nonignorable. These differences were statistically significant for generic domains ($p = .04$, two-sample $t$ tests) but did not achieve statistical significance for HSCT-specific domains ($p = .25$, two-sample $t$ test). We also examined the domain scores for those who voluntarily dropped out before the next assessment and found no clear differences with those who survived to the next assessment (data not shown). As for intermittent missing data, we compared medical conditions of completers who skipped the 3-month assessment ($n = 6$) with those who did not skip, and found that those who skipped tended to have more toxicity (50% vs. 14%, $p = .05$) and more early, severe infection (67% vs. 47%, $p = .42$).

**Methods for Handling Nonignorable Missing Data**

Figure 3 shows different approaches to handle nonignorable missing data. One approach is to model the

---

**Table II. Health-Related Quality of Life (HRQL) Domains By Time: Generic and Hematopoietic Stem Cell Transplantation (HSCT)-Specific HRQL Domains**

<table>
<thead>
<tr>
<th>Time point</th>
<th>$N$</th>
<th>Physical [M (SD)]</th>
<th>Emotional [M (SD)]</th>
<th>Role [M (SD)]</th>
<th>Hassles [M (SD)]</th>
<th>Worry [M (SD)]</th>
<th>Body image [M (SD)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>160</td>
<td>52.3 (30.5)</td>
<td>57.3 (20.9)</td>
<td>67.9 (27.3)</td>
<td>— $^a$</td>
<td>— $^b$</td>
<td>— $^b$</td>
</tr>
<tr>
<td>3 months</td>
<td>137</td>
<td>48.0 (32.1)</td>
<td>59.8 (20.6)</td>
<td>71.5 (27.8)</td>
<td>46.5 (16.3)</td>
<td>32.4 (25.6)</td>
<td>35.6 (23.6)</td>
</tr>
<tr>
<td>6 months</td>
<td>116</td>
<td>66.0 (28.2)</td>
<td>68.8 (18.6)</td>
<td>80.0 (23.4)</td>
<td>34.7 (17.6)</td>
<td>26.5 (23.2)</td>
<td>25.3 (23.6)</td>
</tr>
<tr>
<td>12 months</td>
<td>100</td>
<td>78.1 (25.8)</td>
<td>73.9 (19.0)</td>
<td>82.0 (25.9)</td>
<td>24.8 (13.8)</td>
<td>22.0 (24.0)</td>
<td>17.3 (22.2)</td>
</tr>
<tr>
<td>Survivors'</td>
<td>51</td>
<td>81.3 (25.5)</td>
<td>76.2 (16.0)</td>
<td>85.7 (25.1)</td>
<td>23.5 (11.2)</td>
<td>14.0 (20.0)</td>
<td>13.7 (12.7)</td>
</tr>
</tbody>
</table>

$^a$Higher generic scores connote better functioning while lower HSCT-specific scores connote better functioning on a 0–100 scale.

$^b$HSCT-specific domains assessed at follow-up time points only.

$^c$HSCT survivors from cross sectional study (Parsons et al., 2005).
Figure 1. Child Health Ratings Inventories (CHRI) generic and hematopoietic stem cell transplantation (HSCT) mean scores by transplant type.
trajectories separately for patients with different survival time. The mean trajectories for survivors (≥12 months) are compared with those for children who died between 3 and 6 months and with those for children who died between 6 and 12 months. Also shown in Fig. 3 are the mean trajectories based on simple imputation (death = worst score) and last observation carried forward. It can be seen that these two imputation techniques underestimate HRQL among study completers and overestimate or misrepresent the path for those who died before 12 months. As a result, we chose to use pattern mixture models to address the missing data issue, in which separate trajectories were developed for survivors and nonsurvivors.

**Pattern Mixture Model**

The results of the pattern mixture model for major HSCT complications are presented in Tables III and IV for generic and HSCT-specific domains, respectively. Each coefficient represents the difference in domain scores between survivors and nonsurvivors.

---

2Although the main effect of survival group (with 4 levels: died before 3 months, died between 3 and 6 months, died between 6 and 12 months, survived beyond 12 months) is included in the model, the fitted coefficients and the standard errors are not included in Tables III and IV. The interaction of survival and time was significant, as determined by a three degree of freedom likelihood ratio test, in contrast to the Wald test for other one degree of freedom clinical variables in the table.
scores for those with the identified HSCT complication compared to those without that complication. Reflecting established scoring algorithms, a negative coefficient indicates worse functioning in the generic domains, while a positive coefficient indicates greater problems for the HSCT-specific domains. Overall, the impact of transplant-related complications differed both by domain and by survivor status. To illustrate this, consider the impact of toxicity on children's emotional functioning. Children with "intermediate or poor" overall toxicity had emotional functioning scores that were 15.0 points lower (worse) than those for children with "good" overall toxicity \((p < .001)\). However, the impact of toxicity on physical and role functioning was smaller and not statistically significant. In contrast, the fitted model indicated that the impact of early infection on HRQL was larger in nonsurvivors than in survivors for all three generic domains, even though the difference (characterized by interaction between early infection and survival status) was statistically significant for emotional functioning but not for physical or role functioning. In particular, for survivors, physical functioning was on average 2.1 points lower (worse) for those with serious early infections, as compared to those with localized or no infections. For nonsurvivors, physical functioning for those with serious early infection was 25.0 points lower than those with localized or no infection. For HSCT-specific domains, chronic GVHD resulted in a 27.5-point increase in the HSCT domain of hassles \((p < .001)\) and a 23.2-point increase in body image \((p < .001)\). Paradoxically, chronic GVHD resulted in a 9.9-point decrease in worry. See Tables III and IV for fitted coefficients of other HSCT complications.

After adjusting for all of the HSCT complications as well as for program site, maternal emotional functioning, and recipient's age and gender, the transplant type by time interaction was statistically significant for physical functioning \((p = .003)\), hassles \((p = .01)\), and body image \((p = .03)\), but not for role functioning, emotional functioning, or worry. Examination of the fitted interaction coefficients revealed that the nature of the interaction was similar to that observed in Fig. 1.

In a similar manner, after adjusting for transplant characteristics and complications, program site, maternal emotional functioning, and recipient's age and gender, the survival status by time interaction was statistically significant for role \((p = .04)\) and emotional functioning \((p = .001)\), marginal for physical functioning \((p = .06)\) and body image \((p = .06)\), but nonsignificant for hassles or worry. Examination of the fitted interaction coefficients revealed that the nature of the interaction was similar to that observed in Fig. 3.

Further, statistically significant differences across program sites were noted for hassles \((p < .001)\), but not for other domains. However, because data were not collected prospectively to evaluate institutional practices related to protective isolation and dietary modification in the posttransplant period, we cannot evaluate the impact on perceived hassles.

Of note, after adjusting for the clinical variables, recipient age and mother's emotional health remained significantly associated with HRQL scores for all generic and HSCT domains. Noting that higher generic scores and lower HSCT scores each connote better functioning, the fitted model reveals that HRQL functioning worsened as age increased and as mother's emotional functioning worsened. Specifically, for every one year
Figure 3. Child Health Ratings Inventories (CHRIs) generic and hematopoietic stem cell transplantation (HSCT) mean scores by survival status and by imputation.
increase in recipient age, generic domain scores decreased by 0.61–1.24 points; HSCT-specific domain scores increased 0.42–1.19 points. Similarly, for every one point increase in mother’s emotional functioning, children’s generic scores increased by 0.26–0.33 points and HSCT scores decreased by 0.20–0.42 points (p < .001). In addition, female recipients experienced significantly worse scores in body image (p < .05); no other gender differences were detected.

Discussion
In this report, we describe maternal perceptions of the impact of HSCT on children’s HRQL during the first year following transplant. Given the longitudinal data of 160 participants, we are able to characterize HRQL by generic and HSCT-specific domains for each transplant type (i.e., autologous, related-allogeneic, unrelated-allogeneic) as well as for survivors and nonsurvivors. Our results showed that functioning, although similar at baseline, varied significantly by transplant type, by survival status, and by time posttransplant during the 12-month follow-up period.

The study revealed that, among the study survivors, functioning by the end of the first year had at least recovered to the level of the baseline assessment, which was obtained just before having the HSCT. However, since the pretransplant baseline is likely not to reflect a “true” premorbid baseline, the return to the level of baseline functioning may still indicate erosion in functioning. The recent study of adult HSCT survivors at a mean of 7.0 years post-HSCT indicated significantly worse functioning (physical health, psychological adjustment, and physical function) in HSCT survivors.

Table III. Pattern Mixture Model: Effect of Variables on Generic Health-Related Quality of Life (HRQL) Domain Scores

<table>
<thead>
<tr>
<th>Variable</th>
<th>Physical coefficient (SE)</th>
<th>Emotional coefficient (SE)</th>
<th>Role coefficient (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute GVHD</td>
<td>−7.8 (11.4)</td>
<td>6.5 (6.5)</td>
<td>−6.6 (11.7)</td>
</tr>
<tr>
<td>Chronic GVHD</td>
<td>−12.4 (7.2)</td>
<td>−3.5 (4.6)</td>
<td>−6.9 (7.0)</td>
</tr>
<tr>
<td>Early infection</td>
<td>−2.1 (5.4)</td>
<td>0.3 (2.9)</td>
<td>−0.7 (5.4)</td>
</tr>
<tr>
<td>Late infection</td>
<td>−12.6 (5.2)*</td>
<td>−4.0 (3.1)</td>
<td>−5.9 (5.1)</td>
</tr>
<tr>
<td>Early infection × survival status</td>
<td>−22.9 (14.9)</td>
<td>−20.7 (8.2)*</td>
<td>−23.9 (14.5)</td>
</tr>
<tr>
<td>Maximum toxicity within 3 months</td>
<td>−8.1 (7.0)</td>
<td>−15.0 (3.8)**</td>
<td>−10.9 (6.8)</td>
</tr>
<tr>
<td>Relapse</td>
<td>9.1 (8.2)</td>
<td>−8.9 (5.1)</td>
<td>1.0 (8.1)</td>
</tr>
<tr>
<td>Mother’s emotional health</td>
<td>0.33 (0.08)**</td>
<td>0.29 (0.05)**</td>
<td>0.26 (0.07)**</td>
</tr>
<tr>
<td>Child’s age</td>
<td>−1.24 (0.36)**</td>
<td>−1.13 (0.24)**</td>
<td>−0.61 (0.35)</td>
</tr>
<tr>
<td>Child’s gender</td>
<td>−0.4 (3.0)</td>
<td>2.3 (2.1)</td>
<td>0.9 (2.9)</td>
</tr>
</tbody>
</table>

GVHD, graft versus host disease.
Difference in domain coefficients for those with the identified major hematopoietic stem cell transplantation (HSCT) complications. A negative score indicates a worse functioning compared to those without that complication.
*p < .05. **p < .01. ***p < .001.

Table IV. Pattern Mixture Model: Effect of Clinical Variables on Hematopoietic Stem Cell Transplantation (HSCT)-Specific Health-Related Quality of Life (HRQL) Domains

<table>
<thead>
<tr>
<th>Clinical variable</th>
<th>Hassles coefficient (SE)</th>
<th>Worry coefficient (SE)</th>
<th>Body Image coefficient (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acute GVHD</td>
<td>−5.2 (5.6)</td>
<td>−18.9 (9.8)</td>
<td>−8.6 (9.8)</td>
</tr>
<tr>
<td>Chronic GVHD</td>
<td>27.5 (3.3)**</td>
<td>−9.9 (5.6)</td>
<td>23.2 (6.2)**</td>
</tr>
<tr>
<td>Early infection</td>
<td>3.3 (2.4)</td>
<td>5.7 (3.8)</td>
<td>0.8 (4.3)</td>
</tr>
<tr>
<td>Late infection</td>
<td>4.6 (2.7)</td>
<td>0.1 (3.9)</td>
<td>6.1 (4.5)</td>
</tr>
<tr>
<td>Early infection × survival status</td>
<td>16.6 (6.6)*</td>
<td>5.6 (11.6)</td>
<td>18.4 (11.3)</td>
</tr>
<tr>
<td>Maximum toxicity within 3 months</td>
<td>8.5 (3.1)**</td>
<td>4.6 (5.1)</td>
<td>1.1 (5.4)</td>
</tr>
<tr>
<td>Relapse</td>
<td>−0.6 (4.1)</td>
<td>9.6 (6.8)</td>
<td>−3.5 (6.9)</td>
</tr>
<tr>
<td>Mother’s emotional health</td>
<td>−0.20 (0.04)**</td>
<td>−0.42 (0.07)**</td>
<td>−0.21 (0.08)**</td>
</tr>
<tr>
<td>Child’s age</td>
<td>0.42 (0.20)*</td>
<td>1.19 (0.42)**</td>
<td>0.74 (0.36)*</td>
</tr>
<tr>
<td>Child’s gender</td>
<td>−2.0 (1.7)</td>
<td>−1.9 (3.6)</td>
<td>−7.3 (3.1)*</td>
</tr>
</tbody>
</table>

GVHD, graft versus host disease.
Differences in domain coefficients for those with the identified major HSCT complications compared to those without that complication. A positive score indicates a worse functioning.
*p < .05. **p < .01. ***p < .001.
when compared with healthy, age, and sex-matched controls (Andrykowski et al., 2005).

In contrast to the trajectories for survivors, our results reveal that the HRQL scores of nonsurvivors declined even in the months before death, suggesting a difference in clinical course. Significant differences between survivors and nonsurvivors were found in all three of the generic HRQL domains and HSCT domain of body image, but not in the HSCT domains of worry and hassles. This highlights an important study design consideration. Specifically, the first scheduled posttransplant assessment was at 3 months; this also marked the first time that HSCT-specific functioning was collected. Earlier posttransplant assessments are needed to characterize more robustly the change over time in the HSCT-specific domains, particularly in evaluating the impact of early mortality. We have subsequently implemented earlier assessments in ongoing studies in this population.

Our findings also mirror those of Syrjala (1993) and Chao (1992) by demonstrating worse functioning at the 3-month point following HSCT. In our study, worst functioning and greatest decline in functioning were observed in recipients of unrelated allogeneic transplants at 3 months posttransplant. Controlling for HSCT complications, the interaction between time and transplant type remained significant for physical functioning ($p = .003$), hassles ($p = .01$), and body image ($p = .03$), but not for role functioning, emotional functioning, or worry. This may reflect clinical differences in medication regimens for recipients of unrelated donor transplants, such as the use of corticosteroids for GVHD prophylaxis.

Recipient age was significantly associated with baseline as well as posttransplant HRQL scores. Specifically, older age was associated with worse generic and HSCT-specific functioning. The significance of this finding is that the age effect is over and above the differences in clinical course. The reason for this may be related to the developmental differences in the way in which illness is experienced and expressed by children as they get older.

This study highlights the need to formally address issues related to missing data in longitudinal analyses. In particular, given the likely connection between loss to follow-up and participants’ illness severity, we were concerned that failure to address the nonignorable missingness would produce biased results. For example, if an assessment were missing because the caregiver was too distressed to respond or the child was too sick, the information about missing would likely be directly related to the child’s unobserved HRQL. Through visual inspection of HRQL domain scores over time for those who remained in the study compared with those who died, we concluded that the missing data due to attrition were nonignorable.

We used a pattern mixture model (Fitzmaurice et al., 2001) to handle attrition based on our observation that the trajectory of HRQL differed between survivors and nonsurvivors. This type of modeling allows us to describe the trajectories separately for the two groups. With the simultaneous adjustment for relevant clinical variables, the pattern mixture model also allowed us to explore whether the differences in trajectories could be explained by specific transplant complications. We illustrated that the commonly used imputation methods, including simple imputation (i.e., missingness due to death is imputed by the worst possible score) and “last observation carried forward” (LOCF), which provides a conservative approach to connote absence of change from last observation (Mallinckrodt, Clark, & David, 2001) were not appropriate with nonignorable missing data because the resultant trajectories from these strategies do not directly inform the experience of either the survivors or to those who drop out for other reasons. Several alternative methods have recently been proposed, including multiple imputation in random effects models (Ali & Siddiqui, 2002) and Bayesian approaches (Carpenter, Pocock, & Lamm, 2002). Further research is needed to assess these alternative methods compared with the pattern mixture model, particularly with regard to the joint modeling of survival time and HRQL (Schluchter, 1992).

We were struck by the strong relationship between maternal ratings of the child’s functioning with ratings of her own emotional functioning. One explanation for this is that both measures were completed by the same rater and may reflect the common variance shared between the two mother reported measures. However, the other domains of maternal functioning from the SF-36 were not as strongly correlated with maternal reports of the child’s function. An alternative explanation is the well-documented collateral impact sustained by maternal caregivers during the HSCT process (Manne et al., 2001). This impact, however, may threaten the ability of mothers to serve as “objective” proxy raters. Moreover, given the focus on mothers’ perceptions, this study does not provide any information about how children experience the transplant process or how parents and children’s perception may compare. A direct comparison between parent and child ratings of the child’s functioning after pediatric HSCT is under evaluation in ongoing studies in our research group.

Our results, however, also highlight some important limitations of this study. Much of the initial measurement
work for the CHRI was done in cross-sectional samples, including ours in pediatric HSCT (Parsons et al., in press). In this longitudinal application, we were concerned about the relevance of two items related to school attendance, since children cannot attend school in weeks to months following their transplant. After evaluating the impact of removing these items on the factor structure of the measure, we elected to use the resultant three-factor structure because it made more conceptual sense than the original four-component structure. Further studies will be needed to evaluate the robustness of this structure, as well as the longitudinal stability of the factor structure in this population. We are mindful, however, that such an analysis would require a sample of at least 200–300, to achieve the recommended participants: item ratio of 10:1 (Nunnally & Bernstein, 1994).

Despite the seemingly robust sample of 160 participants in the study, we observed a low rate of severe clinical events. As a result, the point estimates (fitted coefficients) for these severe clinical events typically had large standard errors, which led to unexpected statistical nonsignificance, even though these severe clinical events would be expected to result in worse HRQL. Substantially greater numbers of participants (~400 or more) would be required to have sufficient power to detect significant differences for some of the clinical events for selected HRQL domains. Therefore these results are not meant to imply that there is not an HRQL impact; rather, we are unable to detect its significance with the existing sample size. We are currently collecting data to address both the measurement-related and clinically related issues.

The sample is also limited by its sociodemographic profile, which includes only English speakers who are largely Caucasian, well educated, and insured. Therefore, the generalizability of our findings to other populations must be done cautiously.

Despite these limitations, this study offers the first glimpse of the 12-month trajectory of HRQL for children following pediatric HSCT, based on parent proxy report. Three important predictors of HRQL have emerged from this study: transplant type, survival status, and recipient age. With regard to age, we have shown a linear relationship between increased age and worsening HRQL information about factors influencing HRQL is sorely needed to guide decision-making, counsel and support recipients and their families, educate their health care providers, and guide the development of interventions to ameliorate the impact on HSCT on recipients and their families. Further, based on the observed decline in HRQL for children who did not survive to the end of the first year, specific interventions for the child and family may be needed, distinct from those for survivors.

These results are especially important to pediatric transplant clinicians who have been reluctant to apply previous research on adult transplant recipients to their pediatric patients (Lee et al., 2004). Research is underway to confirm our findings, to evaluate further the impact of clinical events on HRQL, and, perhaps most importantly, learn from the recipients themselves how they perceive the HSCT experience.

Acknowledgments

Sponsored in part by Grant MH 57738, awarded by the National Cancer Institute and the National Institute of Mental Health (WH Redd, PI), a Career Development Award from the American Society of Clinical Oncology (SKP), and the American Cancer Society Research Scholar Grant RSGPB 02-186-01-PBP (SKP).

References


