CIBMTR® CENTER FOR INTERNATIONAL BLOOD & MARROW TRANSPLANT RESEARCH

Final Summary and Recommendations of the Center Outcomes Forum December 18, 2008

Background

In 1986, the National Bone Marrow Donor Registry (managed by the National Marrow Donor Program (NMDP)) was established, with responsibility for the maintenance of an unrelated donor registry for hematopoietic cell transplantation. In 1990, the Transplants Amendment Act made the reporting of center-specific outcomes for unrelated hematopoietic cell transplantation (HCT) mandatory in the United States. This activity has been conducted by the NMDP since 1994. With the Stem Cell Therapeutic and Research Act of 2005, the requirement to report outcomes of HCT by transplant center was broadened to include all allogeneic (related and unrelated) HCTs in the United States. This responsibility rests with the contractor for the Stem Cell Therapeutic Outcomes Database, the Center for International Blood and Marrow Transplant Research (CIBMTR).

The CIBMTR has collaborated closely with the NMDP since 2003 in the generation of center outcomes reports for unrelated donor HCT for the Health Resources and Services Administration (HRSA), a division of the U.S. Department of Health and Human Services. These reports are well-accepted by the HCT community. However, the Stem Cell Act of 2005 substantially expands the patient population to be considered in these analyses. At most centers, the new requirement means that the percentage of patients to be included will at least double. Some centers, those who do not do unrelated donor HCTs, will be included in these analyses for the first time

During this transition phase of the C.W. Bill Young Cell Transplantation Program, CIBMTR, working with the NMDP, the American Society for Blood and Marrow Transplantation (ASBMT) and HRSA, determined that a meeting should be held to review the current approach to center-specific outcomes reporting and to provide recommendations for future reports in the expanded Program. With this purpose, CIBMTR invited representatives of the HCT community (national and international), the ASBMT Quality Outcomes Committee, governmental funding agencies, the solid organ transplant community, patients, private payers, statisticians and experts in hospital and quality outcomes reporting to Milwaukee, Wisconsin on September 26 and 27, 2008. A complete list of invited attendees can be found in Appendix A.

The objectives of the meeting were to review the current state of center-specific outcomes reporting in medicine and transplantation and to openly discuss strengths and limitations of current approaches with the goal of developing recommendations for HCT center outcomes reports that would be

- scientifically valid;
- equitable;
- free from bias;
- useful to the HCT community for improving quality;
- informative for the public.

The Center-Specific Outcomes Analysis Planning Forum was conducted in two parts. The first involved didactic presentations to establish a common background understanding among participants. Presentations covered topics regarding current center outcomes programs in HCT and solid organ transplantation, trends in hospital outcomes reporting in medicine, risk factors known to affect HCT outcomes, risk adjustment in hospital outcomes reporting, using outcomes reporting for quality improvement and presentation of hospital outcomes to the lay community. A panel discussion reviewed the uses of center outcomes analyses from the perspective of payers, patients, centers and the government.

The second part of the forum involved breakout sessions devoted to specific aspects of center-specific outcomes reporting HCT, each charged with generating recommendations for CIBMTR regarding the process and methodology of future reports. What follows is a summary of the discussion and recommendations from those breakout sessions.

CONSIDERATIONS FOR STATISTICAL ANALYSIS OF CENTER EFFECTS

Statistical Modeling

Statistical models used to generate the current center-specific outcomes report were discussed. In brief, each center's expected outcomes are generated based on a regression model adjusting for patient characteristics. A 95% prediction interval is produced for each center, which represents the range of expected outcomes if patients at that center had been transplanted at a generic center in the network. The observed outcomes for each center can be directly compared to this confidence interval to determine if the center is performing at a level consistent with the overall network or under-performing or over-performing compared to the overall network. The width of this confidence interval reflects the number of patients transplanted at that center such that small centers have wider intervals. This means that they are not more likely to be mis-identified as underperforming but also decreases the power to detect a true difference in performance. An alternative proposal was made to consider using *hierarchical modeling*, which may better estimate the influence of the collective set of center effects on overall network outcomes and may be advantageous if the generated center outcomes were to be used to compare performance against a fixed outcome benchmark. For example, a payer group may designate that probability of survival at a center must meet or exceed the absolute value of 75% to be acceptable, without considering random variability introduced by

different sample sizes. If payers were to use the center-specific outcomes as currently generated (using regression modeling techniques) to compare to a fixed benchmark criterion, the effect of center size would not be accounted for appropriately and small centers with greater degrees of expected random variability in survival would be more likely to be inappropriately penalized. Payers expressed willingness to review their current methodologies and felt it was important to align as best as possible with this new process. Hierarchical modeling produces "shrinkage" of estimates of center effects based on a random normal distribution such that estimates tend to be closer to the average with the amount of shrinkage for a given center depending on the center size, i.e., small centers are more likely to be shrunk to the average and less likely to be penalized unfairly. However, one disadvantage of hierarchical modeling is that, due to the shrinkage of center effects, it is less efficient for identifying outlier centers than the current model. The resulting center-specific outcomes are also more relevant to the *overall* network patient population rather than to the specific patient-mix *at a given center*. The primary objective of these center-specific analyses is to provide tools by which centers may assess performance, based upon the patients *actually transplanted* at their center, rather than supporting a benchmarking approach. Therefore, the current modeling approach was considered the most appropriate.

Recommendation:

- Maintain current regression modeling approach. Since the primary objective of the center outcomes report is to provide centers a tool by which to measure and improve outcomes, models that are more sensitive to outliers are appropriate.
- CIBMTR should generate educational materials that clearly explain the hazards of using the output of the current regression model for comparisons to a benchmark using a fixed outcome target.

Previous NMDP center outcomes reports have used *logistic regression modeling* at a fixed time point (1 year). This modeling method accounts for the potential for crossing hazards, as may be seen due to the shift in mortality curves from reduced intensity conditioning. Outcomes reports generated by other organizations have used Cox regression modeling. The rationale for use of logistic regression was supported, and a suggestion was made to consider goodness of fit of the logistic link function.

Recommendation:

- Continue use of logistic regression for survival at a fixed time point, and evaluate goodness of fit of the logistic link function.

Reference Patient Population

A second topic considered by the group was the *patient population used to determine expected outcomes*. Currently, expected outcomes for patients actually transplanted at a given center are computed, so that the observed outcome at that same center can be directly compared to the expected outcome in that same group of patients. An alternative would be to estimate a center effect for each center, and then estimate the center outcome if all patients in the entire network had been transplanted at a given center. This approach would facilitate direct comparisons among centers since the expected outcomes would be

the same for all centers. This has been attempted in an earlier iteration of the NMDP center specific analysis, and was abandoned as too confusing to patients. Additionally, the expected outcomes lacked face validity for transplant centers, since they were not relevant to the actual patient mix at the transplant center. Similar to the discussion regarding hierarchical modeling, this approach is more useful when comparing center outcomes against a fixed benchmark, which is not the intended approach.

Recommendation:

- Continue computation of expected outcomes for patients actually transplanted at each transplant center.

Outcomes measure

The current *outcome measure* assessed in the center-specific outcomes analysis is the probability of survival (at one year, as discussed above). This measure was supported by the group, since it captures most adverse events and is still likely to be sensitive to center effects (as opposed to events that occur later when the patient is no longer under the care of the transplant center). It was suggested that the language could be simplified to "fraction alive at one year" and "expected fraction alive at one year" and which might improve comprehension by public users. There was discussion regarding presentation and inclusion of other outcomes. The group suggested that 100 day mortality reflects the time period most under the influence of centers and its presentation would be a useful measure for centers to review. Since 100 day mortality is likely to differ between myeloablative and reduced intensity conditioning regimens, this measure should be considered informative rather than a benchmark. Time points beyond 1 year survival were considered important and relevant to patients, but a less sensitive measure of center practices.

Recommendations:

- Report probability of survival at one year for the primary center outcomes measure.
- Present simplified language as fractions alive at one year.
- Analyze and present to centers the 100 day mortality stratified by conditioning regimen intensity (myeloablative and reduced intensity).

Handling multiple comparisons

Traditionally, when a statistical analysis involves *multiple comparisons*, the p-value used for statistical significance is adjusted to avoid artifactual significance (often to a p-value of 0.01 or less). Modeling for center outcomes presents a dilemma, since a large number of centers are analyzed simultaneously, but adjusting the p-value considered significant across multiple centers may decrease the ability to detect "true" center effects. Following discussion, it was determined that patients are most interested in the performance of a single center, and transplant centers are most interested in understanding their own performance. For these perspectives unadjusted error rates (5%) represent the most appropriate approach. Similarly, an error rate of 5% maintains reasonable power to identify centers whose performance should be improved. However, caution is warranted not to over-interpret results of center outcomes modeling that are not adjusted for

multiple comparisons, particularly in making comparisons among centers rather than of a single center's observed versus expected performance.

Recommendation:

- Present results of center outcomes *without* adjustment for multiple comparisons.

Model Validation

Validation of modeling techniques over time is important to maintain confidence in the integrity of center outcomes reporting. Current consistency of the model performance over time suggests a stable model. Other techniques to measure goodness of fit of the models include the Brier score, C-statistic, and Hosmer-Lemeshow tests. These goodness of fit tests should also be examined in special subgroups of patients like pediatric and high-risk patients.

Recommendation:

- CIBMTR should include measures of goodness of fit in the modeling process, and present these measures to the groups who periodically review the modeling process.

REVIEW OF OUTCOME MEASURES AND SAMPLE SIZE CONSTRAINTS.

Presentation of survival vs. benchmarking

Current HCT center outcomes reports focus on presenting results of 1 year survival where observed survival is significantly lower or higher than expected. Other organizations establish a *benchmark* using a ratio of observed to expected or the difference between observed and expected. Creating a benchmark was believed to add little additional value compared to presentation of observed survival and whether its deviation from expected is statistically significant.

Recommendation:

- CIBMTR should present observed survival for each center, along with test of significance against expected survival at that center.

Handling small centers

A common question is whether outcomes at small centers should be handled differently than larger centers when considering observed versus expected outcomes, since their sample size leads to wide confidence intervals surrounding expected one year survival. In other words, a substantial number of deaths would need to occur before the mortality would be considered statistically significantly different from expected at a small center. As an example, the Scientific Registry of Transplant Recipients (SRTR) has an approach where any death at a small center is considered for review in their center outcomes report. The HCT community agreed that the wide range of expected patient survival across a broad range of diseases and disease states makes it impractical to consider a single death in a small HCT center as an outlier requiring review. As well, with inclusion of both

related and unrelated donor HCT in the center outcomes reports, fewer centers will be extremely small.

Recommendation:

- Analyze small centers with the same methods regarding outcomes as large centers, and report the same criteria for observed survival and statistical deviation from expected survival.

Time window for inclusion in analysis

The current center outcomes report uses all unrelated transplants within a *5 year time window,* in order to increase the power of the analysis. Criticisms of this approach include concern that the wide time range does not adequately reflect programmatic changes that may occur during the interval, and does not sufficiently focus on most recent results. There was discussion regarding use of different time windows for different audiences, or for different size centers. For example, a shorter time window may be possible for large centers with sufficient sample size, whereas a longer time window might be used for smaller centers. However, variable time windows may lead to confusion and could bias the comparisons. One methodologic suggestion that may overcome shortcomings of the time window approach is the cumulative sums (cusums) continual assessment method. As well, CIBMTR could consider using a 5 year time window of patients across all centers to build a stable adjustment model then base the actual center analysis on a shorter three year time window.

Recommendations:

- Review sample size and model performance for the first related and unrelated center outcomes analysis in 2010 with a 3 year time window.
- Test the cusum method with the 2010 analysis.

CONSIDERATIONS OF PROCESS FOR MODEL REVIEW, FUTURE MODIFICATIONS AND INTRODUCTION OF NEW FACTORS IN THE CENTER OUTCOMES REPORT

Adherence and Schedule for Model Review

The center specific analysis model and outcome should be reviewed by the community within the first year of its publication and then every two years thereafter. It was proposed to use existing groups representing distinct viewpoints within the transplant community to assess the model and outcome. Transplant centers should be represented by the ASBMT Quality Outcomes Committee and perhaps a broader group of transplant center representatives nominated by the ASBMT. Patients' interests should be represented by the CIBMTR Consumer Advocacy Committee, perhaps augmented by patient representatives with particular interest in these statistical reports. Finally, technical aspects of the model should be assessed by a group of biostatistical consultants from transplant centers, cancer centers with experience in survival analysis, and experts in center outcomes analysis from other disciplines. These periodic reviews should assess the model's utility and reliability to analyze center performance. Performance of models

should be evaluated across all centers, but also separately in centers with larger groups of special populations such as pediatric recipients, recipients of cord blood transplants, and recipients of reduced intensity conditioning. The robustness of the model in analyzing large centers compared to small centers should also be considered.

Recommendation:

- Develop a process to perform a multifaceted review of the first combined center outcomes analysis within the first year, followed by biennial review.

Yearly Data Submission, Review and Correction

Centers should have the ability to review and correct the data used to conduct center outcomes analyses. An annual timeline should be established for data submission, compilation, assessment for completeness, and review by the transplant centers, perhaps by April - May 1st each year. This timeline should coincide, where possible, with the CIBMTR Continuous Process Improvement (CPI) process. Data provided to the center should include patient level data for those items to be included in the outcomes analysis. Centers should receive some time (a maximum of 30-45 days was recommended) for error correction and improving completeness of follow-up. After completion of the statistical analysis, perhaps by July 1st, centers should again be given the opportunity to review the outcomes results and offer commentary. At this stage, no modifications of the analysis would be permitted, however centers' comments could be included in the published SCTOD center specific outcomes report.

To facilitate transplant center reviews of submitted data, CIBMTR should consider providing summarized unadjusted outcomes statistics for each center. These reports may include survival estimates in specific disease subsets (AML in first or later remission, lymphoma, myeloma, using different donor types) which centers could quickly assess for face validity. Similarly, CIBMTR should consider developing a query tool that provides center directors the capability to evaluate their submitted data for completeness and accuracy.

Recommendation:

- Develop a cyclic timeline designed to provide centers the opportunity to review and correct their raw data, as well as preview the center outcomes report and provide commentary.

New Elements for the Statistical Analysis

It was acknowledged that changes in the field may require that additional variables be considered in the model building process. As new elements are considered for inclusion into the model or modification of the statistical analysis plan is proposed, a three-step process should be followed to determine the suitability of modifying the center specific analysis model. First, if new variables are being considered for inclusion in the model, completeness and accuracy of data collection for those data elements should be assessed, assuming that these data elements are captured on current CIBMTR TED or Report Forms. Data elements not collected by the CIBMTR but proposed for data collection should be considered in light of the data collection burden and the important international

implications of insuring harmony between the TED forms of the SCTOD and the Med A forms of the EBMT and other international registries. Second, the inclusion of any additional factor in the center outcomes models should be assessed for its impact on the model, the goodness of fit, and any changes in center assessments. Finally, additional data elements considered important for inclusion in the model and any modifications of the analytic model should be discussed with the community and included only after consideration in the review process for the models outlined above. This process is designed to include broad representation from HCT centers, patients, payers, and the government contractor.

Recommendation:

- New elements for data collection and inclusion in center outcomes analyses should be provisionally tested and subject to the center outcomes analysis review process to preserve parsimony of data collection and quality of center-specific outcomes reports.

Special Subsets: Possible Exclusions to Permit Innovation

Certain subsets of patients have an inherently high risk of post-transplant mortality; however, for many HCT still may offer a benefit compared to other available therapies. Many of these patients undergo HCT on early phase investigational studies testing new approaches to improve outcomes. There are legitimate concerns that these important services at HCT centers could be dampened if they are perceived to compromise centers' outcomes reports and thereby affect business practices. An additional consideration for the model should be determination of whether analyses should exclude, or adjust for very high-risk patients being treated on Phase I transplant studies such as those in advanced or resistant relapse, as a special subset in any center specific outcomes assessment. Accommodation of the model for such high-risk transplants was believed to be important to permit innovative techniques to be applied to such patients' care without adversely impacting a center's assessment score and thus discouraging new Phase I investigations that may lead to advances in HCT techniques. The group acknowledged that such adjustment would be difficult in that accurately identifying patient groups to be accorded "special treatment" might be somewhat arbitrary and subject to gamesmanship.

Recommendation:

- Explore reliable and consistent methods to capture participation in legitimate investigative research for high risk indications that will facilitate adjustment in the center outcomes reports.

Center Practice Profile

The primary goal of the center outcomes reports is to provide centers with tools to assess and improve their performance. An additional tool not required by government contract that might be beneficial to this objective is a confidential yearly assessment (in comparison to prior years) of a transplant center profile that includes graphical depictions of treated patients' ages, diagnoses, donor types, and risk mix. A pediatric-specific profile may also be informative for centers with a sizeable component of pediatric recipients. Changes in a center's profile of patients and transplant techniques over time

may be informative in relation to understanding changes in their center-specific outcomes assessment. Additionally, CIBMTR might include an overview of the overall practice profile for the US centers so that centers can see where their practices differ from the norm.

Recommendation:

- CIBMTR should consider developing a center specific patient-, disease- and transplant-related demographic description report to be delivered annually with each center's outcomes report.

AVOIDING UNINTENDED CONSEQUENCES OF OUTCOMES REPORTING

Accounting for patients on research protocols

Patients who are treated on research protocols often represent "high-risk" patients who would not be transplanted using standard transplant approaches. Research protocols that offer HCT to high risk patients represent a platform for innovation of new transplantation techniques. Therefore, if center outcomes reporting provides a disincentive to care for high risk patients, some patients may not derive benefit from HCT, and new innovation may suffer. Impact of clinical trial participation may be particularly evident at small pediatric centers. Suggestions for handling patients treated on research protocols included creation of exceptions for inclusion in outcomes analyses ("carve-outs"), reporting clinical trial participation as a process measure, or adjusting for clinical trial participation at the level of the patient. Creating center outcomes reports using only "standard" transplant patients was considered unacceptable, in the context of a report whose goal is to capture outcomes for all HCT recipients. The most acceptable approach would be to track participation in clinical trials, and adjust patient characteristics based upon this data item. For this to be effective, and to avoid "gaming", an acceptable benchmark to categorize clinical trials meriting exclusion is necessary.
(See also: *Special Subsets: Possible Exclusions to Permit Innovation*)

Recommendation:

- Collect data on registered clinical trials participation, and include participation in Phase I trials as an adjustment factor in center outcomes report.
- Include the number of patients transplanted on registered on Phase I, II and III clinical trials in the center demographic descriptive reports that accompany the outcomes report.

Adequately characterizing patient's risks

Substantial concern exists in the transplant community that current techniques to adjust for patient case-mix severity are inadequate and will leave centers at risk for inappropriate labeling with low performance. The group acknowledged that limitations exist as to ability to collect information on all potential variables that may affect "risk" for adjustment in modeling systems. Better definitions of disease status and burden at the time of HCT should be incorporated into models as they become available. However, a few factors were mentioned as examples for further strong consideration. Cytogenetic

abnormalities continue to emerge as a powerful predictor of outcome for some diseases, including acute leukemias. As better methods to describe cytogenetic abnormalities become available, particularly data driven methods, they should be incorporated into data collection instruments and disease adjustment models. Socioeconomic factors also represent a strong determinant of outcomes. Current models do not account for socioeconomic status (SES). Methods to approximate SES include collection of income, insurance status, education level, or use of zip code data for extrapolation. Zip code data was believed to represent the data element most reliable and readily collectable for use in center-specific outcomes analyses among these choices.

In considering development and recognition of new prognostic variables over time, standard procedures should be developed to guide the process of transitioning from collecting data for research uses, to provisional data collection and testing, and finally standard data collection and use in outcomes reports. IT methodologies for provisional data collection should be explored to facilitate this process. (See also: *Adherence and Schedule for Model Review*)

Recommendation:

- Develop Standard Operating Procedures for reviewing new candidate prognostic variables, collecting new data provisionally to test feasibility and incorporation of new data items into standard data collection and modeling.
- Begin collection of zip code of residence of patients with next forms iteration for incorporation into subsequent model building. Improve data collection of cytogenetic abnormalities for future inclusion in center outcomes analyses.

Which time period for outcome assessment best accommodates different pre-transplant preparative regimen intensities?

Reduced intensity transplant regimens have generally offered benefit to some patients with risk factors that may preclude myeloablative conditioning, or where immunologic effects are most crucial. They complicate center outcomes analysis since reduced intensity regimens are associated with lower mortality in the first 3 months after the procedure, and a higher frequency of events in the 3-9 month time period. However, there was consensus that in general, outcomes should be comparable across a range of conditioning intensities at 1 year after HCT. CIBMTR should continue to assess temporal differences in outcomes by conditioning intensity. Future analyses may consider modeling outcomes separately based on conditioning intensity, with different outcomes reporting periods. (See also: *Outcomes measure*)

Recommendation:

- Continue to model and report center outcomes based on one year survival.

Should pediatric populations be handled differently?

Pediatric centers present specific challenges, including smaller overall size compared to adult centers, and diseases with different risk factors than those in adults. Smaller size leads to wider confidence intervals surrounding expected survival estimates. In general, current center outcomes methodology was believed to be sufficient based upon available data. Future considerations will include whether the time window for pediatric centers

can be reduced to 3 years from 5 years with inclusion of all allogeneic HCT recipients, and whether adjustment for pediatric populations based upon disease factors currently considered in the model will be sufficient. Further research into pediatric specific risk adjustment will inform future center outcomes analyses. (See also: *Time window for inclusion in analysis*)

Recommendation:

- Continue current methods regarding inclusion of pediatric patients. Methods to deal with small centers will apply to pediatric centers.

Considerations to mitigate or prevent unintended consequences

On the surface, concerns that patients with more high risk factors or those with disadvantaged SES may be marginalized from HCT benefits appear valid. However, these potential consequences are not well documented. CIBMTR may be able to track and measure utilization of HCT in high risk or disadvantaged patient populations nationally, and across geographic regions. It may also consider future center surveys to identify practice pattern changes that result from patient selection in response to center outcomes reports. For instance, tracking of transplantation over time may provide understanding of whether certain conditions become less prevalent among transplanted patients, or whether access to care for certain populations is changing over time. As well, care processes and changes in center practices can be evaluated temporally. Documentation of unintended consequences, if they exist, can be provided to national policymakers. Descriptive demographics should be presented to each center annually and may assist them in understanding practice pattern changes. These practice pattern descriptions could also be normed against national descriptive figures. (See also: *Center Practice Profile*)

Recommendation:

- CIBMTR should make efforts to track and measure HCT utilization over time, in particular for disadvantaged populations in order to determine whether center outcomes reporting affects HCT access.

Can centers be protected from punitive responses to center outcomes reports?

The overall goal of center-specific outcomes reporting is to provide centers with performance measurement tools for use in quality improvement. Outcomes reports must be completed with the best possible scientific methodology, and in a fair, unbiased fashion. However, it must be acknowledged that risk modeling is not perfect and processes must be flexible to introduction of new, accepted and validated risk factors. Results must also be presented in a fair and understandable report in the context of quality improvement.

Recommendation:

- CIBMTR should work closely with government, payers, the ASBMT Quality Outcomes committee and patient representatives to produce an equitable report that accounts for patient case-mix, and is presented in clear language that focuses on quality improvement.

TRANSITIONING FROM REPORTING OUTCOMES OF UNRELATED DONOR HCT TO REPORTING OUTCOMES OF ALL ALLOGENEIC HCT

Most of the discussions of this group focused on the appropriate variables to consider in regression models in order to adjust for risk. These topics were also covered in other groups, so there is some repetition.

Patient related factors to consider in models

Previous center outcomes reports aggregated the co-existing disease questions from the NMDP Report Forms into a single categorical variable to adjust for other morbidities. At the recommendation of the ASBMT Quality Outcomes Committee, CIBMTR added the HCT Comorbidity index[1] to the TED forms for data collection. It was agreed that this index should be used in future center outcomes modeling to adjust for differences in comorbidities present at HCT. Graft source is an essential factor to be included when assessing outcomes and consideration was given to modeling outcome separately for each graft source, rather than using a graft covariate in a single model. However, it was agreed that it could be retained in a single model as a covariate, which has the virtue of simplicity, as long as there was appropriate testing for interactions. The group recommended that single and double cord blood transplants be treated as different categories. Substantive discussion ensued regarding collecting information on patient sociodemographic risk factors. Candidate factors include zip code, reported income, "transplant evaluation risk score" (TERS), type of insurance coverage, full mailing address, social work assessment, and self reported health status. Among these, reported income was deemed too hard to capture consistently and several limitations to adequately defining insurance type were mentioned. As well, TERS, social work assessment and self-reported health status were all considered too difficult or impractical for data staff to capture and report. However, the group agreed that zip code was practical and useful for its geographic determinants as well as its ability to be used to estimate SES. (See also: *Adequately characterizing patient's risks*)

Recommendations:

- Replace current aggregate variable for comorbidity with the HCT CI[1]
- Collect and use zip code of residence as a surrogate for SES.

Disease related factors to consider in models

There was considerable discussion regarding inclusion of cytogenetic data at presentation of acute leukemia in the analytic models. The main advantage is that it has been shown to affect likelihood of relapse/progression and death following conventional therapy. Disadvantages include that it is a more inconsistent predictor of posttransplant survival, is difficult for data staff to report and is sometimes not available to the transplant center. The primary role of cytogenetics in center-specific outcomes analyses would be to distinguish among various patients with acute leukemia in CR1, which is a limited group and therefore should have limited impact on classification of a center's outcomes. Consideration was given to simply collecting cytogenetic risk categories (good, intermediate, poor). However, the limitations of changing classification over time and

subjectivity were believed to outweigh the benefit of this approach. European representatives agreed with the difficulty of collecting cytogenetic data. It was noted that cytogenetic data are being collected on the comprehensive Report Forms, and specific WHO entities are being collected on the TED/MED-A. Data from Report Forms may be used for validation studies (see discussion under *New elements for the statistical analysis*) to consider for future inclusion, especially as better techniques to accurately capture cytogenetic data become available. Sensitivity of lymphoma to chemotherapy (esp. Hodgkin Lymphoma) was discussed; it was recognized that these data are already collected and analyzed.

Recommendation:

- Cytogenetic data, though not yet appropriate for routine inclusion in center specific analyses for patients with AML, should be considered for future analyses, using the process to be developed for provisional data collection and inclusion.
- Lymphomas should be considered using categories that account for disease sensitivity.

Transplant related factors to consider in models

Degree of HLA matching was agreed to be an important determinant of HCT outcomes. It was suggested that detailed matching at the allele level would be the appropriate criteria to use in the center outcomes model, including for related HCTs. This was not considered possible in the context of data collection on the TED forms. Previous reports have used the Weisdorf criteria[2] to define matching using highest available resolution data from Report Forms. The match status on the TED forms, which captures the number of mismatches for mismatched related and unrelated donors, was believed to represent sufficient detail for the center outcomes analyses. Cell dose, which has not been a significant factor in previous models, was also discussed. There was interest in capturing CD34 dosing on the TED forms, but recognition that measurement of CD34 in cell products has a high variability and is not performed for all graft types. An alternative for consideration is the nucleated cell dose. These data are not being collected on the current TED/MED-A forms, but should be considered in the future.

Preparative regimen intensity may vary across related and unrelated HCT, as well as with the number of comorbid conditions. CIBMTR currently captures sufficient detail to account for this, either by adjustment in the models, testing for interactions, or analyzing separate models. Current center outcomes models only apply to first unrelated HCT; however a substantial minority of patients receive multiple transplants, particularly those receiving related donor transplants. Consideration should be given to whether future iterations of the center outcomes report should address transplants beyond the first related or unrelated transplant event. Although CIBMTR captures information regarding multiple events, several issues can be foreseen to complicate analyses of multiple HCT events. Subsequent transplants are generally considered to have less good outcomes, may occur at a center different from the prior event, occur within variable time periods following the initial allogeneic HCT, and, although growing in number, continue to represent a small percentage of total transplants. As well, questions are raised as to whether the unit of analysis is each transplant (and the center at which it is performed) or the patient (overall

experience across all HCT). This issue was not resolved in discussion, and should be addressed in focused research studies. Since outcomes of first allogeneic HCT represent a reasonable measure by which to compare performance across centers, this approach will be continued in the center-specific outcomes analyses. Future models should account for prior HCT (auto or allo); however, the best approach is to be determined.

Recommendation:

- Adjust center-specific outcomes analysis for HLA matching based in data captured on the current (2007) TED forms.
- Cell dose data are being collected on Report Forms and can be considered for future data collection on the TED form, using the process for adding new data elements to data collection instruments and center analyses.
- For now, analyses should be restricted to first transplants.

Should a center practice assessment report be generated for each center?

Consistent with the concept that center outcomes reporting has the goal of quality improvement, it was suggested that CIBMTR develop a report for each center that characterizes its HCT practice patterns along with national practice data. Such a profile can provide details regarding numbers and distribution of patient-, disease- and transplant-related factors specific to each center. Specific items of interest include distribution of ages, racial/ethnic groups, types of HCT and graft sources, GVHD prophylaxis, as well as numbers of subsequent transplants and reason. For instance, for patients with prior autologous HCT, the distribution of subsequent transplants performed for recurrence, new disease, graft failure or a planned sequential approach is of interest to centers. This type of quality assessment tool could represent a "value added" service of the CIBMTR. (See also: *Center Practice Profile*)

Recommendation:

- CIBMTR should consider implementing and distributing an HCT practice pattern report to accompany each center's outcomes report annually.

HOW SHOULD THE RESULTS OF THE CENTER OUTCOMES REPORT BE PRESENTED?

What is the intended audience?

Several important audiences can be reached with the center outcomes report or targeted versions of it. The primary audience is the government, whose goal is to facilitate quality improvement and to inform the public. Payer groups are a secondary user. However the center outcomes report created for the CW Bill Young Cell Transplantation Program has the opportunity to establish a new standard for use by payers and reduce duplication of effort by centers to provide outcomes reports to payers. Current reports are sent to HCT centers, and eventually published, but are not specifically sent to payers. Quality ratings, or "certification" of centers is not the responsibility of the CIBMTR, though center-specific outcomes reports may provide a foundation for payers to make such designations.

Recommendation:

- CIBMTR should consider whether the needs of payers can be met by providing a standardized report, based on the required center outcomes report, that can be distributed to the payer community by HCT centers.

What result should be presented to the public, what should be in the report?
Current center-specific patient outcomes presented on the .gov website include raw numbers of patients transplanted by broad disease and disease status categories, as well as the number alive at one year. This may represent insufficient detail to truly inform readers regarding center performance.

Patients generally are interested in understanding whether people like them (disease, age, risk factors) are cared for at any given center and, if so, what their outcomes are at the center. Their need is to understand "Is this center able to care for a patient like me well?" This may be contrary to the current layout of the center outcomes report, which focuses on aggregation of transplants at a center to arrive at a single result of observed performance compared to expectations. This also presents difficulty, as portrayal of risk would require that patient users be able to understand and self-assign to a relative risk group. The contrary perspective is that procedure experience is more relevant than disease experience once a patient requires a transplant. It was also noted that patients are interested in outcomes beyond one year; however, it was acknowledged that outcomes beyond one year may not be fully attributable to the HCT center.

Recommendation:

- Data regarding centers' overall performance, compared statistically to the expected outcome, should be presented as the primary outcome.
- Additional information should be provided in the public section of the report which provides details regarding numbers of patients with particular diseases and disease states, as well as specific risk factors to allow patients to better understand how a center's experience relates to someone like themselves.

Should more than one version of the center outcomes report be generated?
Since there is more than one end user of the outcomes reports (e.g. transplant center, payer, patient, referring physician), the needs of users may be best met by multiple reports. The best approach to this dilemma may be to present a multi-layered report. However, there was consensus that there be one report with the primary content (one year survival with indicator of deviation from expected range), with sub-reports that meet specific needs of other users. For instance, a sub-report for patients should provide explanations of the one year survival and expected survival, with explanations appropriate for a lay audience and transplant volume by disease. The primary layer could contain summary information about the center – its risk adjusted outcomes report in general, its risk score and whether its aggregated outcomes are within the expected statistical range. Subsequent "layers" may contain descriptive data that are more useful for patients, such as number of patients by disease and disease category, and types of HCT that are performed at any given center. It might also contain breakdown of risk as well as tools to assist patients who wish to understand risk categorization. *The patient*

layer should be seen in the context of a tool that facilitates discussion between a patient and their physician at the center. The need for adequate sample size to generate risk adjustment precludes the ability to produce a report at each center on outcomes by disease indication or transplant technique. One limitation of a layered report is the difficulty in displaying them on paper compared to interactive web format.

Recommendation:

- Aside from presentation of the primary outcome as described above, a multi-layered approach may allow for additional descriptive data useful to patients to be presented.

How should the report be presented to achieve greatest comprehension?

Center outcomes reporting is complicated, especially considering factors like statistical modeling, risk adjustment, the large number of variables contributing to outcome and the sophisticated nature of the procedure itself. It is important that, wherever possible, portions of the report intended for non-transplant personnel are targeted towards comprehension at the level of patients. One technique of presenting the data is to use a check box system that simplifies the overall status of centers into graphical symbols that are easily understood. A graphic representation of average risk could also be developed. Consumer Reports ratings systems were discussed as an example. These graphics could be layered in front of numerical details, and accompanied by descriptive explanations for those with more interest. For instance, in an interactive format, a patient could answer a series of questions that provides them their own risk score. Comprehension is more critical than aiming towards a specific reading level; focus groups of patients could be used to target methods of presenting the data.

Although many ideas centered on presenting web-based interactive formats, particularly due to their flexibility, an important consideration is that many potential users may not have internet access. And, many users may wish to print out reports for future reading, or discussion with their provider. Therefore, content posted on a web-site should be designed so that when requested, print format reports can be generated that reproduce the web content in an appropriate paper report.

Recommendation:

- Center outcomes reports should be modified, perhaps using input from targeted patient interviews, so that graphical/symbolic representations can be added to facilitate improved comprehension by patient and other lay groups.

Should center comments about their report be solicited, and should they be displayed in public version of report?

It was generally agreed that centers should have the opportunity to comment upon their results as portrayed in the center outcomes report, and that these comments be displayed in the report. These comments would not be edited by the SCTOD. The timeline for center's review of the reports should include time for centers to generate any comments for inclusion. (see also: *Yearly data submission, review and correction*)

Recommendation:

- Transplant centers should have the opportunity to comment upon their outcomes report, during the appropriate review period leading up to the actual completion.

Should center outcomes reports be used to target centers for intervention?

Inevitably, some centers will have performance that is less than expected after adjustment for risk. The principle objective of measuring center performance is to provide a tool for quality improvement. Several incentives exist for centers to undertake responses to the report. First, they may respond to awareness of performance and make changes based on their own desire to improve outcomes. Second, concern regarding patient awareness may drive improvement. Third, payer-driven processes may lead to practice changes at centers, due to contractual, reimbursement or status designation (center of excellence) concern. This could include government payers. Fourth, an honest broker body, such as FACT, may wish to include center outcomes and quality improvement initiatives based upon the outcomes report more integrally into its accreditation process. Finally, donor networks, such as the NMDP, may use the outcomes report to learn more about under-performing centers on behalf of the interest of donors. Although under-performers are likely to be the focus of interventions, centers with excellent outcomes represent a group of interest since there may be factors that are applicable across centers to improve outcomes.

However, it is the obligation of the CIBMTR to produce an equitable, scientifically valid center outcomes report, recognizing the downstream uses of such a report. Targeted intervention is not a requirement of the contract to generate the report for HRSA.

Recommendation:

- The purpose of center outcomes reports is to provide performance measurement tools by which centers may design quality improvement initiatives. The process of generating and presenting center outcomes reports may drive quality improvement interventions. However, CIBMTR should focus on generating high quality, scientifically valid and unbiased reports.

Appendix A: Attendees of Center-Specific Outcomes Analysis Forum

Name	Organization	Representation
Claudio Anasetti, MD	H. Lee Moffitt Cancer Center	ASBMT / HCT Center
Helen Heslop, MD	Baylor College of Medicine	ASBMT / HCT Center
Bob Krawisz, MBA	ASBMT	ASBMT / HCT Center
Alan Leahigh	ASBMT	ASBMT / HCT Center
Robert Soiffer, MD	Dana-Farber Cancer Institute	ASBMT / HCT Center
Peggy Appel, MHA	Northwest Marrow Transplant Program	ASBMT Quality Outcomes Committee
Roy Jones, MD, PhD	M.D. Anderson Cancer Center	ASBMT Quality Outcomes Committee
Jan Sirilla, RN, MSN	OSU AG James Cancer Hospital	ASBMT Quality Outcomes Committee
Patrick Stiff, MD	Loyola University Medical Center	ASBMT Quality Outcomes Committee
Mary Eapen, MD	CIBMTR/MCW	CIBMTR
Mary Horowitz, MD, MS	CIBMTR/MCW	CIBMTR
Navneet Majhail, MD, MS	NMDP	CIBMTR
J. Douglas Rizzo, MD, MS	CIBMTR/MCW	CIBMTR
Daniel Weisdorf, MD	University of Minnesota	CIBMTR
Stella Davies, MBBS, PhD	Cincinnati Children's Hospital Medical Center	CIBMTR / HCT Center
Stephanie Lee, MD, MPH	Fred Hutchinson Cancer Research Center	CIBMTR / HCT Center
Alois Gratwohl, MD	University Hospital Basel (Switzerland)	EBMT
Per Ljungman, MD, PhD	Karolinska University Hospital (Sweden)	EBMT
Robert Baitty, MPP	HRSA	Government Agency
James Burdick, MD	HRSA	Government Agency
Randy Gale, MPH, MT	HRSA	Government Agency
Linda Griffith, MD, PhD	NIAID	Government Agency
Roy Wu, PhD	NCI	Government Agency
Lois Ayash, MD	Karmanos Cancer Institute	HCT Center
Terri Halverson, RN	Children's Memorial Hospital	HCT Center
H. Kent Holland, MD	BMT Group of Georgia (Northside Atlanta)	HCT Center
Mitchell Horwitz, MD	Duke University BM&SCT Program	HCT Center
Michael Pulsipher, MD	University of Utah	HCT Center
Mohammed Sorror, MD	Fred Hutchinson Cancer Research Center	HCT Center
Douglas Taylor, MD, PhD	University of California-Davis	HCT Center
Tim Hofer, MD	University of Michigan	Hospital Outcomes Reporting
Jeanne McGee, PhD	McGee & Evers Consulting, Inc.	Hospital Outcomes Reporting
Dana Richardson, RN, MHA	Wisconsin Hospital Association	Hospital Outcomes Reporting
Patrick S. Romano, MD, MPH	University of California-Davis	Hospital Outcomes Reporting
John Klein, PhD	CIBMTR/MCW	MCW-Biostatistics
Brent Logan, PhD	CIBMTR/MCW	MCW-Biostatistics
Elizabeth (Beth) Murphy, EdD, RN	NMDP	NMDP
Carol Doleysh, BS, CPA	CIBMTR/MCW	Other Attendees
Waleska Perez, MPH	CIBMTR/MCW	Other Attendees
Kathleen Sobocinski, MS	CIBMTR/MCW	Other Attendees
Patty Steinert, MBA	CIBMTR/MCW	Other Attendees
D'Etta Waldoch Benson, CMP	CIBMTR/MCW	Other Attendees
Paula Watry, RN, PA-C	CIBMTR/MCW	Other Attendees
Becky Lewis, JD	NMDP-Patient Services Committee	Patient Advocate
Barry Schatz	NMDP-Patient Services Committee	Patient Advocate
Carole Flamm, MD, MPH	Blue Cross Blue Shield Association	Payer Group

Name	Organization	Representation
Dennis Irwin, MD	United Resource Networks	Payer Group
Wendy Marinkovich, RN, BSN, MPH	Blue Cross Blue Shield Association	Payer Group
Walter Graham, JD	United Network for Organ Sharing	Solid Organ Program
Randall Sung, MD	SRTR / University of Michigan	Solid Organ Program
Shelly Carter, ScD	EMMES Corporation	Statistical Consultant
Nancy Flournoy, PhD	University of Missouri-Columbia	Statistical Consultant
Ted Gooley, PhD	Fred Hutchinson Cancer Research Center	Statistical Consultant
Glen Heller, PhD	Memorial Sloan-Kettering Cancer Center	Statistical Consultant
Jack Kalbfleisch, PhD	University of Michigan	Statistical Consultant
Chap T. Le, PhD	University of Minnesota	Statistical Consultant
Robert Wolfe, PhD	Scientific Registry of Transplant Recipients	Statistical Consultant

References:

[1] Sorror ML, Maris MB, Storb R, Baron F, Sandmaier BM, Maloney DG, Storer B. Hematopoietic cell transplantation (HCT)-specific comorbidity index: a new tool for risk assessment before allogeneic HCT. Blood 106:2912-2919, 2008.

[2] Weisdorf D, Spellman S, Haagenson M, Horowitz M, Lee S, Anasetti C, Setterholm M, Drexler R, Maiers M, King R, Confer D, Klein J. Classification of HLA-matching for retrospective analysis of unrelated donor transplantation: revised definitions to predict survival. Biol Blood Marrow Transplant 14:748-758, 2008.