How I assess comorbidities prior to hematopoietic cell transplantation

Mohamed L. Sorror
How I Assess Comorbidities Prior to Hematopoietic Cell Transplantation

Mohamed L. Sorror

1Clinical Research Division, Fred Hutchinson Cancer Research Center; and 2Division of Medical Oncology, Department of Medicine, University of Washington School of Medicine, Seattle, WA

Correspondence to:
Mohamed L. Sorror, M.D.  Telephone: 206-667-2765
Clinical Research Division (D1-100)  Fax 206-667-6124
Fred Hutchinson Cancer Research Center  Email: msorror@fhcrc.org
1100 Fairview Avenue North
Seattle, WA 98109-1024
ABSTRACT

The hematopoietic cell transplantation-comorbidity index (HCT-CI) is a comorbidity tool suited for recipients of HCT. The index has been shown to sensitively capture the prevalence and magnitude of severity of various organ impairments before HCT and to provide valuable prognostic information after HCT. Many investigators have validated the discriminative power of the HCT-CI, but others have not. One concern is the consistency in comorbidity coding across different evaluators, particularly in view of the relatively recent addition of the HCT-CI to the transplant-evaluation process. In this article, comorbidity scoring was tested across different evaluators, and only a fair inter-observer agreement rate could be detected. In order to address these issues, a brief training program is proposed here consisting of systematic methodology for data acquisition and consistent guidelines for comorbidity coding that were summarized in a web-based calculator. In a validation patient cohort, this training program was shown to improve the inter-evaluator agreement on HCT-CI scores to an excellent rate with weighted kappa values in the range of 0.89–0.97. This proposed training program will facilitate reliable assessment of comorbidities in the clinic and for research studies leading to standardization of the use of comorbidities in prediction of HCT outcomes.
INTRODUCTION

Organ dysfunctions (comorbidities) were found to be associated with the outcome of treatment of a given primary disease\(^1,2\) and, in particular, cancer.\(^3,4\) In 2005, a hematopoietic cell transplantation-comorbidity index (HCT-CI) was introduced as a measure of organ dysfunction that was suited for recipients of HCT.\(^5\) The HCT-CI was developed from the historical Charlson comorbidity index (CCI)\(^6\) after introducing three conceptual changes: the use of laboratory and organ function tests to redefine pulmonary, hepatic, cardiac, and renal comorbidities; the inclusion of all comorbidities encountered in a cohort of HCT recipients at a single institution; and the estimation of new adjusted hazard ratios for the associations between comorbidities and non-relapse mortality after HCT. These adjusted hazard ratios were then converted into weights that could be summated into a total score.

In validation cohorts of recipients of allogeneic HCT from two different institutions, the HCT-CI was demonstrated to have higher discriminative power than the CCI both for NRM and overall survival.\(^5,7\) Many investigators reported on the valid association between HCT-CI scores and mortality in their respective single-institutions,\(^8-17\) while few others disagreed.\(^18-22\) A discussion of the possible reasons for the lack of complete agreement by investigators on the validity of the HCT-CI is outside the scope of this article. Instead, this article is focused on a single concern that is related to the degree of consistency in assigning comorbidity scores among evaluators. For example, a recent study reported a noticeably higher prevalence of comorbidities compared to other reports.\(^22,23\) As investigators continue to explore the validity of the HCT-CI and to use it in decision-making and prognostication studies, an urgent need has emerged to standardize the methods and guidelines for comorbidity evaluation. A valid and reliable system for comorbidity evaluation would not only ensure the calculation of an accurate total comorbidity score but would also allow the accurate estimation of the prevalence of individual comorbidities, which would be of prime importance in future research addressing roles
of comorbidities in post-HCT complications. Here, a brief training program is proposed consisting of consistent methods for data acquisition from medical records and detailed guidelines for comorbidity assessment that were summarized in a web-based application and a calculator. Validation of the ability of the proposed training program to improve the inter-rater reliability (IRR) of the HCT-CI was described.

The Training Program

I. Methods of Retrospective Assessment of Medical Records for Acquisition of Comorbidity Data

Transplant physicians and physician assistants are conceivably familiarized with the use of well-established measures such as the Karnofsky scale for assessment of performance status\(^{24}\) and the systems used for grading acute GVHD\(^{25,26}\). However, the introduction of new evaluation scales even for a familiar clinical condition, for example chronic GVHD, often requires the development of systematic methodology that could ensure a stepwise pattern of assessment\(^{27}\). Such methodology should not only save time by avoiding back-tracking and duplication of efforts but would also endorse the consistency in the acquisition of data and the evaluation of a given medical condition. The HCT-CI is a distinct assessment tool that is relatively new in the transplant field. Clinicians and study coordinators with varying degrees of experience could complete a comorbidity evaluation in approximately 15 minutes by following the proposed three-step evaluation process (Figure 1).

The landmark date is defined here as the day before HCT that should not to be exceeded for any of the comorbidity-related evaluations. Day -10 was chosen to be the landmark date because all the conditioning regimens for patients who contributed to the development of the HCT-CI\(^\text{5}\) started after that day. As long as medical events and laboratory or organ function tests were done on or before the landmark date (day -10), they are suitable for
the purpose of comorbidity coding. The use of a fixed time point (landmark date) for all evaluations would facilitate the retrieval of laboratory data from computer databases and the use of a web-based calculator (see below), and it would also standardize data collection across institutions. In the circumstance that a patient is given a conditioning regimen that starts before day -10, an evaluator could use the day before the start of conditioning regimen as the landmark date for assessment of comorbidities in that patient.

The following are the steps for the evaluation process:

1. Reading the important sections of the medical records:
   a. Evaluate the nutrition notes to capture the measures of weight and height that were assessed at the closest time point before the landmark date and then to calculate the body mass index (BMI).
   b. Assess the “History and Physical Examination” (H&P) note: it is the note dictated by the transplant physician or physician assistant within a few weeks before the landmark date. This note should include evaluations of: i) the patient’s present, past, social, and family history, ii) review of organ systems, and iii) physical examination of the patient, before requesting the pre-transplant evaluation of laboratory and organ function tests. An evaluator should pay close attention to three major parts in the H&P note:
      i. The past medical history for all details of recent and remote organ dysfunctions.
      ii. The current medication list to aid in the evaluation of some comorbidities and to detect others that might have been inadvertently dropped from the past medical history (for example anti-depressant medications for depression or oral hypoglycemic drugs for diabetes mellitus).
      iii. The final assessment summary for additional details on organ dysfunctions and for information on any planned consults/evaluations.
iv. The H&P note could also be used as a good source for data on other important
   prognostic variables such as prior treatment and performance status scores.

   If the H&P note is exceptionally deficient in details on specific comorbidities, an
evaluator should search for and review any prior notes on organ-specific problems (for
example gastroenterology consult notes on inflammatory bowel disease or a previous
H&P summarized by the patient’s primary oncologist or general medical practitioner) to
confirm the diagnosis of a given comorbidity and to determine whether a specific
treatment was given or not.

c. Examine the note summarized by the transplant physician that describes findings of the
pre-transplant evaluation. In some institutions, this note would also include details on
consenting for clinical trials. This note could be referred to as the “review of data” note.
Other institutions might have this document in a different format. For example, findings
of the pre-transplant evaluations could be summarized in an updated H&P note before
the start of conditioning regimen. It is anticipated that the pre-transplant evaluation
period generally spans 2-3 weeks before the start of the conditioning regimen.

The “review of data” note, most frequently, is an abundant source for:

i. The most recent laboratory data.

ii. Organ function tests.

iii. Finalized recommendations from any requested consults.

iv. Current status and staging of the primary disease.

d. Review notes on any requested consults, for example, a psychiatric consult for
assessment of depression or anxiety, during the pre-transplant evaluation period for:

i. Assessment of severity of a given comorbidity.

ii. Any recommended treatment specific for a given comorbidity.
2. Review of laboratory and organ function tests:
   a. Assess the most updated report of pulmonary function test (PFT) before the landmark date. The PFT report contains details on:
      i. The percentage of measured-to-predicted forced expiratory volume in one second (FEV1, Figure 2).
      ii. The percentage of measured-to-predicted diffusion capacity of carbon monoxide (DLco) after correction for hemoglobin (Figure 2).
   b. Evaluate the echocardiogram or the multi-gated acquisition (MUGA) scan report for:
      i. The percentage of ejection fraction (EF) for adults or shortening fraction (SF) for children.
      ii. Details on the presence and magnitude of severity of any valve abnormality.
      iii. Details on other cardiac comorbidity (for example, dilated cardiomyopathy).
   c. Assess liver function tests between days -24 and -10 (or between days -40 and -10 if only a single value is reported between days -24 and -10) before HCT for values (Figure 3) of:
      i. Alanine aminotransferase (ALT).
      ii. Aspartate aminotransferase (AST).
      iii. Total bilirubin.
   d. Assess serum creatinine values between days -24 and -10 (or between days -40 and -10 if only a single value is reported between days -24 and -10) before HCT.

3. Summary and final assessment:
   a. Double checking:
      i. As an evaluator goes through the different sections of the medical record, it would be helpful to enter each positive finding momentarily in a software spreadsheet tool such as MS Excel, a web-based calculator (see below), or simply by pen and paper.
ii. Once all eight subsections of the two major compartments of data acquisition (Figure 1) are completed

1. Double-check all the positive findings listed in the calculator or the sheet.

2. Fix any incorrect data entry that is not fitting the patient overall presentation as at that time an evaluator would have recent recollection of details of the medical record and quick access to the chart to verify any information.

b. Calculate a total score and assign it to the patient chart

II. Guidelines for Assessment of Comorbidities in the HCT-CI (the Comorbidity Coding Tool)

All clinical and laboratory criteria described in this coding tool are meant for the evaluation of comorbidities specifically per the HCT-CI (Table 1).

1. Arrhythmia (Score 1):

A score of 1 is assigned for any type of arrhythmia that has necessitated the delivery of a specific anti-arrhythmia treatment at any time point in the patient’s past medical history. Examples include atrial fibrillation or flutter, sick sinus syndrome, or ventricular arrhythmias. A score is assigned even if the patient was in normal sinus rhythm at the time of data acquisition or at the landmark date. No score is assigned to transient arrhythmias that never required treatment.

Sometimes, the medical record does not include enough details on the treatment of a prior arrhythmia. In this case, judgment should be based on the clinical significance of the described arrhythmia and the clinical situation accompanying the development and resolution of such arrhythmia. For example, a patient who developed a rapid atrial fibrillation requiring admission and management in the intensive care unit is assigned a score for arrhythmia even if the medical record does not state the type, dose, and duration of treatment given. However, if
the clinical situation after careful review of the medical record raises doubt as to whether or not a treatment was given for an arrhythmia, no score is assigned. An example could be a patient who developed once a paroxysmal atrial fibrillation, with no indication of a rapid ventricular response, that resolved spontaneously with no mention of the use of a specific anti-arrhythmia treatment.

2. Cardiovascular Comorbidity (Score 1):

A maximal score of 1 is assigned for cardiovascular comorbidity in the presence of one or more of the following three clinical presentations.

- **Coronary artery disease**: This is based on the presence of a documented diagnosis of chronic exertional angina, unstable angina, or myocardial infarction at any time point in the patient’s past medical history as stated in the H&P section of the medical record. Information on prior placement of a coronary stent or undergoing a coronary artery bypass graft surgery should support coding this comorbidity.

- **Congestive heart failure (CHF)**: In order to score this clinical presentation, the medical record should have a statement about the development of symptoms/signs of CHF (for example, an exertional or paroxysmal nocturnal dyspnea) that later responded to diuretics, afterload-reducing agents, beta blocker, and/or digitalis at any time point in the patient’s past medical history.

- **Low ejection fraction (EF)**: patients with an EF of \( \leq 50\% \) or a SF (for pediatric patients) of \( \leq 26\% \) by an echocardiogram or a MUGA scan are assigned a score of 1. As noted before, evaluation of this comorbidity item should be restricted to the most recent measurements of EF or SF before the landmark date. Lack of evaluation of EF or SF for an individual patient before transplant does not preclude the calculation of a total HCT-CI score for that patient.
3. **Inflammatory Bowel Disease (Score 1):**

A score of 1 is assigned for this comorbidity based on the presence of a documented prior diagnosis (history of an endoscopic examination of the mucosa with or without confirmatory histology and radiologic findings) of Crohn’s disease or ulcerative colitis requiring treatment at any time point in the patient’s past medical history. If the patient has never received a treatment for this comorbidity, no score is assigned.

4. **Diabetes (Score 1):**

A score of 1 is assigned for this comorbidity based on the diagnosis of diabetes or steroid-induced hyperglycemia requiring continuous treatment with insulin or oral hypoglycemic drugs during the instantaneous period of 4 weeks before the landmark date. No score is assigned for this comorbidity if diabetes could be controlled with diet alone or if a previous treatment for diabetes or steroid-induced hyperglycemia was stopped 4 weeks before the landmark date.

5. **Cerebro-Vascular Disease (Score 1):**

A score of 1 is assigned for cerebro-vascular disease based on a prior diagnosis of transient ischemic attack, subarachnoid hemorrhage, or cerebral thrombosis, embolism, or hemorrhage at any time point in the past medical history. No details on treatment are required for assigning a score for this comorbidity.

6. **Psychiatric Disorder (Score 1):**

A score of 1 is assigned for this comorbidity based on the presence of any mood, anxiety, or other psychiatric disorder requiring continuous treatment during the instantaneous period of 4 weeks before the landmark date. Depression and anxiety are the most common psychiatric disorders encountered in transplant populations, yet other disorders such as schizophrenia or bipolar disorder should also be coded for this comorbidity. Patients who are receiving only "as
needed” medications for any of the above disorders are not assigned a score for this comorbidity.

7. Hepatic Comorbidity (two levels of severity):

As a general rule, assessment of the laboratory tests (a and/or b) has to include at least two values per test on two different days within a period extending between days -24 and -10 before HCT (Figure 3). That period could be extended to be between days -40 and -10 only in the case that liver function tests were done only once between days -24 and -10 before HCT. The laboratory value closest to the landmark date should be the value used in defining the severity of hepatic comorbidity (Figure 3). The upper limit of normal (ULN), for any of the three tests, is determined based on the reference range per the institution laboratory.

**A) Mild Hepatic Comorbidity (Score 1):**

A maximal score of 1 is assigned for mild hepatic comorbidity in the presence of one or more of the following three clinical presentations: a) Elevated total bilirubin to a value higher than the ULN and up to 1.5 times the ULN; b) elevated values of any of the two hepatic transaminase enzymes, ALT or AST, to values higher than the ULN and up to 2.5 times the ULN; or c) a prior diagnosis of an infection with hepatitis B or C at any time point in the patient’s past medical history before the landmark date.

**B) Moderate-Severe hepatic comorbidity (Score 3):**

A maximal score of 3 is assigned for moderate/severe hepatic comorbidity in the presence of one or more of the following three clinical presentations: a) Elevated values of total bilirubin to a level higher than 1.5 times the ULN; b) elevated values of any or both of the two hepatic transaminase enzymes to levels higher than 2.5 times the ULN; or c) A documented
diagnosis of liver cirrhosis at any time point in the patient’s past medical history before the landmark date.

8. Obesity (Score 1):
A score of 1 is assigned for obesity based on a BMI of >35.00 for patients older than 18 years or a BMI-for-age of ≥ 95th percentile for patients of ≤ 18 years of age. Evaluation of this comorbidity is based on the most recent measurement of the BMI (or weight and height needed for the calculation of the BMI) before the landmark date.

9. Infection (Score 1):
A maximal score of 1 is assigned for infection comorbidity in the presence of one or more of the following four clinical presentations: a) a documented infection (for example, by culture or biopsy); b) fever of unknown origin; c) pulmonary nodules suspicious for fungal pneumonia; or d) a positive purified protein derivative test requiring prophylaxis against tuberculosis. Patients must have started a specific anti-microbial treatment before the landmark date with a recommendation, documented in the chart either by the primary team or the infection consult team, to continue the same anti-microbial therapy (or a similar agent) during the days of administration of conditioning regimen and beyond day 0 of HCT.

10. Rheumatologic Comorbidity (Score 2):
A score of 2 is assigned for this comorbidity based on the presence of a documented prior diagnosis of a rheumatologic disease which has required administration of a specific treatment at any time point in the patient’s past medical history. Diagnoses include systemic rheumatologic and connective tissue disorders such as systemic lupus erythematosis, rheumatoid arthritis, Sjögren’s syndrome, scleroderma, polymyositis, dermatomyositis, mixed connective tissues disease, polymyalgia rheumatica, polychondritis, sarcoidosis and vasculitis
syndromes. Patients with undiagnosed polyarthritis, degenerative joint disease, or osteoarthritis are not scored for this comorbidity. Occasionally, a patient might have a clinical pattern of a systemic rheumatologic disease responding to a specific treatment but without a definitive diagnosis. For example, I was consulted once on a patient with an unspecified collagen vascular disease which had presented 4 years earlier and had manifested by iritis, uveitis, bowel disturbances and muscle aches. This unspecified collagen vascular disease responded to low-dose systemic steroids. Even though there was no definitive rheumatologic diagnosis in this case, I erred to the cautious side and assigned a score for this presentation as a rheumatologic comorbidity.

Patients with quiescent rheumatologic diseases who are receiving no treatment in the immediate period before the landmark date are assigned a score for this comorbidity if they have fulfilled the prior criteria.

11. Peptic Ulcer (Score 2):
A score of 2 is assigned for peptic ulcer based on the presence of a prior endoscopic or radiologic diagnosis of gastric or duodenal ulcer, noted in the medical record, at any time point in the patient’s past medical history. Patients with quiescent peptic ulcer who are receiving no treatment in the immediate period before the landmark date are assigned a score for this comorbidity if they have met the prior criteria.

12. Renal Comorbidity (Score 2):
A maximal score of 2 is assigned for renal comorbidity in the presence of one or more of the following three clinical presentations: a) elevated values of serum creatinine to >2 mg/dl or >176.8 umol/L, as detected in at least two laboratory tests on two different days within a period extending between days -24 and -10 before HCT. This evaluation period could be extended to span the time between days -40 and -10 if serum creatinine was evaluated only once between
days -24 and -10 before HCT; b) chronic renal disease requiring weekly dialysis within the instantaneous period of 4 weeks before the landmark date; or c) a documented prior history of renal transplantation at any time point in the patient’s past medical history.

13. Pulmonary Comorbidity (two levels of severity):

As a general rule, assessment of pulmonary comorbidity for the purpose of assigning HCT-CI scores should exclusively rely on PFT results, in particular corrected DLco and FEV1 percentages (Figure 2). A total HCT-CI score should not be calculated in the absence of data on PFT except in the case that PFT could not be done due to technical difficulties, for example, in pediatric patients. Occasionally, patients are assessed by a post-bronchodilator (reversibility) test. In this case, only the pre-bronchodilator values of FEV1 are considered for evaluation of pulmonary comorbidity.

Measured DLco values should first be corrected for the concurrent hemoglobin value using the Dinakara equation (Corrected DLco = uncorrected DLco / (0.06965 × Hemoglobin g/dl)). Then, the corrected value of measured DLco is divided by the predicted value to compute the percentage of DLco. Alternatively, the uncorrected DLco percentage, which is reported in all PFT reports, could be directly corrected for the concurrent hemoglobin value using the Dinakara equation to compute the corrected DLco percentage (Figure 2). Either way will lead to the same final percentage of corrected DLco. The Dinakara equation is favored over other equations such as the one by Cotes et al. because of its more robust ability to account for the effects of anemia, a common sign of the primary hematologic disease, and because it is the equation used by the PFT laboratory at Fred Hutchinson Cancer Research Center (FHCRC) where the HCT-CI was originally developed.

A) Moderate Pulmonary Comorbidity (Score 2):

A maximal score of 2 is assigned for moderate pulmonary comorbidity in the presence of one or more of the following three clinical presentations: a) a percentage of DLco in the range of
66-80% or b) a percentage of FEV1 in the range of 66-80%; both should be the most recent measurements before the landmark date; or c) shortness of breath on slight activity that is attributed to a pulmonary disease and cannot be corrected by blood transfusion for a noticeable anemia as assessed during a clinic visit within the immediate period of two weeks before the landmark date.

B) Severe Pulmonary Comorbidity (Score 3):

A maximal score of 3 is assigned for severe pulmonary comorbidity in the presence of one or more of the following four clinical presentations: a) a percentage of DLco of ≤ 65%; b) a percentage of FEV1 of ≤ 65%; both should be the most recent measurements before the landmark date; c) shortness of breath at rest that is attributed to a pulmonary disease and cannot be corrected by blood transfusion for a noticeable anemia as assessed during a clinic visit within the immediate period of two weeks before the landmark date; or d) the need for intermittent or continuous oxygen supplementation during the immediate period of 4 weeks before the landmark date.

14. Prior Malignancy (Score 3):

A score of 3 is assigned for this comorbidity based on the presence of a prior diagnosis of any solid tumor that required receiving a specific treatment at any time point in the patient’s past medical history regardless of the type of treatment (surgery, radiotherapy, and/or drug therapy). Lymphomas or myelomas that preceded the diagnosis of a myeloid malignancy (for example, acute myeloid leukemia [AML], myelodysplastic syndromes [MDS], or chronic myeloid leukemia [CML]) are assigned a score for this comorbidity. Similarly, myeloid malignancies that preceded the diagnosis of lymphomas or myelomas are assigned a score for this comorbidity. Patients with a prior malignancy from the same lineage of cells of the current malignancy should not be assigned a score for this comorbidity: for example, if a patient had a diagnosis of a Non-Hodgkin that was preceded by a Hodgkin lymphoma or if a patient had a diagnosis of AML that was
preceded by MDS. Melanoma, but not basal or squamous cell carcinoma of the skin, should be assigned a score for this comorbidity. Patients with a prior malignancy that never required a specific treatment are not scored for this comorbidity. Tumors of benign nature are not scored for this comorbidity.

15. Heart Valve Disease (Score 3):

A maximal score of 3 is assigned for heart valve comorbidity in the presence of one or more of the following three clinical presentations: 1) at least a moderate or severe degree of valve stenosis or insufficiency, as determined by echocardiogram, whether that valve was mitral, aortic, tricuspid, or pulmonary; b) prosthetic mitral or aortic valve; or c) symptomatic mitral valve prolapse. Assessment of this comorbidity is limited to the most recent heart evaluation by echocardiogram before the landmark date.

Use of the Guidelines for Prospective Assessment of Comorbidities

The guidelines described in the previous section were meant for retrospective evaluation of comorbidities when the patient has already passed the landmark date of day -10. Retrospective evaluation of comorbidities could be used for prognostic studies or in comparative effectiveness research about the HCT. The HCT-CI could also be evaluated prospectively by clinicians or study coordinators before day -10 for the purpose of risk-benefit assessment, for example at the time of a transplantation consult in order to aid in the decision-making about the intensity of conditioning regimen. In that situation, the same previous guidelines would apply with a change in the landmark date to be the date of the consult. Similarly, investigators assessing comorbidities using the HCT-CI before conventional therapeutic interventions, such as induction chemotherapy for AML, could use the date of comorbidity assessment as the landmark date.

An example: A transplant physician is seeing a patient for a consult on 12/30/2012. Then, the landmark date will be 12/30/2012. The physician should use the values of hepatic
function tests that are done between 12/16/2012 and 12/30/2012 to assess hepatic comorbidity. If only one value of bilirubin is available during this time interval, then the assessment period can be extended to be between 12/01/2012 and 12/30/2012. Also, this patient will have to be continuously treated with an anti-diabetic or an anti-psychiatric treatment for 4 weeks between 12/3/2012 and 12/30/2012 for a score to be assigned for diabetes or psychiatric comorbidity, respectively. If the patient has a diagnosis of an infection before the consult date, it can only be scored as a comorbidity if the prescribed anti-microbial medication is required to be continued for >10 days (until 01/10/2013) after the date of the consult (12/30/2012).

III. The Web-Based HCT-CI Score Calculator

The above explanatory guidelines were adapted into a web-based application and calculator: hctci.org. Evaluators are assigned a registered password-protected access to the web-site, where they can save portions of a patient’s comorbidity data, as they become available in the clinic or during chart review, under a de-identified patient-specific code. Evaluators can access stored data until all comorbidity data are collected and a total score is requested. The calculator contains 15 categories of comorbidities per the HCT-CI (pulmonary and hepatic comorbidities entail two grades of severity), and under each category there are several choices for different clinical presentations. The evaluator is requested to enter the following information: 1) date of the transplant; 2) measures of weight and height to calculate the BMI; 3) two values for each of the laboratory tests for AST, ALT, bilirubin, and creatinine with the corresponding date for each value; 4) percentages for EF, SF, FEV1 and uncorrected DLco with the concurrent hemoglobin value. Options are available if PFT were not done because of younger age; if EF or SF were not done; or if an extended evaluation period (between days -40 and -10) is needed for some laboratory tests. Then, the web-based application will perform the following actions: 1) it will calculate the BMI and determine whether or not a score should be assigned for obesity; 2) it will provide the corrected DLco percentage and will determine, based on percentages of both DLco
and FEV1, whether or not and which score (1 or 3) should be assigned for pulmonary comorbidity; 3) it will determine the score to be assigned for hepatic and renal comorbidities based on the laboratory values and their dates; 4) it will assign scores for other comorbidities based on the selected clinical situations. Finally, a total score could be generated. The web-based application will also provide a summary of all positive comorbidities for a given patient.

If a patient is being evaluated prospectively for comorbidities at a stage preceding the pre-transplant evaluation period (for example during an early consult for HCT), then the evaluator should substitute the date of transplant in the web-based application with a hypothetical date that is 10 days after the date of the consult. Finally, the web-based calculator was tested and validated several times by the PI and the comorbidity evaluation team.

Assessment of the IRR Rates for Validation of the Proposed Training Program

Patients and Methods

Assessment of the IRR was done using data from randomly selected samples of patients who received allogeneic or autologous HCT at FHCRC. This retrospective study was approved by the Internal Review Board of FHCRC and conducted in accordance with the Declaration of Helsinki. Other than the Principal Investigator (PI, MLS), none of the evaluators had any prior experience in evaluating comorbidities, and they had either limited or no experience in HCT.

The assessment of the IRR was done over 3 phases:

*Initial Assessment Phase:* A sample of 88 patients was randomly selected for comorbidity evaluation during this phase. The PI and another evaluator (Evaluator #1) independently collected comorbidity data from medical records of the 88 patients and then assigned the HCT-CI scores. Evaluator #1 was a first-year fellow in the Hematology-Oncology Program. Evaluator #1 used the HCT-CI as it was previously published with no further
assistance in comorbidity coding from the PI. Then, the HCT-CI scores from the PI and Evaluator #1 were forwarded to the Biostatistician for comparison. Additionally, scores assigned by both single evaluators (PI and Evaluator#1) were compared to those previously determined by multiple evaluators in the clinic (Table 2). The “multiple evaluators” were the medical providers who took care of the 88 patients while receiving their transplant and evaluated their comorbidities prospectively.

**Initial Validation Phase:** As described above, the brief training program was developed to achieve substantial agreement on comorbidity coding by different evaluators. Three evaluators contributed to this phase, Evaluator #1, who contributed to the previous phase and two other novice evaluators, Evaluator #2 and Evaluator #3. Evaluators #2 and #3 were graduates of foreign medical schools with no prior clinical or research experience in the United States. An additional sample of 98 patient charts was randomly selected for this phase. The PI printed out the HCT-CI and the documents for the new training program and handled them to the three evaluators. The PI held 60 minutes-long sessions with each evaluator to review the steps for accessing the medical records for data acquisition and to answer any questions about the comorbidity coding tool. Then, the PI and the four evaluators independently collected the comorbidity data from the medical records of the 98 patients and assigned the HCT-CI scores. The scores were then forwarded to the Biostatistician. This phase had two aims: 1) to demonstrate an improvement in the IRR of Evaluator #1 compared to that in the initial phase, and 2) to show that novice evaluators, Evaluator #2 and 3, could demonstrate excellent IRR rates when provided firsthand with the proposed training program.

**Final Validation Phase:** The web-based application and calculator were established including the guidelines for scoring each of the 17 comorbidities. A fourth evaluator, Evaluator #4 was recruited to validate the training program and the web-based application. Evaluator #4 was a first-year medical student at the University of Washington. Among the 98 patients’ charts included in the initial validation phase, a sample of 30 patient charts was randomly selected for
the final validation phase. The PI handled the documents for the training program including the website for the web-based calculator to Evaluator #4. Evaluator #4 independently assigned scores to the 30 charts and the scores were compared to those previously determined by each of the PI and the other three evaluators during the initial validation phase.

**Statistical Methods**

Kappa statistic is a measure used to analyze inter-rater agreement,\(^{30,31}\) and it adjusts for the degree of agreement that would be expected to occur by chance, and is therefore more appropriate than Pearson’s product moment, Spearman’s correlation, or percent agreement.\(^{32}\) It is reported from 0.0 to 1.0. Weighted Kappa statistic (Kw),\(^{33}\) which assigns less weight to agreement as risk categories are further apart, was computed with Fleiss-Cohen weights\(^{34}\) to analyze the magnitude of inter-rater agreement between two raters on assignment of patients to the HCT-CI risk-categories of 0-1, 2, 3, and \(\geq 4\). Standard errors (S.E.) for kappa and Kw statistics were calculated as previously described.\(^{35}\) Kappa statistic could be used to assess the reliability of agreement between either two raters (Cohen’s kappa)\(^{30}\) or multiple raters (Fleiss’ kappa)\(^{31}\), while weighted Kw is reserved for comparisons between two raters.\(^{33}\) While we report here on results on both methods of assessment, we were more interested in comparing the individual results between two raters, using the Kw, than the average scores among multiple raters, using Fleiss’ kappa statistic. The Landis scale was used for interpretation of the magnitude of kappa and Kw statistics where values <0 indicate no agreement; 0.0-0.20, slight; 0.21-0.40, fair; 0.40-0.60, moderate; 0.61-0.80, substantial; and 0.81-1.00, almost perfect agreement, respectively.\(^{36}\) During the initial phase, we considered that a Kw value below 0.60, while acceptable in some settings, would indicate the need to improve the IRR by new methods and guidelines for comorbidity evaluation. By developing the training program, our goal was to achieve a value of Kw >0.80, indicating excellent agreement, among any two evaluators during
the validation phase in order to validate the ability of the proposed training program to improve the inter-observer congruence when assigning the HCT-CI scores.

Results

Initial Assessment Phase: Among the sample of 88 patients, Evaluator #1 could assign scores for 80 patients, which was the final sample used for this phase. Figure 5A showed the magnitude of variations in the frequency of assigning raw scores when comparing the PI, Evaluator #1, and the multiple evaluators. Variations existed among the three sets of evaluators and across all of the raw scores of 0, 1, 2, 3, 4, and ≥5, but they were more pronounced when comparing the scores assigned by multiple evaluators versus each of the two single evaluators. The Fleiss’ kappa statistic (S.E.) for agreement on the average scores among the three groups of raters was 0.38 (0.06) indicating fair agreement. Likewise, among each two-raters, the use of the HCT-CI to score comorbidities without instructive guidelines resulted in only fair inter-rater agreement with Kw values ranging between 0.433 and 0.585 (Table 2). The Kw statistic was slightly better between the two single evaluators (0.585) compared to that between each of the two single evaluators and other multiple evaluators in the clinic (0.552 and 0.433, respectively). Overall, results indicated a modest rate of IRR without comorbidity coding methodology and guidelines.

Initial Validation Phase: Among the sample of 98 patient charts, Evaluators #2, #3, and #4 could assign scores to 90 charts, which was the final sample for this phase. Evaluator#1 showed a substantial improvement in agreement on assigning the HCT-CI scores versus the PI with Kw (S.E.) of 0.91 (0.03). Similarly, the other two novice evaluators demonstrated excellent IRR rates with Kw of 0.89-0.91 with standard errors of 0.03 for both.

Final Validation Phase: Evaluators #4 assigned scores to the total sample of 30 charts. Figure 5B showed a small magnitude of variation in the frequency of assigning the raw scores when comparing the five evaluators, and most of those variations were limited to the highest
scores (4 and ≥ 5). The Fleiss’ kappa statistic for agreement on the average scores among the five evaluators was substantially improved to 0.80 (0.05) compared to the initial assessment phase. The Kw statistics among each group of two-evaluators were all >0.900 indicating almost perfect agreement (Table 2).

Summary
The HCT-CI can be used to capture the magnitude of organ damage before HCT for a given primary hematological disease. It is also an important tool for decision-making in the clinic, for comparative effectiveness studies of conditioning regimens and graft sources, and for adjustment of statistical analyses for prognostic studies. So far, the index has been evaluated in more than 25 publications from transplant centers world-wide. The index is expected to be continuously utilized in the HCT field. The Center of International Blood and Marrow Transplantation Research (CIBMTR) has incorporated the HCT-CI in routine data collection from transplant centers. The CIBMTR will use the index and other variables in the Center-Outcome Analyses designed to compare outcomes across transplant centers and to provide this information to patients, insurance companies, and academic investigators. In addition, studies on further refinements of the HCT-CI that improve prognostication but retain simplicity of the index would require large data from multiple institutions with consistently evaluated comorbidities. To achieve these goals, a consensus on comorbidity evaluation across centers is mandatory. Here, new methods and guidelines were proposed to facilitate consistent comorbidity coding. We have seen a fair degree of inter-observer agreement when novices assessed comorbidities without standardized guidelines. Similar IRR rates are expected among evaluators from different institutions given that comorbidity assessment is a newly introduced subspecialty to the transplant field.

It is important to report the IRR rates for comorbidity indices to ensure accurate comparison of results from clinical trials across institutions. Multiple studies have previously
reported variable IRR rates for several indices. Yet, little systematic effort has been made to improve the IRR for any comorbidity index. Here, efforts were made to enhance the agreement on the HCT-CI scoring by both improving the instrument (the comorbidity coding tool) and training the evaluators (the methodology and the web-based application). The IRR was the lowest (Kw of 0.433) when scores were compared between a single evaluator and multiple untrained evaluators in the clinic, suggesting the need of a training program to prepare experienced comorbidity evaluators at different institutions.

Participants in the current training program had no prior experience in comorbidity coding and limited or no prior experience in allogeneic HCT. Therefore, we expect that the proposed methods and guidelines could function appropriately in training a wide variety of individuals with different qualifications ranging from study coordinators to experienced transplant physicians. Success of the training program was shown by improvements of both Fleiss’ kappa statistics, among multiple evaluators, from 0.380 to 0.800 and Kw values, among two-evaluators, from 0.433-0.585 to 0.890-0.970 when comparing the initial versus the validation phases, respectively. We would expect these values to be maintained or improved upon when applied on center-specific transplant-oriented individuals. The web-based HCT-CI includes a summary of all the explanatory guidelines for comorbidities, and it provides a user-friendly calculator of the scores. The methods, guidelines, and the web-based application constitute together a brief training program that could be used world-wide by evaluators at single institutions to standardize comorbidity coding.
Acknowledgements: The author would like to thank Dr. Barry Storer for his help with the statistical part of inter-rater reliability and Drs. Fabiana Ostronoff, Saima Ijaz, Aisha Al-Khinji, as well as Jennifer McClure for their participation in the assessment of inter-observer agreement on comorbidity coding. The author is also grateful to Helen Crawford, Bonnie Larson, Sue Carbonneau, Joan Vermeulen, and Karen Carbonneau for their administrative assistance with study implementation and manuscript preparation.

Acknowledgement of research support for the study: This study was supported by HL088021 grant from the National Institutes of Health.

AUTHORSHIP

MLS designed, tested, and validated the training program and the web-based calculator, and wrote the manuscript.

AUTHOR DISCLOSURES

The author has no financial conflict of interest relevant to this article.
REFERENCES


Table 1: The Hematopoietic Cell Transplant-Comorbidity Index (HCT-CI)

<table>
<thead>
<tr>
<th>Comorbidities</th>
<th>HCT-CI scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arrhythmia</td>
<td>1</td>
</tr>
<tr>
<td>Cardiovascular comorbidity</td>
<td>1</td>
</tr>
<tr>
<td>Inflammatory bowel disease</td>
<td>1</td>
</tr>
<tr>
<td>Diabetes or steroid-induced hyperglycemia</td>
<td>1</td>
</tr>
<tr>
<td>Cerebro-vascular disease</td>
<td>1</td>
</tr>
<tr>
<td>Psychiatric disorder</td>
<td>1</td>
</tr>
<tr>
<td>Mild hepatic comorbidity</td>
<td>1</td>
</tr>
<tr>
<td>Obesity</td>
<td>1</td>
</tr>
<tr>
<td>Infection</td>
<td>1</td>
</tr>
<tr>
<td>Rheumatologic comorbidity</td>
<td>2</td>
</tr>
<tr>
<td>Peptic ulcer</td>
<td>2</td>
</tr>
<tr>
<td>Renal comorbidity</td>
<td>2</td>
</tr>
<tr>
<td>Moderate pulmonary comorbidity</td>
<td>2</td>
</tr>
<tr>
<td>Prior malignancy</td>
<td>3</td>
</tr>
<tr>
<td>Heart valve disease</td>
<td>3</td>
</tr>
<tr>
<td>Moderate/severe hepatic comorbidity</td>
<td>3</td>
</tr>
<tr>
<td>Severe pulmonary comorbidity</td>
<td>3</td>
</tr>
<tr>
<td><strong>Total Score</strong></td>
<td>=________</td>
</tr>
<tr>
<td>Study Phases</td>
<td>Evaluators</td>
</tr>
<tr>
<td>-------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial Assessment Phase</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PI vs. Evaluator #1</td>
</tr>
<tr>
<td></td>
<td>PI vs. multiple evaluators</td>
</tr>
<tr>
<td></td>
<td>Evaluator #1 vs. multiple evaluators</td>
</tr>
<tr>
<td>Initial Validation Phase</td>
<td>PI vs. Evaluator #1</td>
</tr>
<tr>
<td></td>
<td>PI vs. Evaluator #2</td>
</tr>
<tr>
<td></td>
<td>PI vs. Evaluator #3</td>
</tr>
<tr>
<td>Final Validation Phase</td>
<td>PI vs. Evaluator #4</td>
</tr>
<tr>
<td></td>
<td>Evaluator #1 vs. Evaluator #4</td>
</tr>
<tr>
<td></td>
<td>Evaluator #2 vs. Evaluator #4</td>
</tr>
<tr>
<td></td>
<td>Evaluator #3 vs. Evaluator #4</td>
</tr>
</tbody>
</table>

*Weighted kappa statistic (Kw) was calculated for agreement on HCT-CI risk groups of 0-1, 2, 3, and ≥4.

PI indicates principal investigator.
Figures Legends

**Figure 1.** Schema for the three-step methodology for acquisition of comorbidity data from medical records.

**Figure 2: Pulmonary function tests for 4 different patients.** There are 4 reports of pulmonary function tests for 4 different patients. Each report contains 5 columns of information in the following order: 1) type/name of the test, 2) unit of the test, 3) reference “predicted” value for each test, 4) the pre-bronchodilator measured value for each test, and 5) the pre-bronchodilator percentage, of measured divided by predicted values, for each test. Post-bronchodilator values and percentages were not shown here as they do not contribute to the coding of pulmonary comorbidity per the HCT-CI. All uncorrected values or percentages of measured DLco were corrected for the concurrent hemoglobin values using Dinakara equation \[\text{Corrected DLco} = \frac{\text{uncorrected DLco}}{(0.06965 \times \text{Hemoglobin g/dl})}\].

**A)** This is a patient who had an FEV1 percentage of 97%. The patient also had an uncorrected DLco percentage of 63%. The patient’s concurrent hemoglobin value was 11.1 mg/dl; hence, he had a corrected DLco percentage of 81%. This patient was assigned a score of 0 for pulmonary comorbidity.

**B)** This patient had an FEV1 percentage of 77%. This patient also experienced an uncorrected DLco percentage of 60%. Since the concurrent hemoglobin value was 11.0 mg/dl, then, the corrected percentage of DLco is 78%. This patient was assigned a score of 2 for pulmonary comorbidity due to both the FEV percentage of 77% and the corrected DLco percentage of 78%.

**C)** This is a patient with an FEV1 percentage of 71%. The uncorrected DLco percentage was 41%. The concurrent hemoglobin value was 11.5 mg/dl; therefore, the patient has a
corrected DLco percentage of 52%. This patient was assigned a score of 3 for pulmonary comorbidity due to the corrected DLco percentage of 52%.

D) This patient experienced an FEV1 percentage of 47%. The uncorrected DLco percentage was 40%. The concurrent hemoglobin value was 10.4 mg/dl; therefore, the corrected DLco percentage was 56%. This patient was assigned a score of 3 for pulmonary comorbidity due to both the FEV percentage of 47% and the corrected DLco percentage of 56%.

**Figure 3: Hepatic function tests for 5 patients.** There are 5 reports of hepatic function tests for 5 different patients. Each report contains 2 columns of information in the following order: 1) a scale for days in relation to the hematopoietic cell transplantation (HCT) and 2) the measured values for each test. Name and reference range for each test is provided in the top row of each report. Laboratory values between days -24 and -10 before HCT were the only values considered for evaluation of hepatic comorbidity. The elevated laboratory value that was closest to day -10 was used for assigning the severity score and it was highlighted in yellow. The red asterisk (*) means values were outside the reference range.

A) This patient had 2 normal values of alanine aminotransferase (ALT) at days -17 and -14 before HCT. No score was assigned to the ALT portion of hepatic comorbidity for this patient.

B) This patient had only three values of ALT between days -24 and -10 before HCT, two values were normal and the other was elevated. Even though this patient had multiple elevated values of ALT after day -10, no score was assigned for the ALT portion of hepatic comorbidity since the patient did not have two elevated values of ALT between days -24 and -10.

C) This patient had one normal and two elevated values of aspartate aminotransferase (AST) between days -22 and -12 before HCT. The maximal score, a score of 3, was assigned for hepatic comorbidity in this patient since the closest elevated value of AST to day -10 was
105U/L, which was >2.5 times the upper limit of normal (ULN) for AST based on the reference range. Note that the values after day -12 did not contribute to the evaluation of this comorbidity.

D) This patient had one normal and four elevated values of AST between days -21 and -10 before HCT. A score of 1 was assigned to the AST portion for hepatic comorbidity in this patient given that the closest elevated value of AST to day -10 was 44 U/L, which was <2.5 the ULN for AST. A score of 1 was assigned despite normalization of AST values after day -10. A complete evaluation of other hepatic laboratory tests (ALT and bilirubin) was still required before a decision is made on the maximal score to be assigned for hepatic comorbidity in this patient.

E) This patient had multiple elevated values of total bilirubin between days -24 and -10 before HCT. A score of 1 was assigned for the bilirubin portion of hepatic comorbidity in this patient since the value of total bilirubin at day -10 was 1.9 mg/dl (<2.5 the ULN for total bilirubin). A complete evaluation of other hepatic laboratory tests (ALT and AST) was required before a decision is made on the maximal score to be assigned for hepatic comorbidity in this patient.

**Figure 4: Serum creatinine values for 3 patients.** There are 3 reports of serum creatinine for 3 different patients. Each report contains 2 columns of information in the following order: 1) a scale for days in relation to the hematopoietic cell transplantation (HCT) and 2) the measured values for each test. Reference range for each test is provided in the top row of each report. The red asterisk (*) means values were outside the reference range. The reference range for serum creatinine was 0.3-1.2 mg/dL.

Values of serum creatinine that were tested between days -24 and -10 before HCT were the only values considered for evaluation of renal comorbidity.

A) This patient had three normal and two elevated values of serum creatinine between days -24 and -10 before HCT. No score was assigned for renal comorbidity in this patient since
both of the elevated values were less than 2 mg/dl. The two values of serum creatinine at days -9 and -8, which were higher than 2 mg/dl, did not contribute to the evaluation of this comorbidity since they were assessed after day -10.

B) This patient had consistently elevated values of serum creatinine throughout the whole laboratory report. However, no score was assigned for renal comorbidity in this patient since there was only a single value of >2 mg/dl at day -14, which was not sufficient for coding this comorbidity.

C) This patient had consistently elevated values of serum creatinine throughout the whole laboratory report. A score of 2 was assigned for renal comorbidity in this patient since there were two values of serum creatinine of >2 mg/dl between days -24 and -10 before HCT.

**Figure 5: The frequency of distribution of the raw scores for the HCT-CI by different evaluators.** A) There was a great variation in the frequency of assigning the raw scores of 0, 1, 2, 3, 4, and ≥5 by any of the three groups of evaluators during the initial assessment phase. B) A limited variation could be seen in the frequency of assigning the raw scores of 0, 1, 2, 3, 4, and ≥5 by the five evaluators during the final validation phase.
### Fig 1: Three-Step Process (15 minutes)

<table>
<thead>
<tr>
<th>Medical Notes 8 min</th>
<th>Labs/Tests 6 min</th>
<th>Final Assessment 1 min</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nutrition</td>
<td>Bilirubin/AST/ALT</td>
<td>Double checking</td>
</tr>
<tr>
<td>H&amp;P</td>
<td>Serum creatinine</td>
<td>Total score</td>
</tr>
<tr>
<td>Review of Data</td>
<td>DLco/FEV1</td>
<td></td>
</tr>
<tr>
<td>Consults</td>
<td>Echo/MUGA</td>
<td></td>
</tr>
</tbody>
</table>

**Cumulative Data Acquisition Scale (%)**

- 0%
- 5%
- 20%
- 40%
- 45%
- 55%
- 60%
- 80%
- 90%
- 95%
- 100%
### Fig 2

#### A

<table>
<thead>
<tr>
<th></th>
<th>Ref</th>
<th>Pre Meas</th>
<th>Pre % Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spirometry</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEV1</td>
<td>Liter(s)</td>
<td>2.50</td>
<td>2.42</td>
</tr>
<tr>
<td><strong>Diffusion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hb</td>
<td>gm/dL</td>
<td>11.1</td>
<td>Uncorrected DLco%</td>
</tr>
<tr>
<td>DLCO</td>
<td>mL/mmHg/min</td>
<td>24.2</td>
<td>15.1</td>
</tr>
<tr>
<td>DLCOcorr</td>
<td></td>
<td>19.6</td>
<td>Corrected DLco%</td>
</tr>
</tbody>
</table>

#### B

<table>
<thead>
<tr>
<th></th>
<th>Ref</th>
<th>Pre Meas</th>
<th>Pre % Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spirometry</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEV1</td>
<td>Liter(s)</td>
<td>4.11</td>
<td>3.17</td>
</tr>
<tr>
<td><strong>Diffusion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hb</td>
<td>gm/dL</td>
<td>11.0</td>
<td>Uncorrected DLco%</td>
</tr>
<tr>
<td>DLCO</td>
<td>mL/mmHg/min</td>
<td>38.1</td>
<td>22.7</td>
</tr>
<tr>
<td>DLCOcorr</td>
<td></td>
<td>29.7</td>
<td>Corrected DLco%</td>
</tr>
</tbody>
</table>

#### C

<table>
<thead>
<tr>
<th></th>
<th>Ref</th>
<th>Pre Meas</th>
<th>Pre % Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spirometry</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEV1</td>
<td>Liter(s)</td>
<td>3.82</td>
<td>2.70</td>
</tr>
<tr>
<td><strong>Diffusion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hb</td>
<td>gm/dL</td>
<td>11.5</td>
<td>Uncorrected DLco%</td>
</tr>
<tr>
<td>DLCO</td>
<td>mL/mmHg/min</td>
<td>35.6</td>
<td>14.8</td>
</tr>
<tr>
<td>DLCOcorr</td>
<td></td>
<td>18.4</td>
<td>Corrected DLco%</td>
</tr>
</tbody>
</table>

#### D

<table>
<thead>
<tr>
<th></th>
<th>Ref</th>
<th>Pre Meas</th>
<th>Pre % Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spirometry</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEV1</td>
<td>Liter(s)</td>
<td>3.04</td>
<td>1.44</td>
</tr>
<tr>
<td><strong>Diffusion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hb</td>
<td>gm/dL</td>
<td>10.4</td>
<td>Uncorrected DLco%</td>
</tr>
<tr>
<td>DLCO</td>
<td>mL/mmHg/min</td>
<td>28.0</td>
<td>11.3</td>
</tr>
<tr>
<td>DLCOcorr</td>
<td></td>
<td>15.6</td>
<td>Corrected DLco%</td>
</tr>
</tbody>
</table>
Fig 3

(A) ALT (Reference Range 10–48)

Days before HCT | Test Result
--- | ---
0 | 120 U/L*
-2 | 110 U/L*
-4 | 44 U/L
-7 | 16 U/L
-9 | 23 U/L
-14 | 14 U/L
-17 | 16 U/L

(B) ALT (Reference Range 6–40)

Days before HCT | Test Result
--- | ---
0 | 45 U/L*
-1 | 63 U/L*
-2 | 71 U/L*
-3 | 97 U/L*
-5 | 96 U/L*
-7 | 109 U/L*
-10 | 35 U/L
-15 | 30 U/L
-22 | 91 U/L*

(C) AST (Reference Range 15–40)

Days before HCT | Test Result
--- | ---
0 | 32 U/L
-1 | 31 U/L
-2 | 37 U/L
-5 | 37 U/L
-8 | 45 U/L*
-12 | 105 U/L*
-15 | 187 U/L*
-22 | 26 U/L

(D) AST (Reference Range 15–40)

Days before HCT | Test Result
--- | ---
0 | 24 U/L
-4 | 34 U/L
-6 | 33 U/L
-8 | 37 U/L
-10 | 44 U/L*
-11 | 58 U/L*
-14 | 50 U/L*
-17 | 42 U/L*
-21 | 28 U/L

(E) Bilirubin (Reference Range 0.2–1.3)

Days before HCT | Test Result
--- | ---
0 | 1.2 mg/dL
-2 | 1.3 mg/dL
-4 | 1.4 mg/dL*
-6 | 1.4 mg/dL*
-10 | 1.9 mg/dL*
-11 | 2 mg/dL*
-14 | 1.6 mg/dL*
-17 | 1.6 mg/dL*
-21 | 1.6 mg/dL*
-24 | 1.5 mg/dL*
Fig 5

A

% of patients

HCT-CI scores

Principal Investigator
Evaluator #1
Multiple Evaluators

B

% of patients

HCT-CI scores

Principal Investigator
Evaluator #1
Evaluator #2
Evaluator #3
Evaluator #4