

The use of prognostic models in allogeneic transplants: a perspective guide for clinicians and investigators

Mohamed L. Sorror

Clinical Research Division, Fred Hutchinson Cancer Center and University of Washington School of Medicine, Seattle, WA

Allogeneic hematopoietic cell transplant (HCT) can cure many hematologic diseases, but it carries the potential risk of increased morbidity and mortality rates. Prognostic evaluation is a scientific entity at the core of care for potential recipients of HCT. It can improve the decision-making process of transplant vs no transplant, help choose the best transplant strategy and allows for future trials targeting patients' intolerances to transplant; hence, it ultimately improves transplant outcomes. Prognostic models are key for appropriate actuarial outcome estimates, which have frequently been shown to be better than physicians' subjective estimates. To make the most accurate prognostic evaluation for HCT, one should rely on >1 prognostic model. For relapse and relapse-related mortality risks, the refined disease risk index is currently the most informative model. It can be supplemented with disease-

specific models that consider genetic mutations as predictors in addition to information on measurable residual disease. For nonrelapse mortality and HCT-related morbidity risks, the HCT-comorbidity index and Karnofsky performance status have proven to be the most reliable and most accepted by physicians. These can be supplemented with gait speed as a measure of frailty. Some other global prognostic models might add additional prognostic information. Physicians' educated perceptions can then put this information into context, taking into consideration conditioning regimen and donor choices. The future of transplant mandates (1) clinical investigators specifically trained in prognostication, (2) increased reliance on geriatric assessment, (3) the use of novel biomarkers such as genetic variants, and (4) the successful application of novel statistical methods such as machine learning.

Introduction

In 2005, the hematopoietic cell transplant-comorbidity index (HCT-CI) was developed to diagnose burdensome comorbidities that contribute to increased nonrelapse mortality (NRM) after allogeneic HCT.¹ This work was motivated, then, by 3 reasons. Firstly, although allogeneic HCT is potentially curative, it traditionally carried risks of morbidity and mortality; pre-transplant comorbidities that contribute to these risks need to be accounted for. Secondly, the advent of reduced-intensity regimens urged the need for tools to compare their recipients with those given high-dose regimens. Thirdly, older patients (aged ≥ 60 years) were increasingly offered allogeneic HCT; aging is known to be associated with increased number and severity of comorbidities.²

Transplant outcomes continue to improve,^{3,4} yet, we are far from satisfactory optimization of these outcomes. Data from the Center for International Blood and Marrow Transplant Research (CIBMTR) indicate that individuals aged ≥ 65 years constitute 27% of HCT recipients in 2020.⁵ Therefore, reduced-intensity regimens have allowed older patients with different risk profiles compared with their younger counterparts (Figure 1), to be

offered HCT. In agreement, multicenter transplant trials indicate expansion in the percentage of those with HCT-CI scores of ≥ 3 to be up to 55%.⁷ We also continue to see NRM rates in the range from 33% to 35% among older patients.⁸ Overall survival (OS) rates continue to be ~54% to 58%, at best,⁵ and cumulative incidences of relapse are at 44%, both at 3 years.⁴

Given these data, there have been efforts, and rightly so, to either validate the HCT-CI, modify it, or create new prognostic models⁹⁻²⁰ to improve our ability to predict transplant outcomes. We have learned much over the past 2 decades, but many questions remain. Which tools are best to use in the clinic? How do these tools compare with each other? Should we replace them or use them together? Can we improvise the current tools? (along with many other questions). Herein, I attempt to answer these questions from my perspective and discuss future directions.

Is prognostication a science?

Prognostication is "the relative probabilities of the various outcomes of the natural history of a disease."²¹ It is essential to the science of patient care science and has become an integral part of the day-to-day clinical skills of oncologists while caring

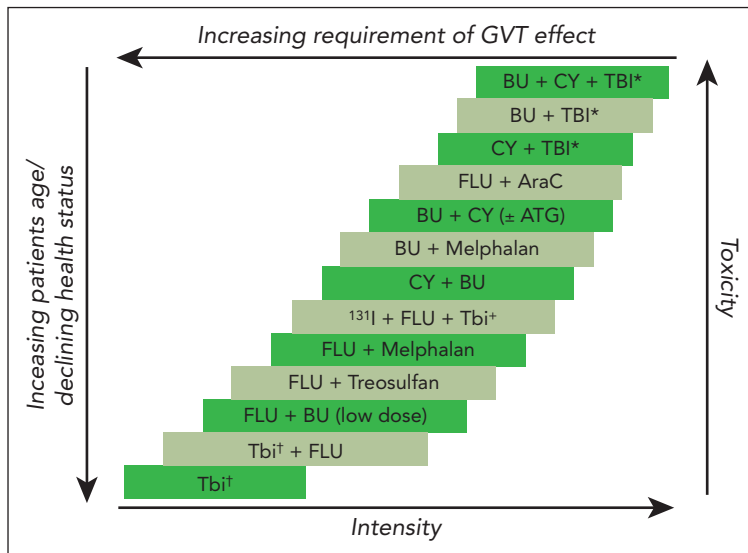


Figure 1. The complexity of transplant decision making. Schematic demonstration of the close correlation between an increase in intensity of conditioning regimens and increased anticipated toxicities from transplant, between a decrease in conditioning intensity and the increased reliance on graft-versus-tumor effect for transplant success, and between decrease in conditioning intensity and increased age and declined health status of those accepted for allogeneic transplants that then result in increased chances for toxicities (despite the reduced conditioning intensity). Adapted from Deeg and Sandmaier.⁶

for their patients. Furthermore, it is a science that requires a set of skills, experiences, and appropriate training. Studies show that the majority of internists (56.8%) feel that they are inappropriately trained in prognostication.²² Prognostication has many advantages varying from aid in choosing appropriate therapy to prevention of complications.²¹ It may be divided into (1) subjective, made directly by the clinician who is taking care of the patient; (2) actuarial, relying on a set of prognostic factors; or (3) combined, using a combination of subjective and actuarial. Prognostication has 2 critical components: (1) foreseeing, the science of formulation of prediction; and (2) foretelling, the science of communicating the prediction.²³ Herein, the focus will entirely be on the science of foreseeing. There are 2 types of foreseeing or prediction formulation: (1) temporal, which covers prediction of the time-to-event; and (2) probabilistic, which covers prediction of the absolute risk of an event.

Finally, prognostication is complex. It covers 6 different domains that could be labeled as the “6 D’s”: (1) death, (2) disease progressions, (3) disability, (4) drug toxicity, (5) dollar (cost), and (6) derivative health (impact on others).^{21,24} Most of the research done so far in the field of HCT focuses primarily on only the first 2 D’s, highlighting how far behind we are in exploring this field of science and its huge impact on our patients and their relatives.

In conclusion, the HCT field has a great need for experts in prognostication to further expand on current research efforts and improve our methods of predicting outcomes.

How important is the use of a prognostic model before HCT?

For the most part, there is a lack of randomized trials that compare allogeneic HCT with other options; this, together with the need to make accurate decisions in the clinic, necessitates the need for prognostic models. As mentioned earlier, actuarial prognostication is exclusively dependent on the use of ≥ 1 prognostic factors. Actuarial is superior to subjective

prognostication. Physicians’ perceptions of OS tend to be optimistically incorrect by a factor ranging from 3 to 5.²⁵ Physicians’ estimates of duration of survival (temporal prediction) for cancer patients are correct for only a third and can occasionally be pessimistic.^{26,27} To further complicate the issue of subjective prognostication, 40.2% physicians were reportedly found to communicate to their patients more optimistic survival estimates than what they actually thought.²⁸ Hence, the majority of patients could potentially be misled twice, mostly with false optimism, from the reality of their prognosis: first through an incorrectly formulated and then an incorrectly communicated prognosis. Physicians’ experiences positively correlate with improved prognostication, whereas the duration and intensity of the doctor-patient relationship had negative impacts on prognostic estimates.²⁵ This suggests a benefit from multidisciplinary care rather than an individual physician decision.

It is therefore not surprising that a recent prospective observational study of patients with acute myeloid leukemia (AML) who were being considered for HCT found a large discordance between physician estimates (6%) and patient estimates (61%) of good (>75%) survival.²⁹ Those estimates were affected by the type of treatment (higher for those receiving HCT relative to those not receiving -HCT) chosen by the physicians for their patients, whereas recipients of HCT and those not receiving HCT had almost identical estimates of survival for themselves.²⁹ In the same trial (NCT01929408), it was found that physicians’ estimates of 1-year survival had a lower area under the receiver operating characteristic curve (area under curve [AUC]) of 0.61 compared with a single prognostic marker, such as the Karnofsky performance status (KPS, 0.65), whereas both were inferior to a composite model of increased age, comorbidity burden, and cytotoxic risks, namely, the AML-composite model (AML-CM had an AUC of 0.74). Physicians’ estimates of both survival and physician-assigned KPS correlated with increasing patient age ($r = .31$; $P < .0001$; and $r = .14$; $P = .001$; respectively), suggesting a potential bias (unpublished results).

Risk aversion and values could differ between physician and patient. A patient’s overestimation of the chances of cure might

not be shared with the physician (poor risk aversion). Risk aversion can be considered in clinical decision models by tools such as standard gamble and time trade-off. Moreover, the uncertainty surrounding outcome forecasts could be quantified (eg, by Bayesian statistics) because treatment uncertainty is as relevant as the risks themselves.^{30,31}

In summary, we cannot rely solely on subjective prognostication; it creates a cascade of problems in patient management. Available prognostic tools in the field, albeit not perfect, are generally better than a physician's guess alone. A compromised and likely effective solution (until formally tested) would be to use actuarial prognostication as the basis of evaluation, but it should be supplemented with the physician's educated judgment.³² Consultation with other physicians might be helpful to avoid biased estimates driven by the physician-patient relationship.

What are the main sectors of prognostication in HCT?

Allogeneic HCT is a complex procedure. Many factors contribute to its appropriate prognostication. There are 3 major sectors: (1) disease-specific prognostication, including disease diagnoses, statuses, and cytogenetic/molecular features. Other less important factors include time from diagnoses and the number of prior therapies. (2) Transplant-related prognostication, which involves the donor type, magnitude of HLA and allele matching, and the source of stem cells. The more recent, and still developing, sector is (3) patient-specific prognostication. Formerly, it relied exclusively on age and KPS; comorbidity burden, other geriatric assessments, and laboratory biomarkers were added later. Global prognostic models have been designed to include selected factors from each sector, whereas combination models combine >1 model (Table 1).

Primary disease-related tools

The disease risk index (DRI) suitably and comprehensively captures the impact of primary diagnosis, disease status, histologic subtypes (for lymphomas), and chromosomal aberrations (for AML, acute lymphoblastic leukemia [ALL], and myelodysplastic syndromes [MDS]) on the odds of OS.¹¹ The DRI creates 4 risk groups with a 4-year OS ranging from 6% to 64%. A study had an independent validation set but did not include information on measures of model performances, such as c-statistic.³⁴ A follow-up validation and model refinement study had minimal improvement in c-statistic estimate over the original DRI (0.643 vs 0.637 for OS prediction).¹³

The DRI provides important prognostic information directly related to the risks of relapse and relapse-related mortality after HCT. It is considered a vital complement to patient-risk assessment tools, such as HCT-CI and KPS, when making decisions in the clinic. It can also be used to compare outcomes of patients with different diagnoses in outcome research studies. Yet, it misses the more advanced molecular prognostic features of certain diseases as well as appraisal of measurable residual disease (MRD); hence, it needs to be complemented.

Disease-specific models include molecular prognostic features for each disease that can be used when counseling patients about the benefits of HCT. These include the European

Leukemia Network Classification 2022³⁵ in AML, the International Prognostic Scoring System, Molecular for MDS,³⁶ the Mutation and Karyotype-enhanced International Prognostic Scoring System for primary myelofibrosis,^{37,38} and the chronic myelomonocytic leukemia-specific Prognostic Scoring System with Molecular Features for chronic myelomonocytic leukemia.³⁹ Some of these models are not fully evaluated for prediction of post-HCT outcomes, but they are widely regarded and used in the clinic because they are based on clinical features that have repeatedly been demonstrated to be prognostic, in addition to molecular genetic mutations that have high sensitivity and specificity for predicting mortality.

MRD at the time of allogeneic HCT constitutes a remarkably important and measurably easy (MRD vs no MRD) prognostic parameter for both AML^{40,41} and ALL.⁴² Yet, standardization of MRD measurement remains a work in progress.^{43,44} This topic is beyond the scope of this article but, in general, timing, type, and size of sampling as well as method of measuring MRD remain controversial.

Patient-related tools

The HCT-CI is the most frequently tested and validated tool in the setting of allogeneic HCT. In its original study, the HCT-CI had c-statistic for prediction of 1- and 2-year NRM of 0.692 and 0.685, respectively. For prediction of OS as a secondary outcome, these figures were 0.661 and 0.657, respectively.¹ Patients with the highest comorbidity scores (≥ 3 ; ~35% to 45% of recipients of HCT) experience significant increases in morbidity,⁴⁵ mortality,⁴⁶ long-term impairments in quality of life,⁴⁷ and use of resources.⁴⁸ Comorbidity scoring has been standardized to ensure improved reproducibility across investigators and institutions.⁴⁹ It has been combined with many other prognostic factors or tools such as KPS,⁵⁰ age,⁵¹ biomarkers,⁵² the European Group for Blood and Marrow Transplantation (EBMT) model,⁵³ and instrumental activities of daily living (IADL),⁵⁴ to improve prognostication. The model has been widely validated,⁵⁵ and it remains the only model that was prospectively validated in 2 large studies, providing a higher level of evidence than retrospective or single-center studies.^{56,57} A lack of validation in a smaller number of studies seemed to result from a lack of essential model components (eg, pulmonary function tests),⁵⁸ a very small patient sample size,⁵⁹ the lack of model performance testing methods (eg, c-statistics),⁶⁰ or a lack of information on inter-rater reliability of score assignment.¹⁷ It is regarded by the majority of surveyed transplant physicians as reflective of patients' overall health.⁶¹

Some studies attempted to modify the model. The only study that suggested improvement in the AUC, pending independent outside validation, was the simplified comorbidity index (SCI) in which the AUC for 1-year NRM was 0.72 compared with 0.657 for the HCT-CI.¹⁶ Interestingly, these figures were reversed when predicting 4-year NRM (0.618 vs 0.644, respectively),¹⁶ suggesting altered impacts of each model over time that could be because of the higher presentation of comorbidities in the HCT-CI. The SCI was developed by removing less-frequent comorbidities and those with P values $>.05$ for association with NRM. As a result, patients with SCI scores of ≥ 3 constituted only 21%, indicating the capturing of a smaller cohort with higher NRM risks compared with the cohort captured by the HCT-CI (47%). The study did not provide information about the

Table 1. Models investigated in recipients of allogeneic HCT and the required information per model with an accompanying link to an online calculator or mobile application whenever available

| HCT-CI comorbidities | PAM | Modified PAM | Modified EBMT | HCT-CI/EBMT | R-DRI* | NRM-J | The SCI | TRM | AML-composite model (AML-CM) | HCT-CR | EASIX |
|----------------------------|--------------------------------------------|----------------------------------|----------------------------------------------|-----------------------------------|--------------|---------------------------------------------|-----------------------------------------------------------------|------------------------------------|-----------------------------------------------------------------|------------------|----------------|
| Arrhythmia | Age (y) | Age: ≥65 | Age: <20, 20-40, and >40 | HCT-CI = 0 and EBMT score = 0-3 | Low | Age: 50-59 and ≥60 | Age: ≥60 | Age (y) | Age: 50-59 and ≥60 | HCT-CI/Age model | Creatinine |
| Cardiac | Donor type | Donor type | Disease status: CR1, CR >1, and no CR | HCT-CI = 0 and EBMT score = 4-7 | Intermediate | Sex: male | Composite cardiac: arrhythmia, cardiac, and heart valve disease | Performance status | All HCT-CI comorbidities | Revised DRI | LDH |
| Inflammatory bowel disease | Disease risk† | Disease risk‡ | Donor: MRD vs others | HCT-CI = 1-2 and EBMT score = 0-3 | High | Performance status: 1 and ≥2 | Hepatic—moderate/severe | White blood cell count | LDH (>200-1000 and >1000) | | Platelet count |
| Diabetes | Conditioning§ | FEV1 | Donor/recipient sex: female, male, or others | HCT-CI = 0 and EBMT score = 4-7 | Very high | HCT-CI: ≥3 | Pulmonary—moderate | Peripheral blood blast percentage | Albumin (<3.5) | — | — |
| Cerebrovascular | FEV1: >80%, 70%-80%, <70% | Patient/donor CMV: -/-, +/-, +/+ | | HCT-CI = ≥3 and EBMT score = 0-3 | — | Donor: unrelated BM, related PB, cord blood | Pulmonary—severe | Type of AML [de novo vs secondary] | 2017 ELN cytomolecular risks (favorable, intermediate, adverse) | — | — |
| Psychiatric | DLco: >80%, 70%-80%, <70% | — | — | HCT-CI = ≥3 and EBMT score = 4-7 | — | — | Renal dysfunction per EGFR¶ | Platelet count | — | — | — |
| Hepatic—mild | Serum creatinine: >106 μmol/L (>1.2 mg/dL) | — | — | — | — | — | — | Albumin | — | — | — |
| Obesity | ALT: >22 μmol/L (>1.3 mg/dL) | — | — | — | — | — | — | Creatinine | — | — | — |

Online calculators: HCT-CI (<http://hctci.org/Home/Calculator> or <http://tgapp.asbmt.org/>); PAM (not available); The modified EBMT (not available); R-DRI (<https://cibmtr.org/CIBMTR/Resources/Research-Tools-Calculators/Disease-Risk-Index-DRI-Assignment-Tool> or <http://tgapp.asbmt.org/>); NRM-J (not available); The SCI (<http://tgapp.asbmt.org/>); TRM (<https://tmcalculator.fredhutch.org/>); AML-CM (<http://amlcompositemodel.org/Home/GetCalculator>); and EASIX (<https://biostatistics.dkfz.de/EASIX/>).

ALT, alanine transaminase; CMV, cytomegalovirus; EASIX, Endothelial Activation and Stress Index; ELN-2022, European Leukemia Network Classification 2022; LDH, lactate dehydrogenase.

*Disease status: *low*: Hodgkin lymphoma (HL) in complete remission (CR), chronic lymphocytic leukemia (CLL) in CR or partial remission (PR), mantle cell lymphoma CR, indolent lymphoma CR or PR, acute myeloid leukemia (AML) favorable cytogenetics CR, chronic myeloid leukemia (CML) chronic phase 1 or 2; *intermediate*: CML advanced phase, mantle cell lymphoma PR, myeloproliferative disease, AML intermediate cytogenetics CR, acute lymphoblastic leukemia (ALL) CR1, T-cell NHL CR/PR, multiple myeloma (MM) CR/very good partial remission (VGPR)/PR, aggressive non-Hodgkin lymphoma (NHL) CR, low-risk myelodysplastic syndromes (MDS) intermediate cytogenetics (early/advanced), low-risk MDS adverse cytogenetics (early), advanced indolent NHL, advanced CLL, aggressive NHL PR; *high*: advanced T-cell NHL, advanced AML favorable cytogenetics, advanced HL, advanced high-risk MDS intermediate cytogenetics, early/advanced high-risk MDS adverse cytogenetics, ALL CR2/CR3, AML adverse cytogenetics CR, advanced mantle cell lymphoma, Burkitt lymphoma CR, advanced MM, advanced low-risk MDS adverse cytogenetics, advanced AML intermediate cytogenetics; *very high-risk*: CML blast phase, advanced ALL, advanced aggressive NHL, advanced AML adverse cytogenetics, advanced Burkitt lymphoma (BL) in PR.

†Disease risks include low-risk diseases included chronic myelogenous leukemia in chronic phase, refractory anemia, aplastic anemia, and the Blackfan-Diamond syndrome. Intermediate-risk diseases included chronic myelogenous leukemia in accelerated or phase after blastic phase, acute leukemia or lymphoma in remission, refractory anemia with excess blasts, chronic lymphocytic leukemia, and paroxysmal nocturnal hemoglobinuria. High-risk diseases included chronic myelogenous leukemia in blastic phase, juvenile chronic myelogenous leukemia, acute leukemia or lymphoma in relapse, refractory anemia with excess blasts in transformation, and myeloma. Solid tumors and nonhematologic diseases were also classified as high-risk diseases.

‡Overall-risk groups were determined based on the risk index developed by Armand et al, which includes disease risk, stage risk, and cytogenetic data for AML and MDS. The poor and very poor MDS cytogenetic risk categories defined by Deeg et al³³ were grouped as high-risk disease, and all other categories were grouped as intermediate-risk disease.

§Myeloablative regimens were categorized based on the dose of total-body irradiation used (12 Gy or 12 Gy). All patients in the nonmyeloablative group received 2 Gy of total-body irradiation.

||The relative change in hazard ratio for each decrease in FEV1 by 10%.

¶Based on estimated glomerular filtration rate (eGFR) mL/min per 1.73 m², using the chronic kidney disease epidemiology collaboration (CKD-EPI) formula for estimating creatinine clearance. An eGFR of ≥90 mL/min per 1.73 m² was considered normal, from 60 to 89.9 mL/min per 1.73 m² was considered to be mildly decreased, and <60 mL/min per 1.73 m² as moderately to severely decreased.

Table 1 (continued)

| HCT-CI comorbidities | PAM | Modified PAM | Modified EBMT | HCT-CI/EBMT | R-DRI [†] | NRM-J | The SCI | TRM | AML-composite model (AML-CM) | HCT-CR | EASIX |
|-------------------------|-----|--------------|---------------|-------------|--------------------|-------|---------|-----|------------------------------|--------|-------|
| Infection | — | — | — | — | — | — | — | — | — | — | — |
| Rheumatologic | — | — | — | — | — | — | — | — | — | — | — |
| Peptic ulcer | — | — | — | — | — | — | — | — | — | — | — |
| Renal—moderate/severe | — | — | — | — | — | — | — | — | — | — | — |
| Pulmonary—moderate | — | — | — | — | — | — | — | — | — | — | — |
| Prior malignancy | — | — | — | — | — | — | — | — | — | — | — |
| Heart valve disease | — | — | — | — | — | — | — | — | — | — | — |
| Pulmonary—severe | — | — | — | — | — | — | — | — | — | — | — |
| Hepatic—moderate/severe | — | — | — | — | — | — | — | — | — | — | — |

Online calculators: HCT-CI (<http://hctci.org/Home/Calculator> or <http://tgapp.asbmt.org/>); PAM (not available); The modified EBMT (not available); R-DRI (<https://cibmtr.org/CIBMTR/Resources/Research-Tools-Calculators/Disease-Risk-Index-DRI-Assignment-Tool> or <http://tgapp.asbmt.org/>); NRM-J (not available); The SCI (<http://tgapp.asbmt.org/>); TRM (<https://tmcalculator.fredhutch.org/>); AML-CM (<http://amlcompositemodel.org/Home/GetCalculator>); and EASIX (<https://biostatistics.dkfz.de/EASIX/>).

ALT, alanine transaminase; CMV, cytomegalovirus; EASIX, Endothelial Activation and Stress Index; ELN-2022, European Leukemia Network Classification 2022; LDH, lactate dehydrogenase.

*Disease status: *low*: Hodgkin lymphoma (HL) in complete remission (CR), chronic lymphocytic leukemia (CLL) in CR or partial remission (PR), mantle cell lymphoma CR, indolent lymphoma CR or PR, acute myeloid leukemia (AML) favorable cytogenetics CR, chronic myeloid leukemia (CML) chronic phase 1 or 2; *intermediate*: CML advanced phase, mantle cell lymphoma PR, myeloproliferative disease, AML intermediate cytogenetics CR, acute lymphoblastic leukemia (ALL) CR1, T-cell NHL CR/PR, multiple myeloma (MM) CR/very good partial remission (VGPR)/PR, aggressive non-Hodgkin lymphoma (NHL) CR, low-risk myelodysplastic syndromes (MDS) intermediate cytogenetics (early/advanced), low-risk MDS adverse cytogenetics (early), advanced indolent NHL, advanced CLL, aggressive NHL PR; *high*: advanced T-cell NHL, advanced AML favorable cytogenetics, advanced HL, advanced high-risk MDS intermediate cytogenetics, early/advanced high-risk MDS adverse cytogenetics, ALL CR2/CR3, AML adverse cytogenetics CR, advanced mantle cell lymphoma, Burkitt lymphoma CR, advanced MM, advanced low-risk MDS adverse cytogenetics, advanced AML intermediate cytogenetics; *very high-risk*: CML blast phase, advanced ALL, advanced aggressive NHL, advanced AML adverse cytogenetics, advanced Burkitt lymphoma (BL) in PR.

†Disease risks include low-risk diseases included chronic myelogenous leukemia in chronic phase, refractory anemia, aplastic anemia, and the Blackfan-Diamond syndrome. Intermediate-risk diseases included chronic myelogenous leukemia in accelerated or phase after blastic phase, acute leukemia or lymphoma in remission, refractory anemia with excess blasts, chronic lymphocytic leukemia, and paroxysmal nocturnal hemoglobinuria. High-risk diseases included chronic myelogenous leukemia in blastic phase, juvenile chronic myelogenous leukemia, acute leukemia or lymphoma in relapse, refractory anemia with excess blasts in transformation, and myeloma. Solid tumors and nonhematologic diseases were also classified as high-risk diseases.

‡Overall-risk groups were determined based on the risk index developed by Armand et al, which includes disease risk, stage risk, and cytogenetic data for AML and MDS. The poor and very poor MDS cytogenetic risk categories defined by Deeg et al³³ were grouped as high-risk disease, and all other categories were grouped as intermediate-risk disease.

§Myeloablative regimens were categorized based on the dose of total-body irradiation used (12 Gy or 12 Gy). All patients in the nonmyeloablative group received 2 Gy of total-body irradiation.

|| The relative change in hazard ratio for each decrease in FEV1 by 10%.

¶Based on estimated glomerular filtration rate (eGFR) mL/min per 1.73 m², using the chronic kidney disease epidemiology collaboration (CKD-EPI) formula for estimating creatinine clearance. An eGFR of ≥90 mL/min per 1.73 m² was considered normal, from 60 to 89.9 mL/min per 1.73 m² was considered to be mildly decreased, and <60 mL/min per 1.73 m² as moderately to severely decreased.

outcomes of those with HCT-CI comorbidities (~26%) but considered them to be score 0 per the SCI. If, collectively, these patients were to have increased NRMs compared with patients with HCT-CI scores of 0, then it would not be advantageous to these patients to be categorized as being at low risk. Although a simplified model sounds attractive from a usage perspective, what we really need is a comprehensive model that appraises all risks for all patients.

Comorbidity evaluation per the HCT-CI is, however, only a part of a larger picture when it comes to pretransplant prognostic evaluation (Figure 2). It is specific for the prediction of NRM but it can also be informative for OS. It is not meant to be the sole model for accurate actuarial prognostication but, rather, to be used in combination with disease-specific models for optimum prediction of outcomes. In addition to the prognostic benefit, it allows design of trials that can target patients who are medically infirm, identified by HCT-CI scores, with novel supportive care methods (NCT03870750). It is not perfect, but efforts should be better directed toward using objective biomarkers or genetic indicators of NRM risks rather than simply eliminating some comorbidities.

We cannot ignore chronological age because increasing age could affect patient preferences and goals of care (cure, length of life, and quality of life).⁶² It could also play a role in donor availability and physicians' comfort in using higher- vs lower-intensity regimens. Yet, large studies have shown the lack of association between chronological age with NRM or OS, when other factors were taken into account.^{46,63,64}

The KPS scale adds important information to prognostication based on comorbidities.^{50,65} It is still widely used by physicians⁶¹ and frequently used to select patients for transplant.⁶⁶ However, there has been an increasing interest in some geriatric assessment domains. Gait speed (4-meter walk test) seems to be the easiest and most efficient geriatric prognostic element in allogeneic HCT^{29,54} and overall health.⁶⁷ Hence, it is highly recommended to consider it together with comorbidity burden (HCT-CI) and performance status (KPS) when evaluating patients in the clinic.

ADL, IADL, cognitive impairment, social vulnerability, quality of life impairments, and depressive symptoms all seem to play important additional roles in selecting patients for transplant and predicting risks of mortality and morbidity.^{29,54,66,68-70} The comprehensive health assessment risk model is designed to analyze the potential additional prognostic impacts of age, HCT-CI, KPS, physical function, IADL, falls, gait speed, cognition, depression, number of medications, weight loss, albumin level, and C-reactive protein level to best predict NRM among older patients (NCT03992352).

Other important prognostic factors

Although HCT-CI, KPS, and gait speed, together with the refined DRI (R-DRI) can provide a wealth of prognostic information, they need to be complemented. As indicated earlier, each tool alone provides a maximum predictive power of 0.65 (with 0.5 and 1.0 indicating the least and best prediction, respectively) of a c-statistic estimate. Using these tools together should improve our ability to construct a reliable actuarial prognostication before HCT. We then need to consider the conditioning regimen that should be used for a specific patient.

This is done while recognizing the opposing impact of each regimen intensity on the chances of developing toxicities, experiencing relapse, or succumbing to NRM.⁷¹ Likewise, we have to also consider the donor type and magnitude of HLA-matching that are available to the patient, with an understanding of the risks of acute and/or chronic graft-versus-host disease (GVHD) that come with each choice, and the subsequent impacts on morbidity and mortality.⁷²⁻⁸⁰ Finally, physician-educated and experience-based articulation of all this information is required to put the prognostic information into context.

Global models

Global models are designed from factors that cover patient-, disease-, and, sometimes, transplant-related variables. The earliest model was the EBMT model that was modified (mEBMT) to take into account changes in practice, such as the increasing age of transplant recipients.⁸¹ It has an AUC of 0.630 for survival.¹⁰ The assessment of mortality (PAM) model comprises 8 different clinical variables (Table 1) reflecting all 3 prognostication sectors.⁹ In its original study, PAM had c-statistic estimates for prediction of survival that ranged from 0.69 to 0.76. When the investigators retested the model in a more recent patient cohort, its performance decreased to 0.62. The PAM investigators then attempted to modify it but, similar to that observed with the HCT-CI, modifications only improved the model's performance slightly from 0.62 to 0.63.¹⁴ The representation of comorbidity burdens and disease cytomic features in the revised-PAM are limited, making the use of independent and more specific tools, such as the HCT-CI and the R-DRI, more prognostically attractive. Yet, using either the mEBMT or revised-PAM could be considered as a supplement to other models to verify potential prognoses (Table 2). However, for the greater part, these models are only academically interesting for analysis of patient data. A similar assertion applies to models such as the NRM-J (c-statistic, 0.67).¹⁵

The Endothelial Activation and Stress Index takes a different approach, focusing on the biomarker equation [(creatinine × lactate dehydrogenase) ÷ thrombocytes] to predict mortality after HCT. It was validated to have a c-statistic estimate of 0.629 for 1-year survival,⁸² but it has the benefit of predicting specific HCT-related complications, such as microangiopathy.

Other global models include those developed in the settings of patients with AML but were used in the field of HCT as well. The treatment-related mortality model (TRM) was shown to discriminate NRM with a c-statistic of 0.661.¹⁷ Combining the TRM with HCT-CI/age and PAM did not move that estimate beyond 0.697, highlighting the limited benefit from statistically summing these models. Interestingly, the TRM was reported to have a c-statistic of 0.76 for predicting early death after induction therapy for AML, but a follow-up validation study found that estimate to be lower at 0.64.⁸³ The AML-CM has c-statistic estimates for survival rates ranging from 0.719 to 0.728,^{18,19} and it was validated to show relatively similar discriminative power in an independent cohort (0.70)⁸³ and was used in comparing HCT vs non-HCT.²⁹

Combination models

Some studies have attempted to combine the HCT-CI with other models. The HCT-CI/IADL combines the HCT-CI and IADL as 2 different measures of patient-specific risks.⁵⁴ It was

Table 2. Prognostication tools in HSCT: classification and recommendations for clinical use

| | Prognostication sector | Prognostication area | Model/factor | Specific diseases |
|--------------------|------------------------|-------------------------|---------------------------------------------|-------------------|
| Required | Patient factors | Comorbidity burden | HCT-CI | — |
| | | Performance status | KPS | — |
| | Disease factors | Disease risk index | R-DRI | — |
| | Transplant factors | Donor | Type and HLA-matching | — |
| | | Regimens | Conditioning intensity and GVHD prophylaxis | — |
| | Physician | Subjective | Educated and experience-based perception | — |
| Highly recommended | Patient factors | Gait speed | 4-meter walk test | — |
| | Disease factors | Disease-specific models | ELN-2022 | AML |
| | | | IPSS-M | MDS |
| | | | MIPSS70+ | Primary MF |
| | | | CPSS-Mol | CMML |
| | | MRD | PCR or multiparameter flow cytometry | AML or ALL |
| Considered | Patient factors | Age | Calendar | — |
| | | Geriatric assessment | IADL | — |
| | | | ADL | — |
| | | | Depression | — |
| | | | Cognitive impairments | — |
| | | | Social vulnerability | — |
| | Mixed factors | Global models | Revised PAM | — |
| | | | Modified EBMT | — |
| | | | EASIX | — |
| Academic interest | Patient factors | Combination models | HCT-CI/IADL | — |
| | Mixed factors | | HCT-CR | — |
| | | | HCT-CI/EBMT | — |
| | Mixed factors | Global models | TRM | — |
| | | | AML-CM | — |
| | | | NRM-J | — |

CPSS-Mol, Chronic Myelomonocytic Leukemia–Specific Prognostic Scoring System with Molecular Features; IPSS-M, International Prognostic Scoring System (Molecular); MF, myelofibrosis; MIPSS70+, Mutation and Karyotype-Enhanced International Prognostic Scoring System; PCR, polymerase chain reaction.

combined with the DRI to form the HCT–composite risk (HCT-CR) with a c-statistic estimate of only 0.62²⁰ and combined with the EBMT with estimate of 0.63 and 0.662 to predict NRM and OS, respectively.⁵³ These models do not show improved performance over single models and they lose the benefit of the granularity of individual weights of each model separately. It is advisable to use prognostic models like the HCT-CT and the R-DRI separately in the clinic, augmented by physician perception, as detailed earlier. Combination models will remain of academic interest and proof that not much is gained by combining models.

What are the main problems behind current study interpretations?

One problem highlighted by many of validations studies is the lack of distinction made between models representing different prognostic sectors, for example, comparing the

performance of a patient-based model like HCT-CI with that of a disease-specific model like the DRI, or comparing either model with a global model like the PAM, mEBMT, or TRM model.^{17,84} These comparisons are not helpful because they simply assume that 1 model can replace another. We need at least 1 model per prognostic area to achieve the most successful actuarial prognostication (Table 3).

It is also obvious that we can achieve a maximum of ~0.7 (but in many times <0.65) power of prediction per c-statistic estimate regardless of modifying definitions of some models, selecting certain components of a model and ignoring the others, combining models, or combining all data that are available in the patient records into a mixed model. These efforts could not produce a reliable, valid, and more efficient tool than, for example, using the HCT-CI, for NRM, and R-DRI for relapse-related mortality. Although academically investigators might be interested in continuing similar efforts, the eventual yield in the clinic is modest.

Table 3. Recommendations for future prognostic studies

| Approaches | Topic | Area | Type | Specifics |
|------------|-------------------------|----------------------------------|-------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------|
| To do | Statistical approaches | Model development and validation | Methods used | C-statistic/AUC |
| | | | | Positive and negative predictive values |
| | | | | Brier scores |
| | | | | Net Reclassification Index |
| | New investigation areas | Novel prognostic factors | Patient factors | Micro-RNAs |
| | | | | SNPs |
| | | | | Organ-specific biomarkers |
| | | | | Sarcopenia/muscle mass |
| | | | | Validated predictive and brief battery of geriatric assessment tools |
| | | Newly designed models | Patient factors | CHARM |
| | | | | DRI updated with molecular and MRD data |
| | | | | Appraisal of complexity of different donor types and HLA-matching degrees |
| | | Novel statistical methods | Cubic splines | |
| | | | | The least absolute shrinkage and selection operator and object-oriented regression FIS GBM Bayesian belief networks Markov models |
| | | | | Principal component analysis and joint decomposition regression |
| | | Novel methodological approaches | Reversibility of prognostic factors | Designing dynamic models that provide different prognostic estimates depending on timing in patient's treatment journey |
| | | | | Decision curve analysis |
| | | | | Net benefit evaluations for prediction models |
| | New specialty | Prognostication | Oncology—palliative | Trained investigators/MDs |
| Not to do | Validation | Studies | Model performances | Comparing models from different prognostic areas |
| | | | | Small sample studies |
| | | | | Different diagnoses |
| | | | | Different transplant settings |
| | Practice | Clinical | Counseling | Ignoring prognostic data |
| | | | | Using physician perception alone |
| | | | | Unilateral decision in complicated situations |

CHARM, comprehensive health assessment risk model; FIS, fuzzy inference system; GBM; gradient boosting machine; SNPs, single nucleotide polymorphisms.

Many of the validation studies do not even provide a c-statistic estimate as the least required measure of model performance. Many of the studies also have small patient cohorts, heterogeneous patient characteristics, and different treatments from the those in the original study for a model.

Moreover, it is important to keep in mind that the balance between undesirable and desirable outcomes cannot be totally mathematically decided. When using prognostic models, we have to weigh in clinically meaningful differences (eg, survival differences of 5% vs 15%) in outcomes and acceptable

thresholds for the kind of decision making (eg, transplant vs no transplant, high-dose vs reduced-intensity regimens, etc) that we are managing. Two previous studies, for example, suggested that an improvement in progression-free survival of at least 10% based on an individual patient's risks are required to justify a transplant.^{85,86}

Future directions and conclusions

In summary, since 2005, the knowledge about prognostic assessment before allogeneic HCT improved significantly through the use of a set of new models. We can now use the HCT-CI (the combined age/augmented HCT-CI) to diagnose comorbidity burden, KPS to diagnose function, and, maybe,

also gait speed to diagnose frailty; and, together, get a relatively good estimate of risks of NRM. We can also use the R-DRI alone or together with disease-specific models and knowledge about MRD to calculate a potentially accurate risk of relapse and risk of relapse-related mortality. These major, reliable, and frequently validated tools can help us make the best decision in the clinic on whether transplantation is the right choice for a patient, and, if so, what conditioning regimen intensity, donor type, and HLA-level matching would work best. An experienced and well-educated physician or, even better, a group of physicians (especially in complicated cases) can put all of this in perspective to make the best informed decision (Figure 2; Table 2). We can also use the aforementioned models to target patients in clinical trials to further improve supportive care;

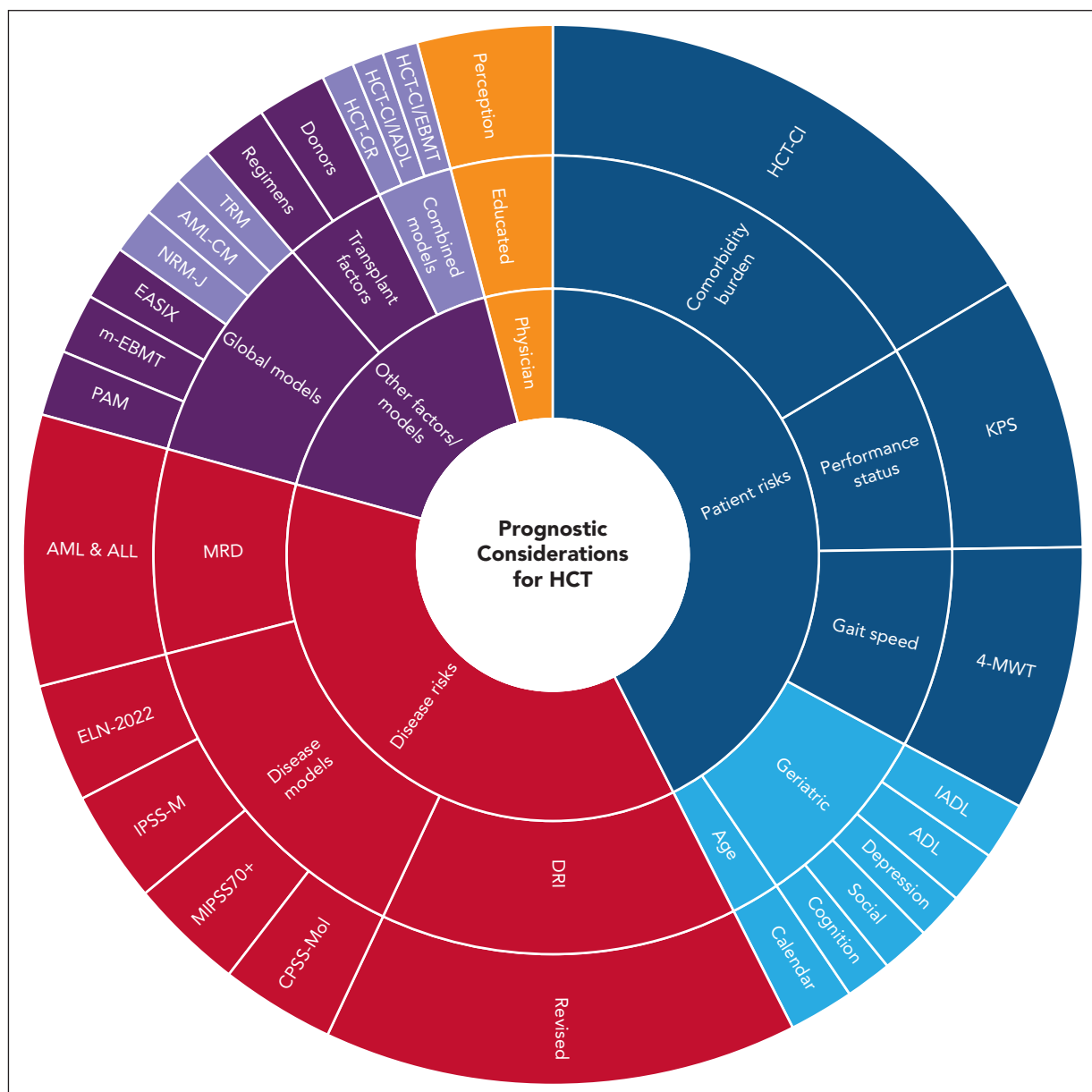


Figure 2. Transplantation prognostic pie chart. Schematic demonstration of different components of actuarial prognostic assessment before allogeneic HCT. Red indicates disease-specific risk models/factors, blue indicates patient-specific risk models/factors, purple indicates other models/factors, and orange indicates the physician's perception. Solid colors indicate validated and clinically important tools/factors that should be used frequently in the clinic for prognostic assessment; faded colors indicate academically interesting but clinically less used tools/factors, some of which require further future study to further enhance interest such as geriatric assessment.

introduce novel interventions; use different, less toxic or more effective conditioning regimens; and design randomized trials to answer the question of whether HCT is the best option for a specific group of patients with identified risks.

However, beyond that, if we only use the factors, information, and methods we have at hand, it seems that we have reached a dead-end regarding the improvement of prognostication in HCT.

Nonetheless, there is much room for improvement. To better predict NRM when considering patient risk factors, there are promising areas that call for further efforts. Molecular biomarkers such as genetic variants are reportedly associated with different HCT-related morbidities, such as infections or GVHD,^{87,88} but also mortality.⁸⁹ There is also interest in the non-HLA genetic variants in unrelated HCT.⁹⁰ Yet, ensuring replication of results remains essential before we are able to incorporate this information in future prognostic models.⁹¹ Investigating micro-RNAs that are known to regulate gene expression as molecular biomarkers for comorbidities and other health impairments could prove to be powerfully prognostic for outcomes.⁹² If we are to truly improve the performance of the HCT-CI, we must use organ-specific and validated biomarkers (eg, galectin-3 and soluble suppression of tumorigenicity 2)⁹³ or tests (eg, VO2 max⁹⁴) specific to each organ comorbidity. This should also further improve the model interrater reliability and avoid lower performance in large registry studies.⁹⁵ There is also interest in incorporating sarcopenia (skeletal muscle wasting) as a predictor of morbidity, mortality, and increased hospital costs^{96,97} in the current patient-driven risk models. Added to that is the interest in endothelial predictors of the pathogenesis of GVHD and their possible direct associations with mortality.⁹⁸ All these potential future efforts require significant funding resources for accurate, successful, and reliable replication and validation and for a potential model with a discriminative power of ≥ 0.8 .

The area of incorporating cancer-specific genetic mutations is already very active for each myeloid malignancy.^{35-37,39} It remains to be seen whether this information can be added to the current R-DRI to aid in selecting the best transplant strategy but also in formulating successful post-HCT maintenance plans. If we are to achieve discriminative power of ≥ 0.8 in diagnosing risks of relapse, then incorporating MRD is imperative.

We also need to revolutionize our statistical approaches and methods. We should shy away from reliance on the *P* value, which is greatly influenced by the number of events or patients, and we should frequently use model performance measures such as c-statistic estimates and/or AUC when developing or validating a model. However, the c-statistic has its own limitation because it relies on the distribution of the important variables in the model.⁹⁹ Therefore, patient sample size and characteristics could play a major role in validation of results. Moreover, the c-statistic does not measure calibration. Hence, we need to report positive and negative predictive values for each model. Other methods that can be used include Brier scores that can provide estimates of prediction accuracy^{100,101} and the net reclassification index that quantifies the improvement in prognostication by adding new biomarkers.¹⁰²

Novel statistical methods are also needed. When building a model, we must recognize that the association between a specific prognostic factor and a specific outcome is not always linear (eg, age and NRM). Hence, using restricted cubic splines are probably the best approach to account for complex and nonlinear associations (best reviewed in Gauthier et al¹⁰³). Machine learning approaches^{104,105} should also be considered because they are now at the forefront of predictive analytics: exploring large and complex health care data sets and generating models for predictive medicine/health. When facing larger numbers of predictors, traditional statistical methods are prone to overfitting (ie, fitting noise), and as a result can lead to degraded performance in validation studies. Approaches such as the least absolute shrinkage and selection operator¹⁰⁶ can be used to select predictors, which are then complemented by object-oriented regression techniques for building a reproducible predictive model. Other artificial intelligence-based methods such as the fuzzy inference system, gradient boosting machine, Bayesian belief networks, and Markov models are of interest for future model development.^{107,108} When confronted with a data set in which the number of variables is larger than the patient sample size (genetic variant data), using principal component analysis and joint decomposition would be advisable.¹⁰⁹

Novel methodological approaches are also of interest. For example, many criteria of prognostic models could change over time depending on at what stage a patient might be along their treatment journey.³¹ Updating current models or creating new models that have dynamic characteristics (ie, providing different prognostic estimates per different times) adaptable to different stages (ie, diagnosis, after induction therapy, before transplant, early after transplant, etc) would greatly improve our abilities to generate accurate decision making.

An example of important posttransplant prognostic criteria would be biomarkers such as the regenerating islet-derived 3 α and the suppressor of tumorigenesis 2. The Mount Sinai Acute GVHD International Consortium has combined both biomarkers into an algorithm that can predict 6-month NRM, acute GVHD, and response to treatment of acute GVHD.^{110,111}

With the relatively large number of currently available models, the need for decision-analytic measures to summarize the performance of models in supporting decision making is heightened. Decision curve analysis (plot of net benefit against threshold probability), introduced in 2006,¹¹² and the net benefit (a weighted sum of true and false positives, the weighting accounting for differential consequences of each) evaluation¹¹³⁻¹²⁰ are recommended by the transparent reporting of a multivariable prediction model for individual prognosis or diagnosis guidelines for prediction models.¹²¹ Similarly, computer clinical decision support system that automates personalized clinical care can be of extreme help.¹²²

Finally, training new investigators in the field of prognostication is important. Importantly, a multidisciplinary collaboration of experts from oncology, palliative care, epidemiology, geriatric medicine, and biostatistics is required to improve prognostication beyond its current status.

Acknowledgments

The author is extremely grateful to the patients who allowed care toward them and who participated in clinical research that made this article possible. The author is especially grateful to Helen Crawford for help in preparing this manuscript and Marivic Jimenez for managing grant administrative tasks. The author is extremely grateful to the research staff who helped in daily activities of studies, including Lori Garrett, Ayah Idris, Kimberly Wands, Judy Allen, and Peg Boyle. The author is tremendously grateful to all the colleagues who provided insights about the topic of this article over the years. These include, among a larger list, the late Elihu Estey as well as Rainer Storb, Barry Storer, Ted Gooley, Benda Sandmaier, David Maloney, Fred Appelbaum, Paul Martin, Stephanie Lee, Joachim Deeg, Jordan Gauthier, Andrew Artz, Brent Logan, Jeannine McCune, Roland Walter, Megan Othus, Lue Zhao, Richard Maziarz, Marcelo Pasquini, and Mary Horowitz.

Research reported in this article and the author were funded, in part, by a Patient-Centered Outcome Research Institute award (CE-1304-7451), a Research Scholar Grant from the American Cancer Society (RSG-13-084-01-CPHPS), an American Society of Hematology Bridge Award, and, grant R01 CA227092 from the National Cancer Institute, National Institute of Health.

Funding organizations had absolutely no role in the design or conduct of the study; collection, management, analysis, or interpretation of the

data; or preparation, review, or approval of the manuscript. The statements and findings in this article are solely the responsibility of the author and do not necessarily represent the views of any funding organization.

Authorship

Contribution: M.L.S. conceptualized the study and drafted the manuscript.

Conflict-of-interest disclosure: M.L.S. reports consultancy for and receiving an honorarium from Jazz Pharmaceuticals.

ORCID profile: M.L.S., [0000-0001-5260-1981](https://orcid.org/0000-0001-5260-1981).

Correspondence: Mohamed L. Sorror, Clinical Research Division (D5-285), Fred Hutchinson Cancer Center, 100 Fairview Ave North, Seattle, WA 98109-1024; email: msorror@fredhutch.org.

Footnote

Submitted 22 November 2022; accepted 8 February 2023; prepublished online on *Blood* First Edition 17 February 2023. <https://doi.org/10.1182/blood.2022017999>.

REFERENCES

- Sorror ML, Maris MB, Storb R, et al. Hematopoietic cell transplantation (HCT)-specific comorbidity index: a new tool for risk assessment before allogeneic HCT. *Blood*. 2005;106(8):2912-2919.
- Extermann M. Measuring comorbidity in older cancer patients [review]. *Eur J Cancer*. 2000;36(4):453-471.
- Gooley TA, Chien JW, Pergam SA, et al. Reduced mortality after allogeneic hematopoietic-cell transplantation. *N Engl J Med*. 2010;363(22):2091-2101.
- McDonald GB, Sandmaier BM, Mielcarek M, et al. Survival, non-relapse mortality, and relapse-related mortality after allogeneic hematopoietic cell transplantation: comparing 2003-2007 vs 2013-2017 cohorts. *Ann Intern Med*. 2020;172(4):229-239.
- Auletta JJ, Kou J, Chen M, Shaw BE. Current use and outcome of hematopoietic stem cell transplantation: CIBMTR US summary slides, 2021. 2021. Accessed 2 November 2022. Available at: <https://www.cibmtr.org/ReferenceCenter/SlidesReports/SummarySlides/Pages/index.aspx>
- Deeg HJ, Sandmaier BM. Who is fit for allogeneic transplantation? *Blood*. 2010;116(23):4762-4770.
- Sorror ML, Sandmaier BM, Storer BE, et al. Allogeneic hematopoietic cell transplantation (HCT) in the eighth decade of life: how much does age matter? *Biol Blood Marrow Transplant*. 2017;23(3):S98-S99.
- Muffy L, Pasquini MC, Martens M, et al. Increasing use of allogeneic hematopoietic cell transplantation in patients aged 70 years and older in the United States. *Blood*. 2017;130(9):1156-1164.
- Parimon T, Au DH, Martin PJ, Chien JW. A risk score for mortality after allogeneic hematopoietic cell transplantation. *Ann Intern Med*. 2006;144(6):407-414.
- Gratwohl A. The EBMT risk score. *Bone Marrow Transplant*. 2012;47(6):749-756.
- Armand P, Gibson CJ, Cutler C, et al. A disease risk index for patients undergoing allogeneic stem cell transplantation. *Blood*. 2012;120(4):905-913.
- Barba P, Martino R, Perez-Simon JA, et al. Combination of the hematopoietic cell transplantation comorbidity index and the european group for blood and marrow transplantation score allows a better stratification of high-risk patients undergoing reduced-toxicity allogeneic hematopoietic cell transplantation. *Biol Blood Marrow Transplant*. 2014;20(1):66-72.
- Armand P, Kim HT, Logan BR, et al. Validation and refinement of the disease risk index for allogeneic stem cell transplantation. *Blood*. 2014;123(23):3664-3671.
- Au BKC, Gooley TA, Armand P, et al. Reevaluation of the pretransplant assessment of mortality score after allogeneic hematopoietic transplantation. *Biol Blood Marrow Transplant*. 2015;21(5):848-854.
- Yanada M, Konuma T, Mizuno S, et al. Predicting non-relapse mortality following allogeneic hematopoietic cell transplantation during first remission of acute myeloid leukemia. *Bone Marrow Transplant*. 2021;56(2):387-394.
- Shouval R, Fein JA, Cho C, et al. The Simplified Comorbidity Index: a new tool for prediction of nonrelapse mortality in allo-HCT. *Blood Adv*. 2022;6(5):1525-1535.
- Zarling LC, Othus M, Sandmaier BM, et al. Utility of the treatment-related mortality (TRM) score to predict outcomes of adults with acute myeloid leukemia undergoing allogeneic hematopoietic cell transplantation. *Leukemia*. 2022;36(6):1563-1574.
- Sorror ML, Storer BE, Nyland J, Estey EH. Revised acute myeloid leukemia composite model using the 2017 european leukemia net risk classification. *JAMA Oncol*. 2019;5(7):1062-1064.
- Sorror ML, Storer BE, Fathi AT, et al. Development and validation of a novel acute myeloid leukemia-composite model to estimate risks of mortality. *JAMA Oncol*. 2017;3(12):1675-1682.
- Kongtim P, Parmar S, Milton DR, et al. Impact of a novel prognostic model, hematopoietic cell transplant-composite risk (HCT-CR), on allogeneic transplant outcomes in patients with acute myeloid leukemia and myelodysplastic syndrome. *Bone Marrow Transplant*. 2019;54(6):839-848.
- Glare P, Christakis NA. Overview: advancing the clinical science of prognostication. In: Glare P, Christakis NA, eds. *Prognosis in Advanced Cancer*. Oxford University Press; 2008:3-12.
- Christakis NA, Iwashyna TJ. Attitude and self-reported practice regarding prognostication in a national sample of internists. *Arch Intern Med*. 1998;158(21):2389-2395.
- Lamont EB. Foreseeing: formulating an accurate prognosis. In: Glare P, Christakis NA, eds. *Prognosis in Advanced Cancer*. Oxford University Press; 1981:25-32.
- Fries JF, Ehrlich GE. *Prognosis: Contemporary Outcomes of Disease*. The Charles Press Publishers; 1981:565.
- Christakis NA, Lamont EB. Extent and determinants of error in doctors' prognoses in terminally ill patients: prospective cohort study. *BMJ*. 2000;320(7233):469-472.

26. Mackillop WJ, Quirt CF. Measuring the accuracy of prognostic judgments in oncology. *J Clin Epidemiol*. 1997;50(1):21-29.
27. Chu C, Anderson R, White N, Stone P. Prognosticating for adult patients with advanced incurable cancer: a needed oncologist skill. *Curr Treat Options Oncol*. 2020;21(1):5.
28. Lamont EB, Christakis NA. Prognostic disclosure to patients with cancer near the end of life. *Ann Intern Med*. 2001;134(12):1096-1105.
29. Sorror ML, Gooley TA, Storer BE, et al. An 8-year pragmatic observation evaluation of the benefits of allogeneic HCT in older and medically infirm AML patients. *Blood*. 2023;141(3):295-308.
30. Simpkin AL, Schwartzstein RM. Tolerating uncertainty - The next medical revolution? *N Engl J Med*. 2016;375(18):1713-1715.
31. Estey E, Gale RP. How good are we at predicting the fate of someone with acute myeloid leukaemia? *Leukemia*. 2017;31(6):1255-1258.
32. Maltoni M, Caraceni A, Brunelli C, et al. Prognostic factors in advanced cancer patients: evidence-based clinical recommendations—a study by the steering committee of the European association for palliative care. *J Clin Oncol*. 2005;23(25):6240-6248.
33. Deeg HJ, Scott BL, Fang M, et al. Five-group cytogenetic risk classification, monosomal karyotype, and outcome after hematopoietic cell transplantation for MDS or acute leukemia evolving from MDS. *Blood*. 2012;120(7):1398-1408.
34. Harrell FE Jr, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. *Stat Med*. 1984;3(2):143-152.
35. Dohner H, Wei AH, Appelbaum FR, et al. Diagnosis and management of AML in adults: 2022 recommendations from an international expert panel on behalf of the ELN. *Blood*. 2022;140(12):1345-1377.
36. Bernard E, Tuechler H, Greenberg PL, et al. Molecular international prognostic scoring system for myelodysplastic syndromes. *NEJM Evidence*. 2022;1(7):EVID0a2200008.
37. Tefferi A, Guglielmelli P, Lasho TL, et al. MIPSS70+ version 2.0: mutation and karyotype-enhanced international prognostic scoring system for primary myelofibrosis. *J Clin Oncol*. 2018;36(17):1769-1770.
38. Gagelmann N, Ditschkowski M, Bogdanov R, et al. Comprehensive clinical-molecular transplant scoring system for myelofibrosis undergoing stem cell transplantation. *Blood*. 2019;133(20):2233-2242.
39. Elena C, Galli A, Such E, et al. Integrating clinical features and genetic lesions in the risk assessment of patients with chronic myelomonocytic leukemia. *Blood*. 2016;128(10):1408-1417.
40. Yin JAL, Grimwade D. Minimal residual disease evaluation in acute myeloid leukaemia. *Lancet*. 2002;360(9327):160-162.
41. Walter RB, Gooley TA, Wood BL, et al. Impact of pretransplantation minimal residual disease, as detected by multiparametric flow cytometry, on outcome of myeloablative hematopoietic cell transplantation for acute myeloid leukemia. *J Clin Oncol*. 2011;29(9):1190-1197.
42. Uzunel M, Mattsson J, Jaksch M, Remberger M, Ringden O. The significance of graft-versus-host disease and pretransplantation minimal residual disease status to outcome after allogeneic stem cell transplantation in patients with acute lymphoblastic leukemia. *Blood*. 2001;98(6):1982-1984.
43. Schuurhuis GJ, Heuser M, Freeman S, et al. Minimal/measurable residual disease in AML: a consensus document from the European LeukemiaNet MRD Working Party. *Blood*. 2018;131(12):1275-1291.
44. Heuser M, Freeman SD, Ossenkoppele GJ, et al. 2021 update on MRD in acute myeloid leukemia: a consensus document from the European LeukemiaNet MRD Working Party. *Blood*. 2021;138(26):2753-2767.
45. Sorror ML, Martin PJ, Storb RF, et al. Pretransplant comorbidities predict severity of acute graft-versus-host disease and subsequent mortality. *Blood*. 2014;124(2):287-295.
46. Sorror ML, Sandmaier BM, Storer BE, et al. Long-term outcomes among older patients following nonmyeloablative conditioning and allogeneic hematopoietic cell transplantation for advanced hematologic malignancies. *JAMA*. 2011;306(17):1874-1883.
47. Sorror ML, Yi JC, Storer BE, et al. Association of pre-transplant comorbidities with long-term quality of life (QOL) among survivors after allogeneic hematopoietic cell transplantation (HCT). *Biol Blood Marrow Transplant*. 2013;19(2):S153.
48. Decook L, Chang YH, Slack J, et al. Association of hematopoietic cell transplantation-specific comorbidity index with resource utilization after allogeneic transplantation. *Bone Marrow Transplant*. 2017;52(7):998-1002.
49. Sorror ML. How I assess comorbidities prior to hematopoietic cell transplantation. *Blood*. 2013;121(15):2854-2863.
50. Sorror M, Storer B, Sandmaier BM, et al. Hematopoietic cell transplantation-comorbidity index and Karnofsky performance status are independent predictors of morbidity and mortality after allogeneic nonmyeloablative hematopoietic cell transplantation. *Cancer*. 2008;112(9):1992-2001.
51. Sorror ML, Storb RF, Sandmaier BM, et al. Comorbidity-age index: a clinical measure of biologic age before allogeneic hematopoietic cell transplantation. *J Clin Oncol*. 2014;32(29):3249-3256.
52. Vaughn JE, Storer BE, Armand P, et al. Design and validation of an augmented hematopoietic cell transplantation-comorbidity index comprising pretransplant ferritin, albumin, and platelet count for prediction of outcomes after allogeneic transplantation. *Biol Blood Marrow Transplant*. 2015;21(8):1418-1424.
53. Elsayy M, Storer BE, Sorror ML. To combine or not to combine: optimizing risk assessment before allogeneic hematopoietic cell transplantation (letter). *Biol Blood Marrow Transplant*. 2014;20(9):1455-1456.
54. Muffy LS, Kocherginsky M, Stock W, et al. Geriatric assessment to predict survival in older allogeneic hematopoietic cell transplantation recipients. *Haematologica*. 2014;99(8):1373-1379.
55. Elsayy M, Sorror ML. Up-to-date tools for risk assessment before allogeneic hematopoietic cell transplantation. *Bone Marrow Transplant*. 2016;51(10):1283-1300.
56. Raimondi R, Tosetto A, Oneto R, et al. Validation of the hematopoietic cell transplantation-specific comorbidity index: a prospective, multicenter GITMO study. *Blood*. 2012;120(6):1327-1333.
57. Sorror ML, Logan BR, Zhu X, et al. Prospective validation of the predictive power of the hematopoietic cell transplantation comorbidity index: a center for international blood and marrow transplant research study. *Biol Blood Marrow Transplant*. 2015;21(8):1479-1487.
58. Xhaard A, Porcher R, Chien JW, et al. Impact of comorbidity indexes on non-relapse mortality. *Leukemia*. 2008;22(11):2062-2069.
59. Castagna L, Furst S, Marchetti N, et al. Retrospective analysis of common scoring systems and outcome in patients older than 60 years treated with reduced-intensity conditioning regimen and alloSCT. *Bone Marrow Transplant*. 2011;46(7):1000-1005.
60. Guilfoyle R, Demers A, Bredeson C, et al. Performance status, but not the hematopoietic cell transplantation comorbidity index (HCT-CI), predicts mortality at a Canadian transplant center. *Bone Marrow Transplant*. 2009;43(2):133-139.
61. Mishra A, Preussler JM, Bhatt VR, et al. Breaking the age barrier: physicians' perceptions of candidacy for allogeneic hematopoietic cell transplantation in older adults. *Transplant Cell Ther*. 2021;27(7):e1617-e17.
62. Dhawale T, Steuten LM, Deeg HJ. Uncertainty of physicians and patients in medical decision making. *Biol Blood Marrow Transplant*. 2017;23(6):865-869.
63. Lim Z, Brand R, Martino R, et al. Allogeneic hematopoietic stem-cell transplantation for

- patients 50 years or older with myelodysplastic syndromes or secondary acute myeloid leukemia. *J Clin Oncol*. 2010;28(3):405-411.
64. McClune BL, Weisdorf DJ, Pedersen TL, et al. Effect of age on outcome of reduced-intensity hematopoietic cell transplantation for older patients with acute myeloid leukemia in first complete remission or with myelodysplastic syndrome. *J Clin Oncol*. 2010;28(11):1878-1887.
 65. Deschler B, Binek K, Ihorst G, et al. Prognostic factor and quality of life analysis in 160 patients aged > or = to 60 years with hematologic neoplasias treated with allogeneic hematopoietic cell transplantation. *Biol Blood Marrow Transplant*. 2010;16(7):967-975.
 66. Sorror ML, Storer BE, Gerds AT, et al. Limitations to receiving allogeneic hematopoietic cell transplantation for treatment of acute myeloid leukemia: A large multi-center prospective longitudinal observational study [abstract]. *Blood*. 2018;132(Suppl 1):1388.
 67. Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. *J Am Med Assoc*. 2011;305(1):50-58.
 68. Deschler B, Ihorst G, Schnitzler S, Bertz H, Finke J. Geriatric assessment and quality of life in older patients considered for allogeneic hematopoietic cell transplantation: a prospective risk factor and serial assessment analysis. *Bone Marrow Transplant*. 2018;53(5):565-575.
 69. Bhandari R, Teh JB, He T, et al. Social vulnerability and risk of non-relapse mortality after allogeneic hematopoietic cell transplantation. *J Natl Cancer Inst*. 2022;114(11):1484-1491.
 70. Lacy M, Fong M, Bolton C, Maranzano M, Bishop M, Artz A. Cognitive functioning of older adults prior to hematopoietic stem cell transplantation. *Bone Marrow Transplant*. 2021;56(10):2575-2581.
 71. Bacigalupo A, Ballen K, Rizzo D, et al. Defining the intensity of conditioning regimens: working definitions. *Biol Blood Marrow Transplant*. 2009;15(12):1628-1633.
 72. Gupta V, Tallman MS, He W, et al. Comparable survival after HLA-well-matched unrelated or matched sibling donor transplantation for acute myeloid leukemia in first remission with unfavorable cytogenetics at diagnosis. *Blood*. 2010;116(11):1839-1848.
 73. Lee SJ, Klein J, Haagenson M, et al. High-resolution donor-recipient HLA matching contributes to the success of unrelated donor marrow transplantation. *Blood*. 2007;110(13):4576-4583.
 74. Woolfrey A, Klein JP, Haagenson M, et al. HLA-C antigen mismatch is associated with worse outcome in unrelated donor peripheral blood stem cell transplantation. *Biol Blood Marrow Transplant*. 2011;17(6):885-892.
 75. Hurley CK, Woolfrey A, Wang T, et al. The impact of HLA unidirectional mismatches on the outcome of myeloablative hematopoietic stem cell transplantation with unrelated donors. *Blood*. 2013;121(23):4800-4806.
 76. Fleischhauer K, Shaw BE, Gooley T, et al. Effect of T-cell-epitope matching at HLA-DPB1 in recipients of unrelated-donor haemopoietic-cell transplantation: a retrospective study [Erratum appears in *Lancet Oncol*. 2012;13(4):e134-e135]. *Lancet Oncol*. 2012;13(4):366-374.
 77. Milano F, Gooley T, Wood B, et al. Cord blood transplant in patients with minimal residual disease. *N Engl J Med*. 2016;375(10):944-953.
 78. Fuchs EJ, O'Donnell PV, Eapen M, et al. Double unrelated umbilical cord blood vs HLA-haploidentical bone marrow transplantation: the BMT CTN 1101 trial. *Blood*. 2021;137(3):420-428.
 79. Gooptu M, Romee R, St Martin A, et al. HLA haploidentical versus matched unrelated donor transplants with post-transplant cyclophosphamide based prophylaxis. *Blood*. 2021;138(3):273-282.
 80. Kollman C, Spellman SR, Zhang MJ, et al. The effect of donor characteristics on survival after unrelated donor transplantation for hematologic malignancy. *Blood*. 2016;127(2):260-267.
 81. Hemmati PG, Terwey TH, le Coutre P, et al. A modified EBMT risk score predicts the outcome of patients with acute myeloid leukemia receiving allogeneic stem cell transplants. *Eur J Haematol*. 2011;86(4):305-316.
 82. Luft T, Benner A, Terzer T, et al. EASIX and mortality after allogeneic stem cell transplantation. *Bone Marrow Transplant*. 2020;55(3):553-561.
 83. Desprez C, Riou J, Peterlin P, et al. Comparison of scoring systems evaluating suitability for intensive chemotherapy in adults with acute myeloid leukemia-a grand ouest against leukemia (GOAL) study. *Leukemia*. 2022;36(10):2408-2417.
 84. Hegde PS, Rybicki L, Serafino S, et al. Evaluation of pre-transplant risk assessments in allogeneic hematopoietic cell transplant. *Bone Marrow Transplant*. 2022;57(6):1031-1033.
 85. Cornelissen JJ, Gratwohl A, Schlenk RF, et al. The European LeukemiaNet AML Working Party consensus statement on allogeneic HSCT for patients with AML in remission: an integrated-risk adapted approach [review]. *Nat Rev Clin Oncol*. 2012;9(10):579-590.
 86. Sorror ML, Estey E. Allogeneic hematopoietic cell transplantation for acute myeloid leukemia in older adults. *Hematology Am Soc Hematol Educ Program*. 2014;2014(1):21-33.
 87. Bochud PY, Chien JW, Marr KA, et al. Toll-like receptor 4 polymorphisms and aspergillosis in stem-cell transplantation. *N Engl J Med*. 2008;359(17):1766-1777.
 88. Martin PJ, Fan W, Storer BE, et al. Replication of associations between genetic polymorphisms and chronic graft-versus-host disease. *Blood*. 2016;128(20):2450-2456.
 89. Noren E, Verma D, Soderkvist P, et al. Single nucleotide polymorphisms in MORC4, CD14, and TLR4 are related to outcome of allogeneic stem cell transplantation. *Ann Transplant*. 2016;21:56-67.
 90. Hahn T, Wang J, Preus LM, et al. Novel genetic variants associated with mortality after unrelated donor allogeneic hematopoietic cell transplantation. *EClinicalMedicine*. 2021;40:101093.
 91. Martin PJ, Levine DM, Storer BE, Sather CL, Spellman SR, Hansen JA. Genetic associations with immune-mediated outcomes after allogeneic hematopoietic cell transplantation. *Blood Adv*. 2022;6(8):2608-2617.
 92. Sorror ML, Gooley TA, Maclean KH, et al. Pre-transplant expressions of microRNAs, comorbidities, and post-transplant mortality. *Bone Marrow Transplant*. 2019;54(7):973-979.
 93. Mueller T, Gegenhuber A, Leitner I, Poelz W, Haltmayer M, Dieplinger B. Diagnostic and prognostic accuracy of galectin-3 and soluble ST2 for acute heart failure. *Clin Chim Acta*. 2016;463:158-164.
 94. Re F, Zachara E, Avella A, et al. Dissecting functional impairment in hypertrophic cardiomyopathy by dynamic assessment of diastolic reserve and outflow obstruction: a combined cardiopulmonary-echocardiographic study. *Int J Cardiol*. 2017;227:743-750.
 95. Penack O, Peczynski C, Mohty M, et al. Association of pre-existing comorbidities with outcome of allogeneic hematopoietic cell transplantation. A retrospective analysis from the EBMT. *Bone Marrow Transplant*. 2022;57(2):183-190.
 96. Morishita S, Kaida K, Tanaka T, et al. Prevalence of sarcopenia and relevance of body composition, physiological function, fatigue, and health-related quality of life in patients before allogeneic hematopoietic stem cell transplantation. *Support Care Cancer*. 2012;20(12):3161-3168.
 97. Ando T, Fujisawa S, Teshigawara H, et al. Computed tomography-defined sarcopenia: prognostic predictor of nonrelapse mortality after allogeneic hematopoietic stem cell transplantation: a multicenter retrospective study. *Int J Hematol*. 2020;121(1):46-56.
 98. Dietrich S, Falk CS, Benner A, et al. Endothelial vulnerability and endothelial damage are associated with risk of graft-versus-host disease and response to steroid treatment. *Biol Blood Marrow Transplant*. 2013;19(1):22-27.
 99. Pencina MJ, D'Agostino RB, Sr. Evaluating discrimination of risk prediction models: the C statistic. *JAMA*. 2015;314(10):1063-1064.

100. Brier GW. Verification of forecasts expressed in terms of probability. *Mon Weather Rev*. 1950;78(1):1-3.
101. Rufibach K. Use of Brier score to assess binary predictions. *J Clin Epidemiol*. 2010; 63(8):938-939.
102. Leening MJG, Vedder MM, Wittteman JCM, Pencina MJ, Steyerberg EW. Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. *Ann Intern Med*. 2014;160(2):122-131.
103. Gauthier J, Wu QV, Gooley TA. Cubic splines to model relationships between continuous variables and outcomes: a guide for clinicians. *Bone Marrow Transplant*. 2020;55(4):675-680.
104. Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning: data mining, inference, and prediction: with 200 full-color illustrations*. Springer; 2001.
105. Yoo C, Ramirez L, Liuzzi J. Big data analysis using modern statistical and machine learning methods in medicine. *Int Neurolog J*. 2014; 18(2):50-57.
106. Tibshirani R. Regression shrinkage and selection via the Lasso. *J Roy Stat Soc B*. 1996;58(1):267-288.
107. Leclerc V, Ducher M, Ceraulo A, Bertrand Y, Bleyzac N. A clinical decision support tool to find the best initial intravenous cyclosporine regimen in pediatric hematopoietic stem cell transplantation. *J Clin Pharmacol*. 2021; 61(11):1485-1492.
108. Taheriyani M, Safaee Nodehi S, Niakan Kalhori SR, Mohammadzadeh N. A systematic review of the predicted outcomes related to hematopoietic stem cell transplantation: focus on applied machine learning methods' performance. *Expert Rev Hematol*. 2022;15(2):137-156.
109. Chen H, Caffo B, Stein-O'Brien G, et al. Two-stage linked component analysis for joint decomposition of multiple biologically related data sets. *Biostatistics*. 2022;23(4): 1200-1217.
110. Hartwell MJ, Ozbek U, Holler E, et al. An early-biomarker algorithm predicts lethal graft-versus-host disease and survival. *JCI Insight*. 2017;2(3):e89798.
111. Srinagesh HK, Ozbek U, Kapoor U, et al. The MAGIC algorithm probability is a validated response biomarker of treatment of acute graft-versus-host disease. *Blood Adv*. 2019; 3(23):4034-4042.
112. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making*. 2006;26(6): 565-574.
113. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21(1): 128-138.
114. Kerr KF, Brown MD, Zhu K, Janes H. Assessing the clinical impact of risk prediction models with decision curves: guidance for correct interpretation and appropriate use. *J Clin Oncol*. 2016;34(21): 2534-2540.
115. Fitzgerald M, Saville BR, Lewis RJ. Decision curve analysis. *JAMA*. 2015;313(4):409-410.
116. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol*. 2015;16(4):e173-e180.
117. Pepe MS, Janes H, Li CI. Net risk reclassification p values: valid or misleading? *J Natl Cancer Inst*. 2014; 106(4):dju041.
118. Holmberg L, Vickers A. Evaluation of prediction models for decision-making: beyond calibration and discrimination. *PLoS Med*. 2013;10(7):e1001491.
119. Localio AR, Goodman S. Beyond the usual prediction accuracy metrics: reporting results for clinical decision making. *Ann Intern Med*. 2012;157(4): 294-295.
120. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ*. 2016; 352:i6.
121. Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162(1): W1-W73.
122. Morris AH, Horvat C, Stagg B, et al. Computer clinical decision support that automates personalized clinical care: a challenging but needed healthcare delivery strategy. *J Am Med Inform Assoc*. 2022; 30(1):178-194.

© 2023 by The American Society of Hematology