medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Title:

Interactive web application for plotting personalized prognosis prediction curves in allogeneic hematopoietic cell transplantation using machine learning

Names of authors:

Hiroshi Okamura,[1] Mika Nakamae,[1] Shiro Koh,[1] Satoru Nanno,[1] Yasuhiro Nakashima,[1] Hideo Koh,[1] Takahiko Nakane,[1] Asao Hirose,[1] Masayuki Hino,[1] and Hirohisa Nakamae.[1]

Authors' affiliations:

[1] Hematology, Graduate School of Medicine, Osaka City University, Osaka, Japan

Contact information for correspondence: Hiroshi Okamura, M.D., Ph.D.

Hematology, Graduate School of Medicine, Osaka City University, 1-4-3 Asahi-machi, Abeno-ku, Osaka 545-8585, Japan

Phone: +81-6-6645-3881

Fax: +81-6-6645-3880

E-mail: h_okamura@med.osaka-cu.ac.jp

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Abstract

Allogeneic hematopoietic cell transplantation (allo-HCT) is a curative treatment option for malignant hematological disorders. Transplant clinicians estimate patient-specific prognosis empirically in clinical practice based on previous studies on similar patients. However, this approach does not provide objective data. The present study primarily aimed to develop a tool capable of providing accurate personalized prognosis prediction after allo-HCT in an objective manner. We succeeded in generating personalized prognosis prediction curves of 1-year overall survival (OS), progression-free survival (PFS), relapse/progression, and non-relapse mortality (NRM) adjusted for new allo-HCT candidates by eight pre-transplant factors using our interactive web application (https://predicted-os-after-transplantation.shinyapps.io/RSF_model/). A random survival forest model using the data of patients who underwent allo-HCT at our institution was applied to develop this application. To assess its predictive performance, the entire cohort (363 cases) was split into a training cohort (70%) to develop the predictive model and test cohort (30%) to confirm its performance time-sequentially. The areas under the receiver-operating characteristic curves for 1-year OS, PFS, relapse/progression, and NRM in test cohort were 0.70, 0.72, 0.73, and 0.77, respectively. In conclusion, the new web application could allow transplant

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

clinicians to inform a new allo-HCT candidate of the objective personalized prognosis prediction and facilitate decision-making.

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

INTRODUCTION

Allogeneic hematopoietic cell transplantation (allo-HCT) is an advanced therapeutic intervention required for high-risk malignant hematological disorders [1]. Various complex and intertwined pre-transplant factors affect patient prognosis after allo-HCT [2-5]. Transplant clinicians need to estimate individual prognosis after allo-HCT in clinical practice by combining not only certain indices that can stratify the prognosis after allo-HCT, such as refined Disease Risk Index (DRI-R), EBMT risk score, and Hematopoietic Cell Transplantation Comorbidity Index (HCT-CI) [6-8], but also individual factors (e.g., age, performance status [PS], disease status, or the number of transplantations) and transplant procedures (e.g., donor source or conditioning regimen). However, the predictive capacities of these prognostic indices remain suboptimal [2]. Furthermore, prognosis prediction based on the empirical integration of multiple patient-specific factors by the clinician is not always objective due to the complexity of allo-HCT treatment.

A random survival forest (RSF) is an ensemble machine learning method for right-censored time-to-event data analysis. Recently, RSF has become increasingly popular for survival data analysis in various medical fields because of its high-precision predictions [9-11]. It is possible to develop a prognosis prediction model integrating multiple factors to compute patient-specific prognosis prediction

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

curves such as survival or cumulative incidence adjusted for individual factors of a new patient using RSF [12]. Although the most popular model for survival analysis is the Cox proportional hazard (Cox PH) regression, RSF has some advantages over the Cox PH model. First, RSF can easily deal with non-linear effects, correlated parameters, and variable interaction. Second, while the Cox PH model may be limited due to the proportional hazards assumption, RSF is non-parametric and completely independent of model assumptions [11, 13]. Therefore, survival prediction analysis by RSF may provide novel insights into the allo-HCT field.

Although objective and accurate patient-specific prognosis prediction using a machine learning model can be quite informative for clinicians and patients in clinical decision-making, it has not been applied in a clinical setting. A user-friendly tool for constructing accurate personalized prognosis prediction curves after allo-HCT is lacking. Such a tool could allow transplant clinicians to inform patients of patient-specific prognosis prediction objectively derived from past patient data, and build an optimal personalized treatment strategy.

Here, we developed a web application tool with a graphical user interface (GUI) capable of plotting the personalized survival and cumulative incidence prediction curves after allo-HCT adjusted by eight patient-specific factors, which are known as prognostic predictors, using RSF and assessed their predictive performances.

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

PATIENTS AND METHODS

Patients and Definition

This study, comprising retrospective prognostic modeling, analyzed a group of consecutive patients who underwent allo-HCT at our institution between January 2008 and November 2017. All pre-transplant factors were recorded in the medical chart within 28 days prior to conditioning. The conditioning regimen containing either total body irradiation in fractionated doses greater than 8 Gy, an oral busulfan dose of 9 mg/kg or more, an intravenous busulfan dose of 7.2 mg/kg or more, or a melphalan dose of 140 mg/m^2 or more was defined as myeloablative in line with a previous report [14]. Human leukocyte antigen (HLA) compatibility was defined by DNA typing for HLA-A, HLA-B, HLA-C, and HLA-DR. This study was approved by the Osaka City University Hospital Certified Review Board.

Supportive Care

Prophylactic antibiotics, either levofloxacin or polymyxin B oral tablets, an antifungal agent, and acyclovir were routinely administered from the start of conditioning. Trimethoprim-sulfamethoxazole was also administered from the start of conditioning to 2 days prior to allo-HCT and after neutrophil engraftment to prevent

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Pneumocystis jirovecii-induced pneumonia. Granulocyte colony-stimulating factor treatment was generally initiated from the day following HCT to neutrophil engraftment. Ursodeoxycholic acid was administered from the start of conditioning to prevent sinusoidal obstruction syndrome [15].

Predictors and Outcomes

Eight pre-transplant factors known as prognostic predictors for allo-HCT, including recipients' age, DRI-R, HCT-CI, PS, donor source (related bone marrow [rBM], related peripheral blood [rPB], unrelated bone marrow [uBM], cord blood [CB], or haploidentical peripheral blood [haplo-PB]), HLA compatibility, conditioning intensity (myeloablative conditioning [MAC] or reduced intensity conditioning [RIC]), and the number of allo-HCT were used as predictive variables [2-5].

The predictive objectives were prediction of 1-year overall survival (OS), 1-year progression-free survival (PFS), and 1-year cumulative incidences of relapse/progression and non-relapse mortality (NRM). An OS event was defined as death from any cause. PFS was defined as survival without relapse/progression. NRM was defined as death without relapse/progression. We considered cases that never achieved complete remission after allo-HCT as relapse/progression on day 1 after allo-HCT. Relapse/progression and NRM were treated as competing events with

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

one another. The probabilities of OS and PFS were calculated using the Kaplan–Meier method. Cumulative incidences of relapse/progression and NRM were calculated using Gray's method.

Statistical Analysis

The following is a summary of the analysis outline: (i) preprocessing data quality assurance, (ii) development of a web application tool, and (iii) assessment of the predictive performance for OS, PFS, relapse/progression, and NRM in accordance with the Type 2b (nonrandom split-sample development and validation) of prediction model studies covered by the Transparent Reporting of a prediction model for Individual Prognosis Or Diagnosis (TRIPOD) statement [16]. We had no missing data. The number of trees to develop the RSF model was set as 500. A log-rank splitting rule was used to grow survival trees of a forest.

We used R package randomForestSRC [17] for analysis of the RSF model, R package timeROC [18] for calculating AUCs, and R package shiny/R for development of the web application. All analyses were performed using R version 3.5.1. We adhere to the TRIPOD statement [16].

Performance Assessment of Predictive Model

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

All predictive performances were calculated using the time-dependent area under the receiver-operating characteristic curve (AUC) [18]. First, the entire cohort was split into a training cohort (70%) and test cohort (30%) time-sequentially. Second, the optimal number of variables randomly selected as candidates for splitting a node, which is a hyperparameter in the RSF model, for each outcome was tuned using 5-fold cross-validation method in the training cohort [19]. It was defined as the one with the highest AUC value. Third, we assessed the predictive performances of the RSF, Cox PH model, DRI-R, and HCT-CI for 1-year OS and PFS, and of RSF, DRI-R, and HCT-CI for 1-year relapse/progression and NRM in test cohort. We plotted time-sequential AUCs per month from 3 to 12 months after allo-HCT in the test cohort to assess the continuous predictive performances in each prognosis prediction curve. The schema of the assessment method for these predictive models is shown in Supplementary Fig. 1.

Development of a Web Application Tool

We developed a web application tool capable of plotting the personalized prognosis prediction curves of 1-year overall survival (OS), 1-year progression-free survival (PFS), and 1-year cumulative incidences of relapse/progression and non-relapse mortality (NRM) adjusted by eight patient-specific factors of a new

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

allo-HCT candidate, using RSF predictive model trained by the entire cohort.

RESULTS

Patients and Transplantations

Three hundred and eighty-four patients underwent allo-HCT. One patient who underwent transplant from unrelated peripheral blood stem cells was excluded due to insufficient sample for analysis. Since DRI-R cannot be scored when there is benign or rare disease, 21 patients were also excluded (aplastic anemia, 11; Idiopathic cytopenia of undetermined significance, 1; chronic active Epstein–Barr virus disease, 4; plasma cell leukemia, 1; chronic myelomonocytic leukemia, 2; myelodysplastic/myeloproliferative neoplasms, unclassifiable, 1). In total, 363 patients were analyzed in this study. The clinical characteristics of the entire cohort are listed in Table 1. The median follow-up period for the survivor in the entire cohort was 51.9 months (range; 2.4-127.9 months). The 1-year OS and PFS after allo-HCT of the entire cohort were 62.7% (95% confidence interval [CI]; 57.5-67.5) and 49.6% (95% CI; 44.3-54.6). The 1-year cumulative incidences of relapse/progression and NRM in the entire cohort were 40.5% (95% CI; 35.4-45.5) and 9.9% (95% CI; 7.1-13.3), respectively. The clinical characteristics and outcomes in the training cohort (254/363) and test cohort (109/363) are listed in Supplementary Table 1.

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Interactive Web Application for Plotting Personalized Prognosis Prediction Curves

The interactive web application that shows the personalized 1-year prognosis prediction curves after allo-HCT is shown in Fig. 1. Users choose an outcome they intend to predict among OS, PFS, relapse/progression, or NRM in the select box at left-top (a). Next, after users set eight patient-specific factors in the left side-panel (b) and click the “prediction” button (c), the personalized 1-year prognosis prediction curve is displayed in the main panel (d). For example, Fig. 1 shows the personalized 1-year OS prediction curve after allo-HCT for a patient whose eight pre-transplant factors are as follows: age, 50; DRI-R, high; PS, 1; HCT-CI, 1; conditioning intensity, MAC; HLA compatibility, matched; donor source, rPB; the number of allo-HCT, 1. This patient’s 1-year OS rate after allo-HCT is predicted at 59.1%. Users can also see the personalized 1-year PFS curve and cumulative incidence prediction curves of relapse/progression and NRM after allo-HCT. We provided an online interface to use this web application (https://predicted-os-after-transplantation.shinyapps.io/RSF_model/) and its source code (https://github.com/HiroshiOkamura-OCU/Prognosis_prediction_curves_in_allo-HCT).

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

The predictive performances for each 1-year prognosis model in the RSF trained by the entire cohort is shown in Supplementary Table 2. They were calculated using the 5-fold cross-validation method to prevent overfitting.

Assessment of Predictive Performance

The AUCs calculated by the 5-fold cross-validation in the training cohort for 1-year OS, PFS, relapse/progression, and NRM per hyperparameter in the RSF model are shown in Table 2. The optimal number of variable candidates for splitting a node in the RSF model for 1-year OS, PFS, relapse/progression, and NRM were 7, 2, 2, and 5, respectively.

The AUCs for 1-year OS, PFS, relapse/progression, and NRM in the test cohort are shown in Table 3. Time-sequential AUCs plotted per month from 3 to 12 months after allo-HCT for OS, PFS, relapse/progression, and NRM in the test cohort are shown in Fig. 2. The predictive performances of RSF for OS and PFS were comparable to those of the Cox PH model at any time-point. The predictive performances of RSF for every outcome were superior to those of DRI-R or HCT-CI alone consistently at any given time-point.

DISCUSSION

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

We succeeded in developing an interactive web application tool with GUI capable of accurately, objectively, and rapidly generating the four types of personalized 1-year prognosis prediction curves after allo-HCT in just a few steps using machine learning. To the best of our knowledge, there have been no previous reports regarding the development of a similar interactive tool that can provide various types of patient-specific prognosis prediction curves in allo-HCT.

Some predictive indices of prognosis after allo-HCT such as DRI-R, HCT-CI, EBMT risk score, and pre-transplantation assessment of mortality score have been reported [2, 6-8, 20]. Although these indices can stratify prognosis after allo-HCT, transplant clinicians need to empirically consider various additional factors, such as age, PS, donor source, or conditioning regimen, to estimate patient-specific prognosis in clinical practice. Our web application tool can calculate and plot personalized prognosis prediction curves more objectively than a clinician's empirical estimate and more accurately than any single prognosis index.

Shouval et al. developed a website to predict prognosis after allo-HCT using the alternating decision tree (ADTree), which is a machine learning model [21]. Users can obtain personalized prognostic predictive information such as day 100 mortality by providing certain individual pre-transplant factors for a new transplant candidate. However, the ADTree model cannot treat time-to-event data, whereas the RSF is a

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

machine learning algorithm for survival analysis that can take the right-censored data or competing risks into consideration [12, 22]. Hence, our web application can not only provide prognosis prediction value for a fixed time point but also plot continuous survival and cumulative incidence prediction curves until 1 year after allo-HCT (Fig. 1). Moreover, users can simulate various transplant settings for a particular patient intuitively with just a few clicks using mobile devices such as tablets and smartphones as well as computers because our web application was developed using shiny/R library to ensure interactivity. Consequently, our web application can help transplant clinicians inform a new allo-HCT candidate of their personalized prognosis estimate in clinical practice, and it can provide both the clinician and candidate with valuable information to help make clinical decisions.

Shouval et al. also reported on the prognosis prediction performance after allo-HCT by ADTree using registry data in European Group for Blood and Marrow Transplantation (EBMT) and Italian Group for Bone Marrow Transplantation (GITMO) and showed that the time-dependent AUCs for 2-year OS were 0.66 and 0.65, respectively [21, 23]. Their prediction model developed using large amounts of patient data is quite relevant, because it can be used widely as a generalized prediction model in allo-HCT fields. However, various biases related to transplant institutes or countries are inherent in allo-HCT. For example, there are differences in the

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

indications for allo-HCT, transplant experiences, patient characteristics, and strategy for donor source selection, conditioning regimen, or transplant-related complications among the transplant institutes or countries. Therefore, some reports have shown that there is a significant difference in prognosis after allo-HCT among transplant centers or countries [24, 25]. The prediction model developed using data from a single institution is protected from bias based on the differences among the transplant institutes as long as it is used within that institute, because the patient characteristics or transplant procedures similar to those of the training cohort by which the prediction model was developed are likely to be inherited by future patients at that institute. Furthermore, the noise of data collected from a single institute may be less than that of multicenter data, because the methods of clinical assessment or data collection could be consistent in a single institute. Hence, an in-house developed prognosis prediction model may have more advantages for prognosis predictive performance after allo-HCT in a particular institute than the prediction model based on data sourced from multiple institutions. Supporting this idea, Fuse et al. also showed that the predictive performance of relapse in a single institute where the predictive model was developed using ADTree was better than that in another institute using the identical predictive model (AUC; 0.75 vs 0.67) [26]. We have published the source code of our web application on the Internet. Each transplant center can also develop

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

a web application using an in-house prognosis prediction model with its past patient data.

To assess the predictive performance, we developed predictive models with the training cohort, which is the former group generated by splitting the entire cohort into two groups time-sequentially, and their predictive performances were assessed in the test cohort, the latter group. This assessment method accords with the Type 2b of prediction model studies covered by the TRIPOD statement [16]. As a result, the predictive performances for 1-year OS, EFS, relapse/progression in the test cohort were worse than those in the training cohort (Table 2, 3). These results could be attributed to the difference in the year of allo-HCT between the training and test cohorts. The predictive performances of the RSF model trained by the entire cohort, which was used in the web application, is shown in supplementary table 2. However, the predictive performances for the new allo-HCT candidates in this web application may not be as high as those shown in supplementary table 2 because the allo-HCT procedures and patient characteristics may have changed over time.

This study has several limitations. First, the predictive performance of our web application for new allo-HCT candidates in other institutes is unknown since our model has not obtained external validation. However, as mentioned above, we suppose it would be suitable for a transplant center, which contains many allo-HCT

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

data, to develop an in-house model for more accurate prognosis prediction. Therefore, we provide the source codes, which will help each transplant institute to develop an in-house web application for personalized prognosis prediction after allo-HCT and assess its predictive performance. Second, as a retrospective analysis, this web application may be susceptible to selection bias, which may limit the interpretation of the prediction results. The predictive accuracy is unreliable in cases where user-provided transplant settings were not included in the training cohort for model development. Third, a different combination of pre-transplant factors might be considered to develop a better predictive model. We restricted the number of pre-transplant predictors to eight to ensure the usability of this web application tool. These refined eight pre-transplant parameters could be easily and non-invasively obtained at a low cost. However, there may be other combinations, including molecular minimal residual disease, that are better for prognosis prediction. Therefore, further studies are required.

In conclusion, we developed an interactive web application tool capable of plotting four kinds of personalized prognosis prediction curves after allo-HCT objectively through the integration of eight pre-transplant factors using RSF, and confirmed their promising predictive performances. The predictive information obtained by this web application could allow transplant clinicians to inform a new allo-HCT candidate of

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

their objective personalized prognosis prediction in clinical practice and could support the clinician's and the candidate's decision-making.

Acknowledgments

We wish to thank the Hematopoietic Cell Transplant Coordinator of Osaka City University Hospital, Ms. Umemoto, and Ms. Tanaka for assistance with data entry. This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant (number 17K09017).

Conflict of interest:

All authors have no conflicts of interest to disclose.

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

References

1. Copelan EA. Hematopoietic stem-cell transplantation. N Engl J Med 2006; **354**: 1813-1826.
2. Carreras E., Rambaldi A. Evaluation and counseling of candidates. In: Carreras E, Dufour C, Mohty M, Kröger N (eds). The EBMT Handbook, Hematopoietic Stem Cell Transplantation and Cellular Therapies. Springer Nature Switzerland AG: Cham, Switzerland, 2019, pp 1443–1454
3. Scott BL, Pasquini MC, Logan BR, Wu J, Devine SM, Porter DL et al. Myeloablative versus reduced-intensity hematopoietic cell transplantation for acute myeloid leukemia and myelodysplastic syndromes. J Clin Oncol 2017; **35**: 1154-1161.
4. Milano F, Gooley T, Wood B, Woolfrey A, Flowers ME, Doney K et al. Cord-blood transplantation in patients with minimal residual disease. N Engl J Med 2016; **375**: 944-953.
5. Fleischhauer K, Hsu KC, Shaw BE. Prevention of relapse after allogeneic hematopoietic cell transplantation by donor and cell source selection. Bone Marrow Transplant 2018; **53**: 1498-1507.
6. Armand P, Kim HT, Logan BR, Wang Z, Alyea EP, Kalaycio ME et al. Validation and refinement of the Disease Risk Index for allogeneic stem cell transplantation.

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Blood 2014; **123**: 3664-3671.

7. Sorror ML, Maris MB, Storb R, Baron F, Sandmaier BM, Maloney DG et al. Hematopoietic cell transplantation (HCT)-specific comorbidity index: a new tool for risk assessment before allogeneic HCT. Blood 2005; **106**: 2912-2919.
8. Gratwohl A. The EBMT risk score. Bone Marrow Transplant 2012; **47**: 749-756.
9. Shouval R, Ruggeri A, Labopin M, Mohty M, Sanz G, Michel G et al. An integrative scoring system for survival prediction following umbilical cord blood transplantation in acute leukemia. Clin Cancer Res 2017; **23**, 6478-6486.
10. Zdilar L, Vock DM, Marai GE, Fuller CD, Mohamed ASR, Elhalawani H et al. Evaluating the effect of right-censored end point transformation for radiomic feature selection of data from patients with oropharyngeal cancer. JCO Clin Cancer Inform 2018; **2**, 1-19.
11. Ingrisch M, Schoppe F, Paprottka K, Fabritius M, Strobl FF, De Toni EN et al. Prediction of (90)Y radioembolization outcome from pretherapeutic factors with random survival forests. J Nucl Med 2018; **59**: 769-773.
12. Mogensen UB, Ishwaran H, Gerds TA. Evaluating random forests for survival analysis using prediction error curves. J Stat Softw 2012; **50**: 1-23.
13. Wang H, Li G. A selective review on random survival forests for high dimensional data. Quant Biosci 2017; **36**: 85-96.

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

14. Giralt S, Ballen K, Rizzo D, Bacigalupo A, Horowitz M, Pasquini M et al. Reduced-intensity conditioning regimen workshop: defining the dose spectrum. Report of a workshop convened by the center for international blood and marrow transplant research. Biol Blood Marrow Transplant 2009; **15**: 367-369.
15. Nakamae M, Yamashita M, Koh H, Nishimoto M, Hayashi Y, Nakane T et al. Lung function score including a parameter of small airway disease as a highly predictive indicator of survival after allogeneic hematopoietic cell transplantation. Transpl Int 2016; **29**: 707-714.
16. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015; **162**: W1-73.
17. Ishwaran H, Gerds TA, Kogalur UB, Moore RD, Gange SJ, Lau BM. Random survival forests for competing risks. Biostatistics 2014; **15**: 757-773.
18. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Stat Med 2013; **32**: 5381-5397.
19. Hastie T, Tibshirani R, Friedman J. 7.10 Cross-Validation. In: Hastie T, Tibshirani R, Friedman J (eds). The Elements of Statistical Learning, Data Mining, Inference,

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

and Prediction, 2nd edn. Springer: New York, USA, 2009, pp 241-248

20. Parimon T, Au DH, Martin PJ, Chien JW. A risk score for mortality after allogeneic hematopoietic cell transplantation. Ann Intern Med 2006; **144**: 407-414.
21. Shouval R, Labopin M, Bondi O, Mishan-Shamay H, Shimoni A, Ciceri F et al. Prediction of allogeneic hematopoietic stem-cell transplantation mortality 100 days after transplantation using a machine learning algorithm: a European Group for Blood and Marrow Transplantation Acute Leukemia Working Party Retrospective Data Mining Study. J Clin Oncol 2015; **33**: 3144-3151.
22. Arai Y, Kondo T, Fuse K, Shibasaki Y, Masuko M, Sugita J et al. Using a machine learning algorithm to predict acute graft-versus-host disease following allogeneic transplantation. Blood Adv. 2019; **3**: 3626-3634.
23. Shouval R, Bonifazi F, Fein J, Boschini C, Oldani E, Labopin M et al. Validation of the acute leukemia-EBMT score for prediction of mortality following allogeneic stem cell transplantation in a multi-center GITMO cohort. Am J Hematol 2017; **92**: 429-434.
24. Marmor S, Begun JW, Abraham J, Virnig BA. The impact of center accreditation on hematopoietic cell transplantation (HCT). Bone Marrow Transplant 2015; **50**: 87-94.
25. Gratwohl A, Sureda A, Baldomero H, Gratwohl M, Dreger P, Kröger N et al.

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Economics and outcome after hematopoietic stem cell transplantation: A retrospective cohort study. EBioMedicine 2015; **2**: 2101-2109.

26. Fuse K, Uemura S, Tamura S, Suwabe T, Katagiri T, Tanaka T et al. Patient-based prediction algorithm of relapse after allo-HSCT for acute leukemia and its usefulness in the decision-making process using a machine learning approach. Cancer Med 2019; **8**: 5058-5067.

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Figure Legends

Fig. 1 Interactive web application for plotting the personalized prognosis prediction curves after allo-HCT

The view and contents of the web application and an example of the personalized 1-year OS prediction curve which was displayed interactively for a patient (age, 50; DRI-R, high; PS, 1; HCT-CI, 1; conditioning intensity, MAC; HLA compatibility, matched; donor source, rPB; the number of allo-HCT, 1.) are shown.

allo-HCT, allogeneic hematopoietic cell transplantation; DRI-R, refined Disease Risk Index; HCT-CI, Hematopoietic Cell Transplantation Comorbidity Index; HLA, human leukocyte antigen; MAC, myeloablative conditioning; NRM, non-relapse mortality; OS, overall survival; PFS, progression-free survival; PS, performance status; RIC, reduced intensity conditioning; rPB, related peripheral blood.

Fig.2 Time-sequential predictive performances

Time-sequential AUCs of RSF, Cox PH model, DRI-R, and HCT-CI for OS (a) and PFS (b) and those of RSF, DRI-R, and HCT-CI for relapse/progression (c) and NRM (d) are plotted per month from 3 to 12 months after allo-HCT.

allo-HCT, allogeneic hematopoietic cell transplantation; AUCs; area under the

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

receiver-operating characteristic curves; Cox PH, Cox proportional hazard; DRI-R, refined Disease Risk Index; HCT-CI, Hematopoietic Cell Transplantation Comorbidity Index; NRM, non-relapse mortality; OS, overall survival; PFS, progression-free survival; RSF, random survival forest.

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Table 1. Patient Characteristics

Characteristic	All patients (n = 363)
Median age [range]	46 [17-69]
Diagnosis (%)	
AML	170 (47)
ALL	63 (17)
MDS	46 (13)
CML	13 (4)
ML	43 (12)
Others	28 (8)
DRI-R (%)	
low	19 (5)
int	176 (49)
high	130 (36)
very high	38 (11)
HCT-CI (%)	
0	147 (41)
1	67 (19)
2	62 (17)
over 3	87 (24)
Conditioning (%)	
MAC	219 (60)
RIC	144 (40)
HLA disparity (%)	
matched	139 (38)
mismatched	224 (62)
Source of stem cell (%)	
rBM	8 (2)
rPB	54 (15)
uBM	107 (30)
CB	80 (22)
Haplo PB	114 (31)
PS (%)	
0	185 (51)
1	159 (44)
over 2	19 (5)

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .

Number of allo-HCT (%)	
1	289 (80)
2	62 (17)
3	12 (3)
Year of allo-HCT (%)	
2008-2014	251 (69)
2015-2017	112 (31)
Prognosis [95% CI]	
1-year OS (%)	62.7 [57.5-67.5]
1-year PFS (%)	49.6 [44.3-54.6]
1-year relapse/progression (%)	40.5 [35.4-45.5]
1-year NRM (%)	9.9 [7.1-13.3]
Median follow-up months for the survivor [range]	51.9 [2.4-127.9]

Abbreviations: AML, acute myeloid leukemia; ALL, acute lymphoblastic leukemia; allo-HCT, allogeneic hematopoietic cell transplantation; BM, bone marrow; CI, confidence interval; CML, chronic myeloid leukemia; DRI-R, refined Disease Risk Index; Haplo, HLA-haploidentical; HCT-CI, Hematopoietic Cell Transplantation Comorbidity Index; HLA, human leukocyte antigen; MAC, myeloablative conditioning; MDS, myelodysplastic syndrome; ML, malignant lymphoma; NRM, non-relapse mortality; OS, overall survival; PB, peripheral blood; PFS, progression free survival; PS, performance status; RIC, reduced intensity conditioning.

Table 2. Predictive performance in the training cohort per hyperparameter in the RSF model

The number of variable candidates (hyperparameter)	AUC			
	1-year OS	1-year PFS	1-year relapse/progression	1-year NRM
1	0.7710	0.7482	0.7394	0.6124
2	0.7774	**0.7606**	**0.7542**	0.6502
3	0.7806	0.7600	0.7504	0.6684
4	0.7828	0.7582	0.7452	0.6712
5	0.7866	0.7548	0.7422	**0.6838**
6	0.7856	0.7548	0.739	0.6776
7	**0.7878**	0.7544	0.7348	0.6740
8	0.7864	0.7540	0.7330	0.6794

Abbreviations: AUC, area under the receiver-operating characteristic curve; NRM, non-relapse mortality; OS, overall survival; PFS, progression free survival; RSF, random survival forest.

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

Table 3. Predictive performance in the test cohort

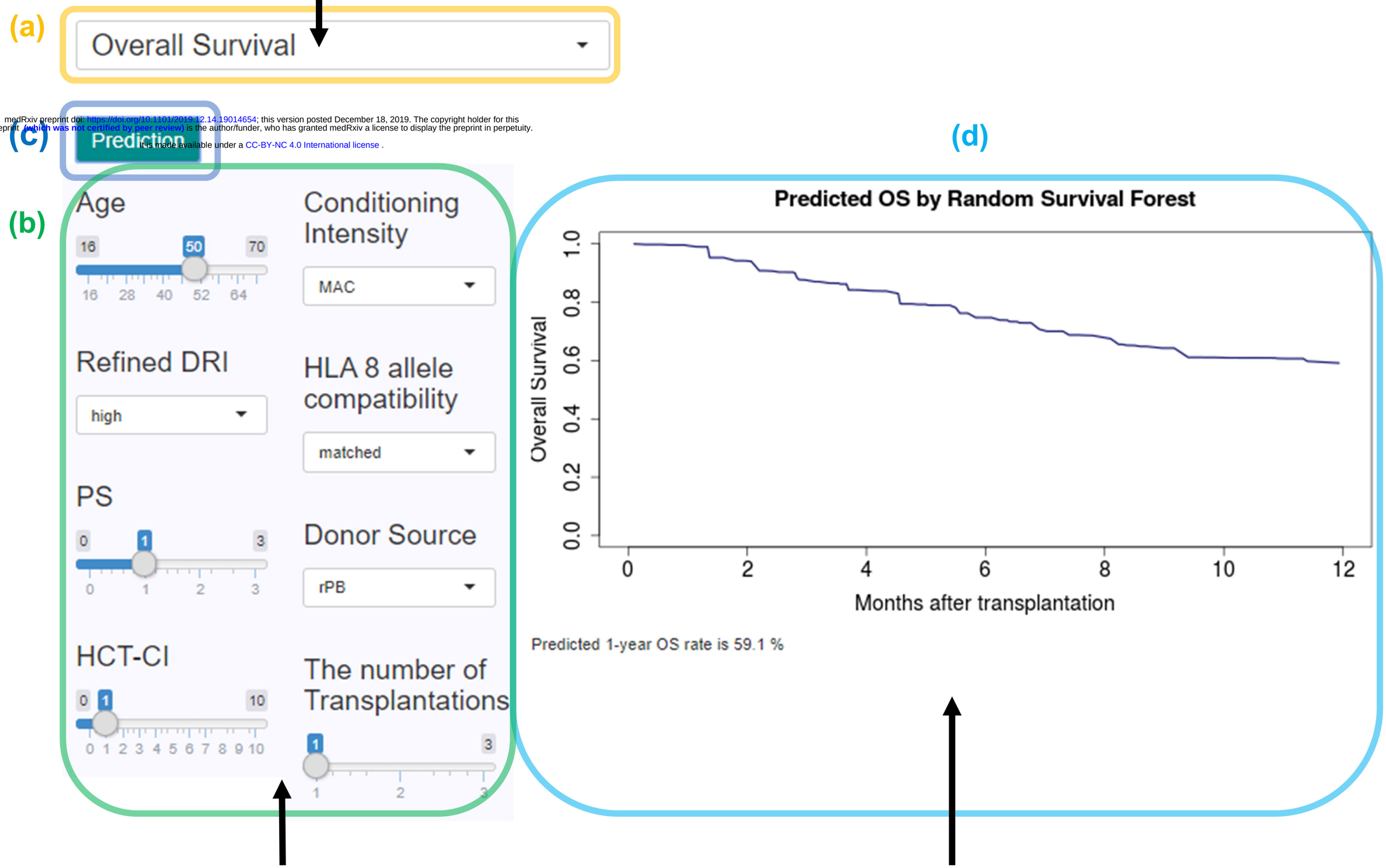
Predictive model	AUC			
	1-year OS [95% CI]	1-year PFS [95% CI]	1-year relapse/progression [95% CI]	1-year NRM [95% CI]
RSF	0.70 [0.60-0.80]	0.72 [0.62-0.82]	0.73 [0.63-0.84]	0.77 [0.64-0.90]
Cox PH	0.71 [0.60-0.81]	0.70 [0.59-0.80]	-	-
DRI-R	0.65 [0.55-0.75]	0.67 [0.57-0.77]	0.72 [0.62-0.81]	0.39 [0.24-0.53]
HCT-CI	0.55 [0.44-0.66]	0.54 [0.44-0.65]	0.51 [0.42-0.64]	0.59 [0.43-0.75]

Abbreviations: AUC, area under the receiver-operating characteristic curve; Cox PH, Cox proportional hazard; DRI-R, refined disease risk index; HCT-CI, hematopoietic cell transplantation comorbidity index; NRM, non-relapse mortality; OS, overall survival; PFS, progression free survival; RSF, random survival forest.

medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

Fig. 1

Users choose an outcome they intend to predict either OS, PFS, relapse/progression, or NRM

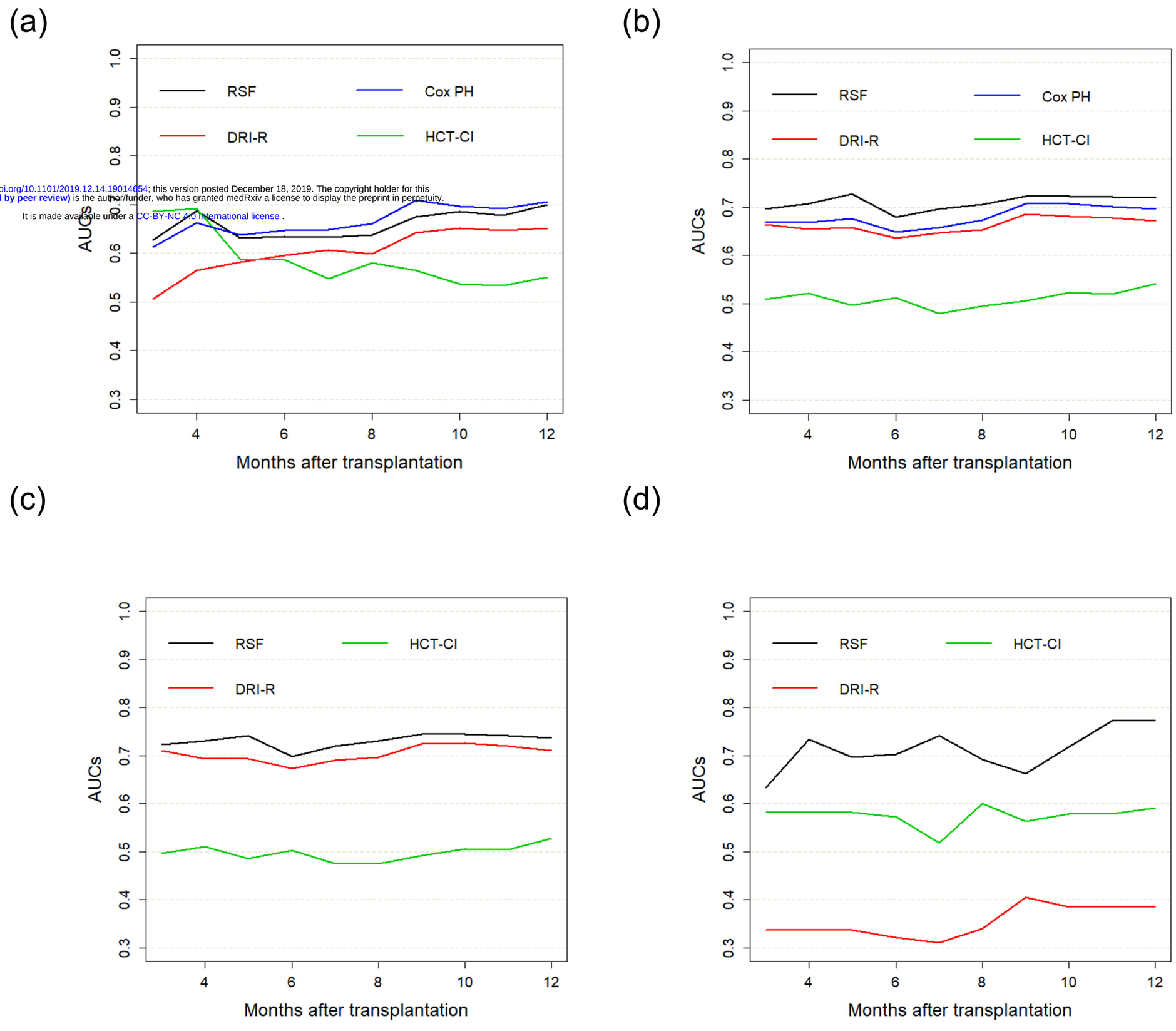


medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Users set eight patient-specific pre-transplant factors

The personalized 1-year prognosis prediction curve is displayed

Fig. 2



medRxiv preprint doi: https://doi.org/10.1101/2019.12.14.19014654; this version posted December 18, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY-NC 4.0 International license .