A genetic score for the prediction of beta-thalassemia severity

Fabrice Danjou,[1] Marcella Francavilla,[1] Franco Anni,[1] Stefania Satta,[1] Franca-Rosa Demartis,[1] Lucia Perseu,[2] Matteo Manca,[2] Maria Carla Sollaino,[2] Laura Manunza,[1] Elisabetta Mereu,[1] Giuseppe Marceddu,[1] Serge Pissard,[3] Philippe Joly,[4] Isabelle Thuret,[5] Raffaella Origa,[1] Joseph Borg,[6] Gian Luca Forni,[7] Antonio Piga,[8] Maria Eliana Lai,[2] Catherine Badens,[9] Paolo Moi,[1,2] Renzo Galanello[1,2]

[1]Dipartimento di Sanità Pubblica, Medicina Clinica e Molecolare, Università degli Studi di Cagliari, Italy; [2]Ospedale Regionale per le Microcitemie, ASL8, Cagliari, Italy; [3]Laboratoire de Génétique, Hôpital Henri Mondor, Créteil, France; [4]Laboratoire de Biochimie et Biologie Moléculaire, Hôpital Edouard Herriot, Lyon, France; [5]French reference center for Thalassemia and Service d'Onco-Hématologie Pédiatrique, Assistance Publique Hôpitaux de Marseille, France; [6]Department of Applied Biomedical Science, Faculty of Health Sciences and Laboratory of Molecular Genetics, Department of Physiology and Biochemistry, University of Malta, Msida, Malta; [7]Ematologia - Centro della Microcitemia ed Anemie Congenite, Ospedale Galliera, Genova, Italy; [8]Università di Torino, Italy; and [9]French reference center for Thalassemia and Department of Genetics, Assistance Publique Hopitaux de Marseille and Aix-Marseille University, France

ABSTRACT

Clinical and hematologic characteristics of beta(β)-thalassemia are determined by several factors resulting in a wide spectrum of severity. Phenotype modulators are: *HBB* mutations, *HBA* defects and fetal hemoglobin production modulators (HBG2:g.-158C>T polymorphism, *HBS1L-MYB* intergenic region and the *BCL11A*). We characterized 54 genetic variants at these five loci robustly associated with the amelioration of beta-thalassemia phenotype, to build a predictive score of severity using a representative cohort of 890 β-thalassemic patients. Using Cox proportional hazard analysis on a training set, we assessed the effect of these loci on the age at which patient started regular transfusions, built a Thalassemia Severity Score, and validated it on a testing set. Discriminatory power of the model was high (C-index=0.705; R^2=0.343) and the validation conducted on the testing set confirmed its predictive accuracy with transfusion-free survival probability ($P<0.001$) and with transfusion dependency status (Area Under the Receiver Operating Characteristic Curve=0.774; $P<0.001$). Finally, an automatized on-line calculation of the score was made available at *http://tss.unica.it*. Besides the accurate assessment of genetic predictors effect, the present results could be helpful in the management of patients, both as a predictive score for screening and a standardized scale of severity to overcome the major-intermedia dichotomy and support clinical decisions.

Introduction

Clinical and hematologic characteristics of beta(β)-thalassemia are determined by several factors resulting in a wide spectrum of severity from no need to dependence on regular blood transfusions; they mainly include the type of disease causing mutation and the capacity production of alpha(α)- and gamma(γ)-globin chains. Mutations in the *HBB* gene determine lack or low level of β-globin chains synthesis in the erythropoietic cells, causing an imbalance of the α- to β-globin chains ratio. This imbalance is especially evident in patients homozygous for *HBB* mutations in whom the accumulation of unbound α-globin chains, forming highly toxic aggregates that precipitate in the erythroid precursor of the bone marrow, leads to ineffective erythropoiesis resulting in severe anemia due to low red blood cell survival from ineffective erythropoiesis and hemolysis. The clinical presentation is widely variable because the amount of unbound α-globin chains can be modified by both the capacity to produce α-globin chains (*HBA* genes variants) and the capacity to produce γ-globin chains (*HBG2* gene modulators) that can bind available α-globin chains to form effective fetal hemoglobin (HBF).[1,2] The severity of mutations in the *HBB* gene and defects of the HBA genes were the first determinants of the phenotype variability of β-thalassemia to be discovered. The third determinant to be identified was the XmnI polymor-phism of the *HBG2* promoter (HBG2:g.-158C>T) which is widely included in the diagnostic workup for thalassemia patients.[3,4] More recently genome-wide association studies contributed to the definition of very important trans-acting modifiers of the production of fetal hemoglobin such as the *BCL11A* gene and the *HBS1L-MYB* intergenic region.[5–10]

Previous studies investigated the contribution of known genetic modifiers to the major-intermedia phenotypic classification demonstrating that a great proportion of this status is effectively genetically determined.[11,12] This condition makes β-thalassemia one of the few diseases with complex phenotype that could be accurately predicted on a genetic basis, opening the path for the clinical application of genetic prediction to many other diseases. Recently, we demonstrated the feasibility of a more accurate prediction of the hematologic severity of β-thalassemia, using known genetic modifiers to predict the age of the start of regular blood transfusions in a cohort of homozygous β-39 individuals (HBB:c.118C>T).[13] This approach has two main benefits: 1) the severity is measured on a simple and reproducible scale that precisely assesses the overall spectrum of hematologic severity, therefore overcoming the major-intermedia dichotomy; and 2) the factors considered are exclusively genetic allowing prediction from a unique measurement already available during pregnancy, and have a wide range of clinical applications, such as screening, genetic scoring and assistance in therapeutic decisions.[13]

The aim of the present study was to build and validate a predictive score of severity based on known genetic markers, using a large representative cohort of β-thalassemic patients. In six centers from three countries of the Mediterranean basin, we recruited 890 patients for whom the date at which regular transfusion started was accurately registered. The whole cohort was genotyped for six markers at five loci that were already robustly associated with both HBF levels and/or the amelioration of β-thalassemia phenotype.[6,14–18] Using Cox proportional hazard analysis for the age of regular transfusion start, we built a Thalassemia Severity Score (TSS) from genetic markers, tested its predictive ability, and, finally, made it available online at: *http://tss.unica.it*

Methods

Study sample and phenotye assessment

For this multi-center study, an international cohort of 890 β-thalassemia patients was recruited from the pediatric and adult departments of the Microcythemia Hospital of Cagliari, the San Luigi Gonzaga Hospital of Turin, the Hematology Center for Microcythemia and Congenital Abnormalities, at Galliera Hospital, Genoa, Italy, the French Reference Center for Thalassemia from Marseille, and the Thalassaemia and Molecular Genetics Clinic of the Mater Dei Hospital from Malta.

Overall data were pooled into a single dataset composed of 49.9% females and 50.1% males, representing 45 β-thalassemic mutations (65 β-thalassemic genotypes) from 88.24% transfusion-dependent thalassemia (TDT) and 11.76% non-transfusion-dependent thalassemia (NTDT) patients. The time elapsed between birth and the start of regular transfusions was continuously distributed and reflected the variability of the hematologic severity among uncensored thalassemia patients (TDT, median transfusion-free survival time of 11 months), whereas the time between birth and last follow up was taken into account for censored patients (NTDT, median survival time 44 years). Data relative to the transfusion program of the patients were collected through the WebTHAL computerized clinical records database (*http://www.thalassemia.it*) in use for the daily management of patients in Italian centers, and through direct transmission of data regarding the other two centers. Criteria to start transfusion were: hemoglobin level lower than 7 g/dL for more than two weeks in absence of infections, moderate to severe spleen enlargement, initial skeletal changes and/or poor growth. Patients were considered as regularly transfused when undergoing more than eight blood transfusions a year.

This retrospective study was registered with local ethical committees and conducted in accordance with the Declaration of Helsinki; all patients gave informed consent for DNA analysis and research study.

Marker selection and genotyping

We selected a set of genetic markers robustly associated with the disease severity of β-thalassemia to use as genetic predictors: the type of *HBB* gene mutation, the HBG2:g.-158C>T polymorphism, the type and number of *HBA* genes defects, two SNPs from

Table 1. Genetic modulators of beta-thalassemia phenotype: loci characteristics.

Loci	Position (GRCh37.p10)		Categories	Nb of cases	Frequencies
HBB gene defects	chr11:5246696-5248301	mild/mild	β^+_mild/β^+_mild	13	1.60%
			β^+_silent/β^0_severe	4	0.49%
			β^+_mild/β^0_mild	2	0.25%
			β^0_mild/β^0_mild	1	0.12%
		mild/severe	β^+_mild/β^0_severe	62	7.61%
			β^+_mild/β^+_severe	20	2.45%
			β^0_mild/β^0_severe	5	0.61%
		severe/severe	β^0_ severe /β^0_ severe	603	73.99%
			β^+_ severe /β^0_ severe	71	8.71%
			β^+_ severe /β^+_ severe	34	4.17%
HBA genes defects	chr16:226679-227520		No defect	560	65.19%
			1 defect	230	26.78%
			(from 3.7kb-I, 4.2kb and HphI deletions or NcoI polymorphism)		
			2 defects	69	8.03%
			(from MED-I, 20.5kb, 3.7kb-I/II/III, 4.2kb and HphI deletions or NcoI polymorphism)		
HBG2	rs7482144 (HBG2:g.-158C>T)	chr11:5276169	C/C (-/-)	812	94.53%
			C/T (-/+)	41	4.77%
			T/T (+/+)	6	0.70%
HBS1L-MYB intergenic region	rs9399137	chr6:135419018	T/T	556	64.73%
			C/T	292	33.99%
			C/C	11	1.28%
BCL11A	rs1427407	chr2:60718043	G/G	535	62.28%
			G/T	284	33.06%
			T/T	40	4.66%
	rs10189857	chr2:60713235	A/A	418	48.66%
			A/G	310	36.09%
			G/G	131	15.25%

Table 2. Distribution of genetic modulators of beta-thalassemia severity in the studied cohort.

Loci			Median time to first transfusion (in months) of TDT patients [5th-95th percentile]	NTDT proportion [corresponding folds increased proportion]*
HBB gene defects	mild/mild		54 [27 - 698]	60.0% [5.10]
	mild/severe		19 [4 - 131]	14.9% [1.27]
	severe/severe		10 [2 - 75]	10.1% [0.85]
HBA genes defects	no defect		10 [2 - 72]	7.0% [0.59]
	1 defect		12 [3 - 82]	18.7% [1.57]
	2 defects		26 [6 - 118]	27.5% [2.34]
HBG2	rs7482144	C/C [–/–]	11 [3 - 77]	10.5% [0.90]
	(HBG2:g.-158C>T)	C/T [–/+]	35 [3 - 131]	34.1% [2.90]
		T/T [+/+]	56 [8 - 265]	16.7% [1.42]
HBS1L-MYB intergenic region	rs9399137	T/T	10 [2 - 77]	8.8% [0.75]
		C/T	13 [3 - 93]	15.7% [1.34]
		C/C	28 [3 - 32]	54.5% [4.64]
BCL11A	rs1427407	G/G	9 [2 - 81]	8.2% [0.70]
		G/T	14 [3 - 77]	16.8% [1.43]
		T/T	22 [6 - 70]	22.5% [1.91]
	rs10189857	A/A	10 [2 - 66]	14.8% [1.26]
		A/G	11 [3 - 81]	9.5% [0.81]
		G/G	13 [3 - 95]	11.5% [0.97]

*In this last column, for each class of predictor we report the percentage value of NTDT patients among the overall cohort, as well as the ratio of this percentage (for this specific class of predictor) over the percentage of NTDT patients in the entire cohort (11.76%). It promotes a better understanding of the effect and direction of each single predictor on NTDT proportions (a ratio greater than 1 is equivalent to a greater proportion of NTDT patients than average, and conversely for ratio smaller than 1).

the second intron of *BCL11A* gene (rs1427407 and rs10189857) and one from the *HBS1L-MYB* intergenic region (rs9399137). These markers have been independently associated to both HBF levels and the amelioration of β-thalassemia phenotype in different studies (for details see Table 1, *Online Supplementary Methods* and *Online Supplementary Table S1*).[3,4,6,13–18]

DNA was extracted with standard methods from venous peripheral blood. Mutation analysis of the β-globin gene was performed by direct DNA sequencing and discovered 45 different mutations (Table 1 and *Online Supplementary Table S1*). α-globin gene defects considered [3.7 kb rightward type I, II and III (NG_000006.1:g.34164_37967del3804, HbVar ID: 1078 and 1079, respectively), 4.2 kb leftward (HbVar ID: 1079), Mediterranean type I (NG_000006.1:g.24664_41064del16401), 20.5 kb (NG_000006.1:g.15164_37864del22701) large deletions, as well as HphI (HBA2:c.95+2_95+6delTGAGG) small deletion and NcoI (HBA2:c.2T>C) polymorphism, see Table 1] were determined using GAP-PCR or restriction enzyme digestion for large deletions and other defects, respectively.[19] The HBG2:g.-158C>T (rs7482144) polymorphism was determined as previously described, while other SNPs were directly genotyped using TaqMan SNP genotyping assay (Applied Biosystems, Warrington, UK).[20]

Statistical analysis

Accurate data quality controls were performed on the dataset, including consistency check of data and missing value analysis, after which 31 samples were removed, leaving 859 samples with complete data. All statistical procedures were performed using SPSS v.20.0 software package (SPSS Inc., Chicago, Illinois, USA).

The cohort was divided into a training set (71% selected randomly) and a testing set (the remaining 29%). The training set was used to build a Cox proportional hazard model for the age of regular transfusion start using genetic markers as predictors, and the testing set was used to test the predictive ability of the model using Cox proportional hazard analysis, Kaplan-Meier survival

analysis and ROC curves for TDT-NTDT status prediction. The choice of Cox proportional hazard model was driven by the fact that death for all causes is a negligible competing event with respect to start of transfusion in β-thalassemia patients.[21] For score elaboration, all variables were treated as categorical: sex was codified 0 when female and 1 when male, HBB mutations were codified upon their severity as reported in literature (*http://globin.bx.psu.edu/hbvar*) in mild/mild, mild/severe or severe/severe (from 1 to 3), *HBA* gene defects were classified as 0, 1, or 2 according to the number of deleted or mutated copies of the *HBA* gene, while for each SNP a variable was defined with the value of 0, 1, or 2 according to the number of copies of the less frequent allele (Table 1). For survival analysis, center of origin was used as a stratification variable, and patients were considered uncensored when blood transfusion occurred or censored when blood transfusion did not occur. We report Cox and Snell R^2 due to low proportion of early censoring as well as Harrell's concordance index (C-index) to assess how well the model performed. Finally, after validation on the testing set to check prediction accuracy, we derived a 0-10 scale of severity by standardizing the sum of Cox model hazard ratios as detailed in *Online Supplementary Methods*.

The Thalassemia Severity Score web-tool (*http://tss.unica.it*) is freely available to any user to calculate the TSS and plot the transfusion-free survival curve corresponding to the entered genotypes. The web tool calculates the TSS by transforming the sum of preloaded linear predictor scores as reported in the *Online Supplementary Methods*. The calculated TSS is then associated to one of the four statistically different survival curves observed in the studied cohort for a graphical presentation of the prediction.

Results

The distribution of selected genetic markers in the cohort is presented in Table 2. The genotypes from *HBB* gene mutations were 65 in total and the most frequent

Table 3. Results of the Cox proportional hazards model on training set.

Modificators of age at regular transfusion start			P		Hazard Ratio (folds of later transfusion start probability)	Cox and Snell R^2 (Harrell's C-index)
Sex		female	3.35E-02	reference	reference	*Sex:* 0.010 (0.532)
		male		3.35E-02	0.825 (1.212)	
HBB gene defects		severe/severe	6.51E-14	reference	reference	*HBB* gene defects: 0.105 (0.581)
		mild/severe		5.33E-09	0.373 (2.684)	
		mild/mild		1.05E-08	0.032 (30.899)	
HBA genes defects		No defect	3.16E-13	reference	reference	*HBA* genes defects: 0.056 (0.569)
		1 defect		6.86E-05	0.646 (1.549)	
		2 defects		9.79E-13	0.252 (3.971)	
HBG2	rs7482144	C/C (-/-)	1.46E-08	reference	reference	*HBG2:* 0.060 (0.532)
	(HBG2:g.-158C>T)	C/T (-/+)		5.41E-08	0.221 (4.522)	
		T/T (+/+)		8.02E-05	0.052 (19.360)	
HBS1L-MYB intergenic region	rs9399137	T/T	5.86E-09	reference	reference	*HBS1L-MYB* 0.039 (0.549)
		C/T		8.69E-08	0.598 (1.673)	
		C/C		6.16E-04	0.158 (6.320)	
BCL11A	rs1427407	G/G	2.26E-06	reference	reference	*BCL11A:* 0.042 (0.563)
		G/T		5.75E-04	0.702 (1.425)	
		T/T		1.09E-05	0.362 (2.759)	
	rs10189857	A/A	2.71E-02	reference	reference	
		A/G		9.00E-01	0.987 (1.013)	
		G/G		1.40E-02	0.689 (1.451)	

Cox and Snell R^2 and Harrell's C-index were respectively 0.343 and 0.703 considering all modificators together, while 0.133 and 0.607 considering only gamma-globin modificators (HBG2, HBS1L-MYB and BCL11A).

were HBB:c.118C>T/HBB:c.118C>T (65.5%), HBB:c.93-21G>A/HBB:c.118C>T (5.6%) and HBB:c.92+6T>C/HBB:c.118C>T (4.1%). The most common *HBA* genes defect was the 3.7 kb rightward deletion (22.0% with one deletion and 6.5% with two), among the four defects observed. Regarding other loci, the minor allele frequencies were 3% for rs7482144 (HBG2:g.−158C>T), 18% for rs9399137 (*HBS1L-MYB* intergenic region), 21% for rs1427407 and 40% for rs10189857 (*BCL11A*). *HBB* genes defects frequencies differed between centers. In particular, in Italian centers we mainly observed samples homozygous for HBB:c.118C>T and heterozygous HBB:c.93-21G>A/HBB:c.118C>T and HBB:c.92+6T>C/HBB:c.118C>T; in non-Italian centers the HBB:c.92+6T>C/HBB:c.92+6T>C genotype prevailed in Malta, and the HBB:c.93-21G>A/HBB:c.93-21G>A genotype in Marseille. In Sardinia, the mutation observed was almost exclusively HBB:c.118C>T/HBB:c.118C>T.

Results from the Cox proportional hazard model built on a random sample of 71% of the cohort (training set, n=613) are presented in Table 3. The type of *HBB* mutation had the strongest effect on the hematologic severity of β-thalassemia with a 30-fold increased risk of transfusion start for severe/severe phenotype *versus* mild/mild (mild/mild: HR=0.032; $P<0.001$), followed by HBG2:g.−158C>T polymorphism (homozygous T/T: HR=0.052; $P<0.001$), rs9399137 (*HBS1L-MYB*; homozygous C/C: HR=0.158; $P<0.001$), *HBA* gene defects (two defects: HR=0.252; $P<0.001$), rs1427407 (*BCL11A*; homozygous T/T: HR=0.362; $P<0.001$), rs10189857 (*BCL11A*; homozygous G/G: HR=0.689; $P=0.014$) and sex (males: HR=0.825; $P=0.033$). The discriminatory power of the model was high (C-index=0.703; $R^2=0.343$) and most of it was attributable to β-globin (C-index=0.581, $R^2=0.105$) and to modulators of the γ-globin chain production (HBG2:g.−158C>T, *BCL11A* and *HBS1L-MYB* loci: C-index=0.607, $R^2=0.133$), whereas the remaining was attributable to defects in the production of α-globin chains and to sex (C-index=0.569, $R^2=0.056$ and C-index=0.532, $R^2=0.010$, respectively).

Validation of the model was conducted on the testing set (29% of the cohort, 246 cases) to confirm its predictive accuracy. Cox proportional hazard model highly associated the linear predictor score with transfusion-free survival probability ($P<0.001$). As expected, the model also demonstrated strong ability to predict TDT *versus* NTDT status (Area Under the ROC Curve: AUC=0.769; $P<0.001$).

A score, called Thalassemia Severity Score (TSS), was derived from the linear predictor score distribution to obtain a scale of increasing severity ranging from 0 to 10. As for the original model mentioned above, when applied to the testing set, as expected, the derived score also highly correlates with the probability of transfusion-free survival ($P<0.001$) and TDT *versus* NTDT status (AUC=0.769; $P<0.001$). Finally, four significantly different survival curves can be drawn from the distribution of TSS ($P<0.001$) (for both sets see Figure 1). Details of the four groups are presented in *Online Supplementary Table S2*. The TSS calculation has been automated and is freely available at *http://tss.unica.it*.

Discussion

This multicenter study involving 890 β-thalassemia patients of six different centers from three countries of the Mediterranean basin demonstrated the ability to accurate-

ly score and predict the hematologic severity of the disease using genetic markers and the patient's age at the start of regular transfusion.

Besides the accurate assessment of the effect of each genetic predictor on the disease, the present results could also be helpful in patient management. The prediction ability of the score could make a significant contribution to pre-conception screening by adapting genetic counseling according to score; however, its use after birth as a hematologic severity scale should be intended as a support to the wide range of clinical information available to medical staff.

In the present study, we created an easy-to-interpret score that uses a scale of increasing severity ranging from 0 to 10. The TSS can be used as a pre-natal predictive score to refine genetic counseling by completing β-thalassemia pre-conception and pre-natal screening whenever available, and to evaluate the severity of the disease in support of the clinically difficult decisions as to whether to start or not start a transfusion program. After birth, it could be used as a standardized scale to better define the hematologic severity of patients and overcome the major-intermedia dichotomy. Such a scale could be of great support when deciding whether to perform a stem cell transplantation: the latter are more efficient if performed early in life (especially with cord blood) but the disease phenotype can take four or even five years to stabilize, and in such a situation the possibility of anticipating the phenotype could be essential. The TSS could also be useful to sustain clinical decisions after the transfusion program has started, to further develop more complete severity scales that would incorporate clinical measurements with the TSS, or even to identify outlier patients by checking their clinical severity status against expected hematologic severity based on genetic background. Another application we envisage is the possibility to better categorize thalassemia patients enrolled in clinical trials for erythropoiesis stimulating factors. Using only genetic markers, the TSS needs a single assessment, is already available during pregnancy, has life-long validity, and could, therefore, be a good solution to maintain an informative profile as to patient's predisposition for severity. This is important as, in order to improve quality of life, transfusions are increasingly used for patients who, in the past, would not have been transfused, and this will probably result in a drastic reduction in non-transfused cohorts in the future.

The benefits of using only genetic predictors are important, but there are also drawbacks. First, some centers might not have the necessary facilities to genotype the markers. Second, not all markers have proven their causality: if *HBB* and *HBA* genes defects are evidently responsible for lack or malfunction of the coded proteins, the direct implication and mechanisms of action of rs1427407, rs10189857, rs9399137 and even rs7482144 have not yet been fully demonstrated. These markers could be surrogates of the causative variants, and even though the final statistical model would not vary much due to the great amount of genetic association signal explained by these variants, the score should be up-dated if a different variant proved to be causative. Third, even if the populations in which these variants were successfully associated represent all the main ethnic groups, not all the populations have been tested and some particular linkage disequilibrium structures might render the model less accurate in untested populations.[7–10,14–18] However, considering only
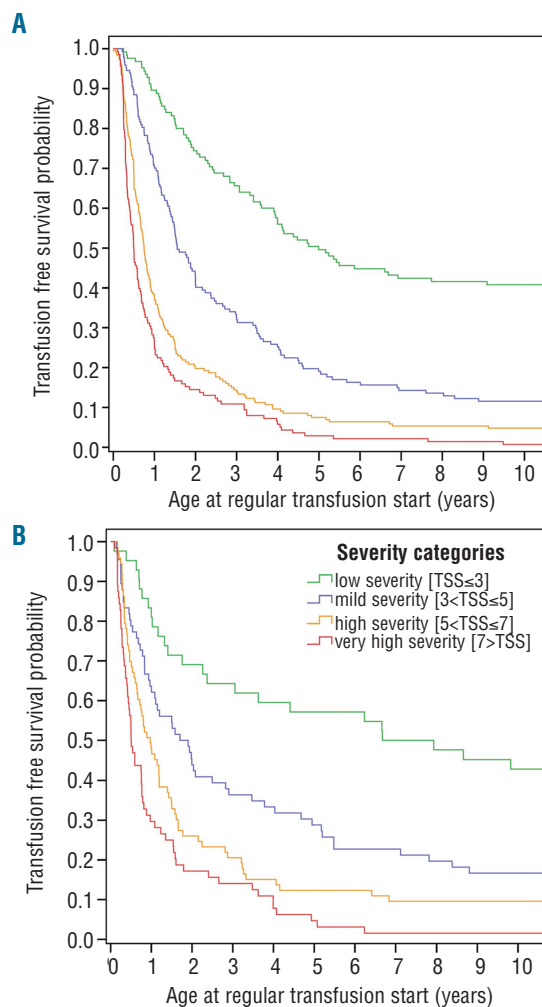
Figure 1. Kaplan-Meier analysis of transfusion-free survival prediction by TSS in training (**A**) and testing (**B**) sets.

countries with efficient health systems and screening programs, this retrospective study is a robust tool against mortality-driven bias. In such a context, mortality rates before ten years old are negligible (0 in our cohort), while over 90% of study patients were diagnosed, followed-up and treated before this age.[21,22] However, only cross validation from prospective studies, also including other thalassemia populations, will ensure full external validity of the score.

Authorship and Disclosures
Information on authorship, contributions, and financial & other disclosures was provided by the authors and is available with the online version of this article at www.haematologica.org.

References

1. Cao A, Galanello R. Cao A, Galanello R. Beta-Thalassemia. In: Pagon RA, Bird TC, Dolan CR, Stephens K, editors. GeneReviews [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2000 [updated 2010].

2. Danjou F, Anni F, Galanello R. Beta-thalassemia: from genotype to phenotype. Haematologica. 2011;96(11):1573-1575.

3. Gilman J, Huisman T. DNA sequence variation associated with elevated fetal G gamma globin production. Blood. 1985; 66(4):783-787.

4. Labie D, Dunda-Belkhodja O, Rouabhi F, et al. The -158 site 5' to the G gamma gene and G gamma expression. Blood. 1985;66(6):1463-1465.

5. Garner C, Tatu T, Reittie JE, et al. Genetic influences on F cells and other hematologic variables: a twin heritability study. Blood. 2000;95(1):342-346.

6. Lettre G, Sankaran VG, Bezerra MAC, et al. DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. Proc Natl Acad Sci USA. 2008;105(33):11869-11874.

7. Sedgewick AE, Timofeev N, Sebastiani P, et al. BCL11A is a major HbF quantitative trait locus in three different populations with beta-hemoglobinopathies. Blood Cells Mol Dis. 2008;41(3):255-258.

8. So C-C, Song Y-Q, Tsang ST, et al. The HBS1L-MYB intergenic region on chromosome 6q23 is a quantitative trait locus controlling fetal haemoglobin level in carriers of beta-thalassaemia. J Med Genet. 2008; 45(11):745-751.

9. Creary LE, Ulug P, Menzel S, et al. Genetic variation on chromosome 6 influences F cell levels in healthy individuals of African descent and HbF levels in sickle cell patients. PloS One. 2009;4(1):e4218.

10. Menzel S, Thein SL. Genetic architecture of hemoglobin F control. Curr Opin Hematol. 2009;16(3):179-186.

11. Galanello R, Sanna S, Perseu L, et al. Amelioration of Sardinian beta0 thalassemia by genetic modifiers. Blood. 2009; 114(18):3935-3937.

12. Badens C, Joly P, Agouti I, et al. Variants in genetic modifiers of β-thalassemia can help to predict the major or intermedia type of the disease. Haematologica. 2011; 96(11):1712-1714.

13. Danjou F, Anni F, Perseu L, et al. Genetic modifiers of β-thalassemia and clinical severity as assessed by age at first transfusion. Haematologica. 2012;97(7):989-993.

14. Menzel S, Garner C, Gut I, et al. A QTL influencing F cell production maps to a gene encoding a zinc-finger protein on chromosome 2p15. Nat Genet. 2007; 39(10):1197-1199.

15. Galarneau G, Palmer CD, Sankaran VG, et al. Fine-mapping at three loci known to affect fetal hemoglobin levels explains additional genetic variation. Nat Genet. 2010;42(12):1049-1051.

16. Solovieff N, Milton JN, Hartley SW, et al. Fetal hemoglobin in sickle cell anemia: genome-wide association studies suggest a regulatory region in the 5' olfactory receptor gene cluster. Blood. 2010;115(9):1815-1822.

17. Bhatnagar P, Purvis S, Barron-Casella E, et al. Genome-wide association study identifies genetic variants influencing F-cell levels in sickle-cell patients. J Hum Genet. 2011; 56(4):316-323.

18. Farrell JJ, Sherva RM, Chen Z-Y, et al. A 3-bp deletion in the HBS1L-MYB intergenic region on chromosome 6q23 is associated with HbF expression. Blood. 2011; 117(18):4935-4945.

19. Galanello R, Sollaino C, Paglietti E, et al. Alpha-thalassemia carrier identification by DNA analysis in the screening for thalassemia. Am J Hematol. 1998;59(4):273-278.

20. Sutton M, Bouhassira EE, Nagel RL. Polymerase chain reaction amplification applied to the determination of beta-like globin gene cluster haplotypes. Am J Hematol. 1989;32(1):66-69.

21. Ladis V, Chouliaras G, Berdoukas V, et al. Survival in a large cohort of Greek patients with transfusion-dependent beta thalassaemia and mortality ratios compared to the general population. Eur J Haematol. 2011;86(4):332-338.

22. Cao A, Kan YW. The prevention of thalassemia. Cold Spring Harb Perspect Med. febbraio 2013;3(2):a011775.