Supplemental Online Materials

eTable 1: Distribution of *IFNL4* polymorphisms rs368234815 and rs12979860 in this study and European populations in the 1000 Genomes Project

	#rs368234815-TT/dG, %				**rs12979860-C/T, %**		
Genotypes	NCI set		*Europeans	Genotypes	DBMT validation set		*Europeans
	Recipients N=316	Donors N=404	N=503		Recipient N=1225	Donor N=1245	N=503
Genotype Frequency							
TT/TT	49.4	46.2	47.7	C/C	44.9	44.7	48.1
TT/dG	38.9	43.5	42.1	C/T	43.4	43.0	41.9
dG/dG	11.7	10.3	10.1	T/T	11.8	12.4	9.9
Allele Frequency							
TT	68.8	65.9	68.8	C	66.6	66.1	69.1
dG	31.2	34.0	31.2	T	33.4	33.8	30.9

#In the 1000 Genomes project, rs368234815-TT/dG is represented by rs74597329-T/G and rs11322783-T/-
*Europeans from the 1000 Genomes – 5 populations of European ancestry (EUR) include Utah Residents (CEPH) with northern and western European ancestry (CEU), Toscani in Italy (TSI), Finnish in Finland (FIN), British in England and Scotland (GBR), and Iberian population in Spain (IBS).
Source: http://grch37.ensembl.org/Homo_sapiens/Variation/Population?db=core;r=19:39738655-39739655;v=rs74597329;vdb=variation;vf=584701978
http://grch37.ensembl.org/Homo_sapiens/Variation/Population?db=core;r=19:39738287-39739287;v=rs12979860;vdb=variation;vf=584538288

eTable 2: Distribution of *IFNL4*-rs117648444 in donors of the NCI and validation DBMT sets and European populations in the 1000 Genomes Project

rs117648444, %			
Genotypes and alleles	NCI set	DBMT validation set	*Europeans
	Donors N=404	Donor N=1241	N=503
G/G	81.1	82.1	80.8
G/A	17.3	17.1	17.2
A/A	1.2	0.8	2.0
G	89.8	90.6	89.4
A	9.9	9.3	10.6

*Europeans from the 1000 Genomes – 5 populations of European ancestry (EUR) include Utah Residents (CEPH) with northern and western European ancestry (CEU), Toscani in Italy (TSI), Finnish in Finland (FIN), British in England and Scotland (GBR), and Iberian population in Spain (IBS).
Source: http://grch37.ensembl.org/Homo_sapiens/Variation/Population?db=core;r=19:39738078-39739078;v=rs117648444;vdb=variation;vf=584897605

eTable 3: Designation and frequencies of the IFN-λ4-protein variants in HCT donors from the NCI and DBMT validation sets

IFN-λ4-P70S protein status	IFN-λ4 activity	NCI Total=403 N (%)	DBMT Total=1241 N (%)	rs368234815/ rs12979860	rs117648444
IFNL4-Null	Null	177 (43.9)	556 (44.8)	TT/TT or C/C	G/G
Only IFN-λ4-S70	Weak	56 (13.9)	163 (13.1)	TT/dG or C/T	A/G
				dG/dG or T/T	A/A
At least one copy of IFN-λ4-P70	Strong	170 (42.2)	522 (42.1)	TT/dG or C/T	G/G
				dG/dG or T/T	A/G
				dG/dG or T/T	G/G

See also eFigure 1

eTable 4: Genotype frequencies of the *IFNL4* polymorphisms rs117648444 and rs368234815 in HCT donors from the NCI set and rs117648444 and rs12979860 in DBMT validation set

IFNL4 polymorphisms	rs117648444		
	G/G	G/A	A/A
NCI Set			
rs368234815	N (%)		
TT/TT	177 (43.9)	0	0
TT/dG	127 (31.5)	51 (12.7)	0
dG/dG	24 (6.0)	19 (4.7)	5 (1.2)
DBMT Validation			
Rs12979860	N (%)		
C/C	556 (44.8)	0	0
C/T	380 (30.6)	153 (12.3)	0
T/T	83 (6.7)	59 (4.7)	10 (0.8)

eTable 5: Risk of non-relapse mortality and overall survival associated with recipient *IFNL4* genotype - presence of rs368234815-dG allele in NCI set and rs12979860-T allele in DBMT

	Non-relapse Mortality (NRM)		Overall survival (OS)	
	N events/total	HR[a] (95% CI) P-value	N events/total	HR[a] (95% CI) P-value
Discovery NCI set	93/318	1.12 (0.84-1.51) p=0.44	172/319	1.01 (0.81-1.26) p=0.96
DBMT validation set	390/1223	1.11 (0.95-1.28) P=0.19	844/1225	1.07 (0.97-1.19) P=0.17
DBMT combined	481/1493	1.12 (0.98-1.27) P=0.10	999/1495	1.07 (0.97-1.17) P=0.16

[a] Additive genetic models adjusted for: donor and recipient age, GvHD prophylaxis, use of total body irradiation, and stratified by graft type

eTable 6: Expression of all human interferons based on analysis of single-cell RNA-sequencing of 32,333 CD34+/CD38- bone marrow/progenitor cells from three healthy individuals

Interferon	Transcript IDs	Positive cells, N	Positive cells of the total cells (N=32,333), %	Positive cells of the total IFN-expressing cells (N=261), %
IFNL1	ENSG00000182393	23	0.071	8.8
IFNL2	ENSG00000183709	3	0.009	1.1
IFNL3	ENSG00000197110	11	0.03	4.2
IFNL4	**ENSG00000272395**	**217**	**0.67**	**83.1**
IFNA1	ENSG00000197919	0	0	0.0
IFNA2	ENSG00000188379	0	0	0.0
IFNA4	ENSG00000236637	0	0	0.0
IFNA5	ENSG00000147873	0	0	0.0
IFNA6	ENSG00000120235	0	0	0.0
IFNA7	ENSG00000214042	0	0	0.0
IFNA8	ENSG00000120242	0	0	0.0
IFNA10	ENSG00000186803	0	0	0.0
IFNA13	ENSG00000233816	0	0	0.0
IFNA14	ENSG00000228083	0	0	0.0
IFNA16	ENSG00000147885	0	0	0.0
IFNA17	ENSG00000234829	0	0	0.0
IFNA21	ENSG00000137080	0	0	0.0
IFNB1	ENSG00000171855	4	0.012	1.5
IFNB2	ENSG00000136244	3	0.009	1.1
IFNE	ENSG00000184995	0	0	0.0
IFNG	ENSG00000111537	0	0	0.0
IFNK	ENSG00000147896	0	0	0.0
Total		**261**	**0.8**	**100**

eTable 7: Risk of non-relapse mortality and overall survival associated with donor *IFNL4* genotype (rs12979860) by CMV serostatus

CMV donor-recipient seropositive		
	HR (95% CI)*, *P*	
T/T vs. C/C	1.99 (1.10-3.59) 0.02	1.22 (0.80-1.84) 0.34
C/T vs. C/C	1.41 (0.91-2.17) 0.11	0.93 (0.69-1.25) 0.63
CMV donor-recipient seronegative		
T/T vs. C/C	1.70 (0.95-3.06) 0.07	1.40 (0.95-2.08) 0.08
C/T vs. C/C	1.24 (0.83-1.85) 0.23	1.27 (0.87-1.45) 0.35
Positive donor CMV serostatus		
T/T vs. C/C	1.69 (1.09-2.64) 0.01	1.16 (0.84-1.6) 0.36
C/T vs. C/C	1.26 (0.91-1.75) 0.14	0.97 (0.78-1.22) 0.84
Negative donor CMV serostatus		
T/T vs. C/C	1.70 (1.21-2.38) 0.001	1.35 (1.06-1.73) 0.01
C/T vs. C/C	1.19 (0.93-1.51) 0.16	1.09 (0.93-1.29) 0.26

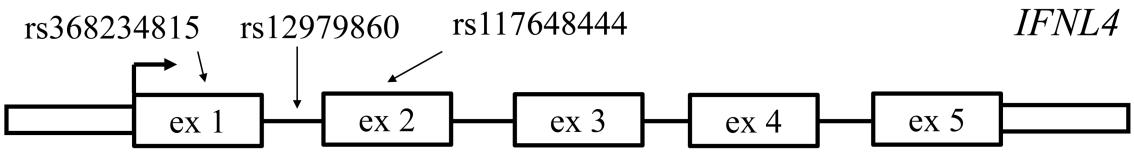
*Models are adjusted for donor and recipient age, GvHD prophylaxis, use of total body irradiation, and stratified by graft type

eTable 8: Number of events under study by donor *IFNL4* genotype

Events	TT/TT[#] (N=177)	TT/dG (N=179)	dG/dG (N=48)	C/C[#] (N=556)	C/T (N=535)	T/T N=154)
	Number of events					
Overall survival*	91	102	30	390	370	115
	NCI set rs368234815			DBMT validation rs12979860		
Non-relapse mortality	38	59	18	165	182	60
	DBMT overlapping set rs12979860			DBMT validation rs12979860		
Causes of Death at 1 year post-HCT	C/C (N=164)	C/T (N=145)	T/T (N=39)	C/C	C/T	T/T
	Number of events					
Relapse death	27	20	5	153	122	38
Infection death	7	3	4	26	39	15
GvHD death	6	10	4	27	38	13
Organ failure death	2	9	2	40	39	7

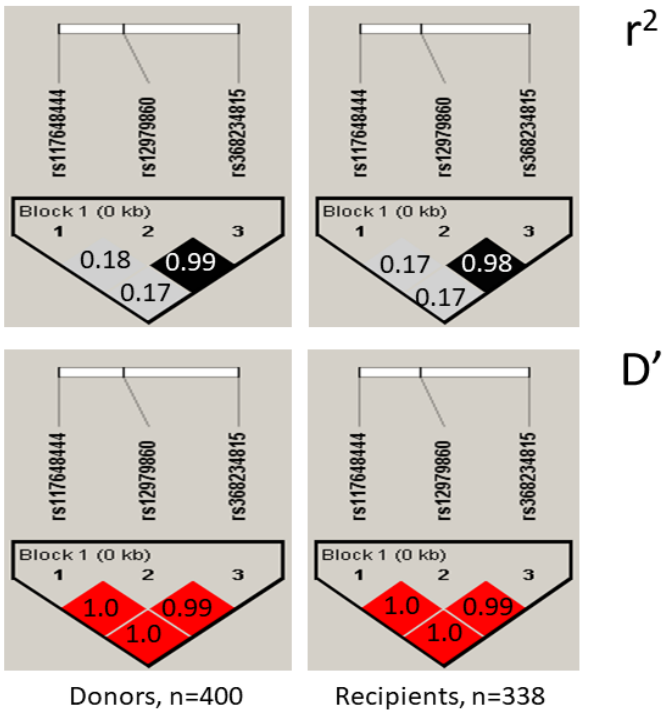
[#] *IFNL4*-Null

*****event: death from any cause



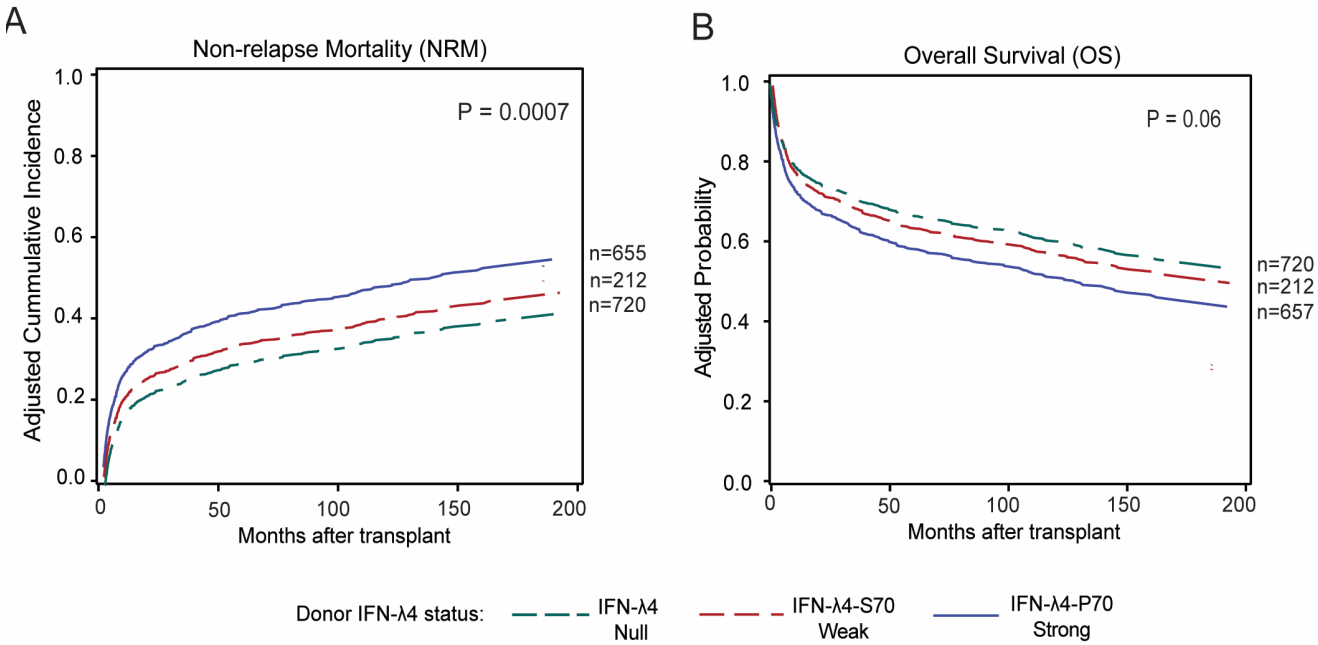
Haplotype combination	Haplotypes	rs368234815 alleles	rs12979860 alleles	rs117648444 alleles	IFN-λ4 protein status	IFN-λ4 activity
1	Haplotype 1	TT	C	G	*IFNL4*-Null	No IFN-λ4
	Haplotype 2	TT	C	G		
2	Haplotype 1	TT	C	G	*IFNL4*-positive	Only IFN-λ4-S70 Weak
	Haplotype 2	dG	T	A		
3	Haplotype 1	dG	T	A		
	Haplotype 2	dG	T	A		
4	Haplotype 1	TT	C	G	*IFNL4*-positive	At least one copy of IFN-λ4-P70 Strong
	Haplotype 2	dG	T	G		
5	Haplotype 1	dG	T	A		
	Haplotype 2	dG	T	G		
6	Haplotype 1	dG	T	G		
	Haplotype 2	dG	T	G		

eFigure 1: Genetic variants within the IFNL4 gene explored in this study. Shown are individual alleles of three genetic variants and their haplotypes. Six haplotype combinations represent three functional states of IFN-λ4 protein – Null, Weak and Strong.

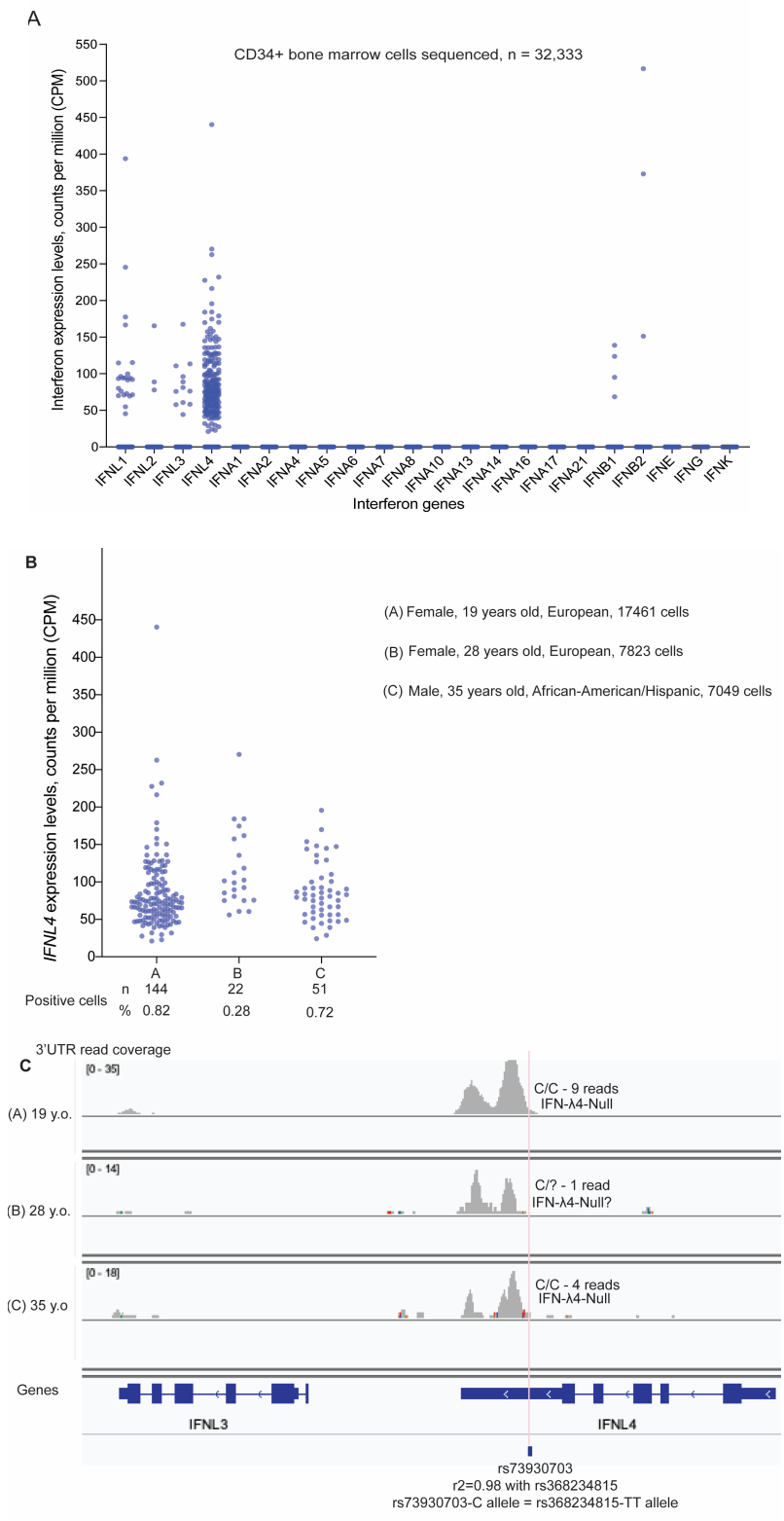


eFigure 2: Linkage disequilibrium (LD) plots for the *IFNL4* polymorphisms in the NCI set

The analysis and plotting were performed using Haploview software (http://www.broad.mit.edu/mpg/haploview/).
The color scheme and values on upper and lower row plots represent pairwise r^2 and D' coefficients, respectively. Only individuals with genotype data for all three polymorphisms studied were used for analysis.



eFigure 3: Donor IFN-λ4-P70S status and adjusted non-relapse mortality (A), and overall survival (B) after unrelated donor HCT for patients with acute leukemia in the combined DBMT set



eFigure 4: Single-cell RNA sequencing (scRNA-seq) of 32,333 CD34+ bone marrow cells from 3 healthy donors (HIV, HCV and HBV-negative). **A**) mRNA expression levels of all human interferons

are shown as counts per million (CPM) in each of the sequenced cells (n = 32,333). **B)** Distribution of the sequenced *IFNL4*-expressing bone marrow cells between these donors. **C**) Integrative Genomics Viewer (IGV) plot showing expression of *IFNL4* but not *IFNL3* mRNA in bone marrow cells of these donors. Digital gene expression profiling was done using short 3'UTR reads generated with 10X Genomics system (see methods for details). Genotypes of rs73930703 located in 3'UTR of *IFNL4* were scored directly from scRNA-seq reads. All three individuals are likely to be *IFNL4*-Null because rs73930703-C allele is linked with rs368234815-TT, with r^2=0.98 in Europeans in the 1000 Genomes Project. Expression of *IFNL4* mRNA is detected in all individuals since rs368234815 does not affect mRNA expression of *IFNL4*, but IFN-λ4 protein will be produced only from *IFNL4* mRNA transcripts with the rs368234815-dG allele. Data source: https://www.ebi.ac.uk/gxa/sc/experiments/E-HCAD-6/downloads

Supplemental Methods: Mean Risk Stratification (MRS) concept and calculation

Absolute risk (AR) and the absolute risk difference (RD) are widely regarded as the key metrics for evaluating the utility of diagnostic or prognostic testing.[1,2] However, the RD does not account for the frequency of test positivity. If testing positive were extremely rare, then even a large RD may not contribute significant health impact at the population level.

We have previously proposed Mean Risk Stratification (MRS) as a measure of diagnostic test performance that accounts for both absolute risk of a disease or an outcome (i.e. prevalence and predictive values) and test positivity.[3,4] MRS is a measure of risk stratification: the average absolute change in outcome risk revealed by using a test. In this paper, the outcome is risk of non-relapse mortality (NRM) and the test is the dichotomized donor *IFNL4* genotype.

MRS is derived by the following argument: without *IFNL4* test results, each patient is assigned the same population-average risk of NRM, denoted as π=P(D+). Knowledge of donor *IFNL4* genotype affects predicted NRM risk for a patient in relation to population-average. Two outcomes are possible:

1. *IFNL4* test is positive with probability t=P(M+). The person's risk increases from π=P(D+) to Positive Predictive Value (PPV=P(D+|M+)), an increase of $PPV - \pi$.
2. *IFNL4* test is negative with probability P(M–)=1– t. The person's risk decreases from π=P(D+) to complement of Negative Predictive Value: $cNPV$=1–NPV=P(D+|M–). The person's risk decreases by $\pi - cNPV$.

We will consider 2 definitions for *IFNL4* test-positivity: having either 1 or 2 copies of the rs368234815-dG allele; here, we used the proxy highly linked rs12979860-T allele that was available

for all the DBMT set; genotype data for rs368234815 was only available for the small set of DBMT set that overlapped with the NCI set.

MRS is a weighted average of the change in risk among those patients who receive bone marrow transplant from donors who test positive *vs.* negative, respectively:

$MRS = \{PPV - \pi\}t + \{\pi - cNPV\}(1 - t)$.

MRS is the average difference between predicted post-test individual risk and population-average (pre-test) risk. Specifically, here MRS is the average change in risk of non-relapse mortality (NRM) of the recipient revealed by knowing donor's *IFNL4* genotype. For example, a 10% MRS means that the post-test risk is ±10 percentage points different from pre-test risk, on average. Larger MRS indicates that the test is more informative about a person's risk, and may justify changing clinical management.

MRS interprets the risk difference in light of test positivity

The risk difference for NRM between *IFNL4* test-positives and negatives is denoted as $RD = PPV - cNPV$. MRS is also a function of *RD* and *IFNL4* test-positivity t:

$MRS = 2t(1-t) \times RD$

Thus, the test with greater *RD* might not also have greater MRS if the test positivity is rare. In particular, a rarely positive test will have small MRS, regardless of *RD*. We will use MRS to compare *RD*s for the 2 possible cutpoints for *IFNL4* genotypes.

How high does MRS need to be to suggest clinical usefulness?

MRS, as a measure of risk-stratification, does not account for the benefits, harms, and costs of interventions. Furthermore, the value of MRS depends on the severity of the disease outcome. For

example, a test providing an MRS of 1% for a relatively mild disease is likely to be less clinically useful than a test providing 1% MRS for risk of death (all other things being equal). Thus, no single MRS value can suggest clinical usefulness across diseases and outcomes.

However, we can consider the question by calculating the implicit MRS underlying an observed change in medical practice. As an example and reference point, we used a *bona fide* clinically significant change in donor selection based on the identification of 8/8 HLA matching as optimal donor selection criteria. That is, requiring an 8/8 HLA match (*versus* 7/8) changes the risk of NRM, for which there is an underlying MRS. This might represent a minimal MRS required to use a test of similar benefit and cost to HLA genotyping, such as *IFNL4* genotyping, for donor selection.

To calculate the MRS for 7/8 *vs.* 8/8 HLA matching, we used the hazard ratio of 1.4 for 7/8 *vs* 8/8 HLA matching for hazard of non-relapse mortality from the Lee et al paper that provided critical evidence underlying the change to requiring 8/8 HLA matching.[5] We define "test-positive" as the riskier condition, which is 7/8 HLA matching: in Lee et al, 34.9% of those with either 7/8 or 8/8 HLA matches were 7/8 matches (t). [5] For the baseline risk of NRM in those with 8/8 HLA matching (which is "test-negative"), we use the overall NRM risk observed in our study (32.5%=*cNPV*) because the risk observed in Lee et al (39.1%) was based on historic data with older treatment regimens. For those with 7/8 HLA matching, we estimate that risk of NRM (*PPV*) is (32.52%/67.48% × 1.40) / (1 + 32.52%/67.48% × 1.40) = 40.3% (note that the HR=1.40 is treated like an odds-ratio (not a relative risk) because NRM is common). This information suffices to calculate MRS for 7/8 vs 8/8 HLA matching via the risk difference ($RD = PPV - cNPV$ = 40.3% – 32.5% = 7.8%):

$MRS = 2t(1-t) \times RD = 2 \times 34.9\% \times 65.1\% \times 7.8\% = 3.5\%$.

Because 8/8 HLA matching is the current clinical standard, its MRS=3.5% for NRM implicitly sets a standard as the minimal MRS required to adopt a test with the harms and costs of HLA testing.

We also calculate MRS for 7/8 vs. 8/8 HLA matching, for the outcome of overall mortality. We will use the HR=1.25 from Lee et al[5] and combine with the risk of overall mortality in those with 8/8 HLA matching in our study (67.67%=*cNPV*). For those with 7/8 HLA matching, we estimate that risk of overall mortality (*PPV*) is (67.67%/32.33% × 1.25) / (1 + 67.67%/32.33% × 1.25) = 72.3%. The risk difference is $RD = PPV - cNPV$ = 72.3% – 67.7% = 4.68%. The MRS for overall mortality is $MRS = 2t(1-t) \times RD$ = 2 × 34.9% × 65.1% × 4.68% = 2.1%.

Comparison of MRS to AUC

Although AUC is normally calculated for continuous biomarker, AUC for a binary biomarker equals the average of sensitivity and specificity.[6] MRS can be written as a function of the dichotomous AUC:

$MRS = 4(AUC-0.5) \times \pi(1-\pi)$,

where π is disease prevalence. Thus MRS interprets the AUC in light of disease prevalence. Note that a high AUC, but for a rare disease, might imply small MRS. For a perfect biomarker (AUC=1), the maximum MRS is achieved: $2\pi(1-\pi)$. However, if a disease is too rare, even a perfect biomarker has a small MRS.

Vice versa, a modest AUC, but for a common and important outcome, might imply a meaningfully large MRS. Given the MRS and π, it is easy to solve the above equation to calculate the AUC: $AUC = 0.5 + MRS/\{4\pi(1-\pi)\}$.

For the NRM in 8/8 *vs* 7/8 HLA matching, the AUC=0.539, and for overall mortality the AUC=0.525. These AUCs are very small, and most such biomarkers would not be considered useful. However, NRM and overall mortality are the most important outcomes, and unfortunately are quite common over 10 years (π=35% and π=69% respectively). Because 8/8 HLA matching is now the standard, the MRSs of 3.5% and 2.1% are implicitly meaningful. This is an example situation where a very low AUC suffices to be clinically useful.

Supplemental Results: Mean Risk Stratification (MRS) for *IFNL4* genotyping *versus* HLA matching

MRS for non-relapse mortality

There are 2 possible cutpoints for the 3 *IFNL4* genotype pairs for donor selection (presented scenarios are based on genotypes of the functional polymorphism *IFNL4*-rs368234815): based on recommending any TT genotype (i.e. TT/TT or TT/dG; "any TT" or avoiding dG/dG) or recommending only TT/TT ("only TT, *IFNL4*-Null"). e**Figure 5** compares the performance of these two cutpoints for risk of non-relapse mortality (NRM). For absolute reduction in NRM, "any TT" appeared superior to "only TT" (11.0% vs. 6.4%). However, "any TT" only rejects dG/dG donors, which is only 12.1% of donors; thus only 12.1% of patients will benefit from *IFNL4* genotype information to reduce their risk of NRM. In contrast, "only TT" rejects 54.8% of donors by accepting only TT/TT donors, thus allowing 54.8% of patients to benefit from *IFNL4* genotyping information. Thus, in spite of having the smaller absolute risk reduction, "only TT" has superior MRS to "any TT" (3.2% vs. 2.4%).

The MRS means that the selection of TT/TT (*IFNL4*-Null) donors decreases risk of NRM by 3.2% on average. In comparison, the selection of 8/8 HLA match donors decreases risk of NRM by 3.5% on average (see **Supplemental Methods**). Since 8/8 HLA matching is the current standard, its MRS of 3.5% was *implicitly* considered useful enough for clinical use to reduce NRM. *IFNL4* genotyping has lower MRS than HLA matching for reducing NRM, but perhaps not by a large amount, suggesting possible clinical usefulness for including *IFNL4* genotyping as an additional test for HCT donor selection.

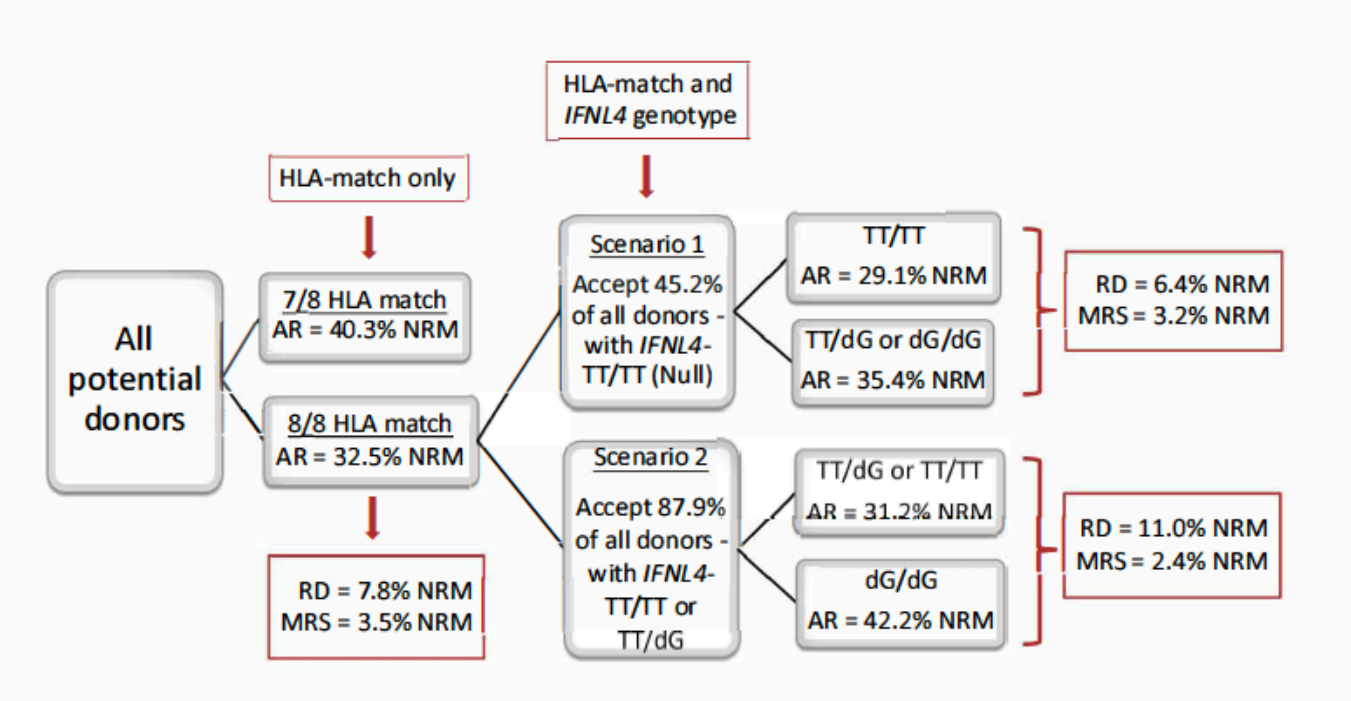
Requiring 8/8 HLA match donors has modest AUC=0.539 for NRM (see **Supplemental Methods**). Similarly, *IFNL4* "only TT" and "any TT" have even lower AUCs (0.535 and 0.526, respectively). However, the MRS for *IFNL4* genotyping is meaningful because NRM is an important outcome and unfortunately is quite common over 10 years (π=33%). Thus the very low AUC could be clinically useful simply because NRM is so important and common.

MRS for post-HCT mortality

Supplemental eFigure 6 compares the performance of the two *IFNL4* genotyping cutpoints for risk of overall mortality. For absolute reduction in overall mortality, "any TT" appeared superior to "only TT" (4.4% vs. 2.9%). However, "any TT" only rejects dG/dG donors, which is only 12.1% of donors; thus only 12.1% of patients will benefit from *IFNL4* genotyping information to reduce their risk of mortality. In contrast, "only TT" rejects 54.8% of donors by accepting only TT/TT donors, thus allowing 54.8% of patients to benefit from *IFNL4* genotyping information. Thus, in spite of having the smaller absolute risk reduction, "only TT" has better MRS to "any TT" (1.4% vs. 0.9%).

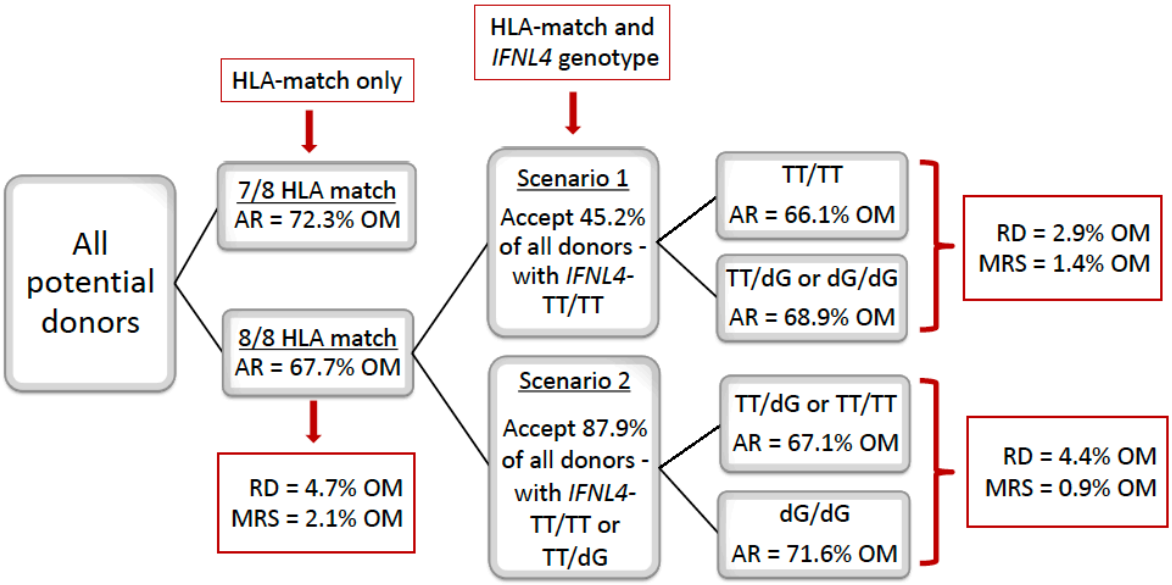
The MRS means that the selection of TT/TT donors reduces the overall mortality by 1.4% on average. In comparison, the selection of 8/8 HLA match donors reduces the overall mortality by 2.1% on average (see **Supplemental Methods**). Since 8/8 HLA matching is the current standard, its MRS of 2.1% was *implicitly* considered useful enough for clinical use to reduce mortality. *IFNL4* genotyping may have lower MRS than HLA matching for overall mortality, but perhaps not by a large amount.

Requiring 8/8 HLA match donors has modest AUC=0.525 for overall mortality (see **Supplemental Methods**). Similarly, *IFNL4* "only TT" and "any TT" have even lower AUCs (0.517 and 0.511 respectively). However, the MRS for *IFNL4* genotyping is meaningful because overall mortality is an important outcome and unfortunately is quite common over 10 years (π=68%). Thus the very low AUC could be clinically useful simply because NRM is so common.



Supplemental eFigure 5. Schematic diagram comparing different scenarios of donor selection based on HLA-matching and IFNL4 genotype and their effects on absolute risk measures for non-relapse mortality after HCT.

Abbreviations: AR (Absolute Risk); RD (Risk Difference); MRS (Mean Risk Stratification); NRM (Non-relapse mortality)



Supplemental eFigure 6. Absolute Risk (AR) measures for mortality and different scenarios of donor selection based on HLA matching and *IFNL4* genotyping in the DBMT combined set. Abbreviations: AR (Absolute Risk); RD (Risk Difference); MRS (Mean Risk Stratification); OM (overall mortality)

Footnotes

Risk stratification for donor selection based on HLA matching alone:

Our study was conducted among those with 8/8 HLA-matched donors and mortality risk was 67.7%. We calculate 72.3% risk for 7/8 HLA-matched donors by multiplying the odds of 67.7% by the hazard ratio of 1.25 (as observed in Lee et al.[6]), and convert to a probability: (67.67%/32.33% × 1.25) / (1 + 67.67%/32.33% × 1.25) = 72.3%. Risk difference (RD) for HLA 8/8 vs. 7/8 match was calculated as 72.3% − 67.7% = 4.7%. The mean risk stratification (MRS) for HLA 8/8 vs. HLA 7/8 was calculated based on RD=4.7% and the fraction of patients in the 7/8 HLA group (t=34.9%): $MRS = 2t(1–t)\times RD = 2 \times 34.9\% \times 65.1\% \times 4.7\% = 2.1\%$.

Risk stratification for donor selection based on 8/8 HLA matching and *IFNL4* genotype:

Scenario 1 is to accept donors only with *IFNL4*-TT/TT genotypes (45.2% of all donors). Scenario 2 is to accept donors with *IFNL4*-TT/TT or TT/dG genotypes (87.9% of all donors). MRS for Scenario 1 is calculated based on *RD*=2.9% and proportion of patients in TT/TT genotype group (*t*=45.2%): $MRS = 2t(1-t) \times RD = 2 \times 45.2\% \times 54.8\% \times 2.9\% = 1.4\%$. MRS for Scenario 2 is calculated based on *RD*=4.5% and proportion of the combined TT/dG and dG/dG genotype group (*t*=87.9%): $MRS = 2t(1-t) \times RD = 2 \times 87.9\% \times 12.1\% \times 4.4\% = 0.9\%$. See **Supplemental Methods** for details.

Supplemental References

1. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol*. 2004 May 1;159(9):882-90.
2. Wentzensen N, Wacholder S. From differences in means between cases and controls to risk stratification: a business plan for biomarker development. *Cancer Discov*. 2013 Feb;3(2):148-57.
3. Katki, HA and Schiffman, M. A novel metric that quantifies risk stratification for evaluating diagnostic tests: The example of evaluating cervical-cancer screening tests across populations. *Prev Med,* 2018*, 110*, 100-105.
4. Katki, HA. Quantifying risk stratification provided by diagnostic tests and risk predictions: Application to population mutation screening. *Stat Med*. 2019 Jul 20;38(16):2943-2955. doi: 10.1002/sim.8163.
5. Lee SJ, Klein J, Haagenson M, Baxter-Lowe LA, Confer DL, Eapen M, Fernandez-Vina M, Flomenberg N, Horowitz M, Hurley CK, Noreen H, Oudshoorn M, Petersdorf E, Setterholm M, Spellman S, Weisdorf D, Williams TM, Anasetti C. High-resolution donor-recipient HLA matching contributes to the success of unrelated donor marrow transplantation. *Blood*. 2007 Dec 15;110(13):4576-83.
6. Cantor SB, Kattan MW. Determining the area under the ROC curve for a binary diagnostic test. Med Decis Making. 2000;20(4):468-470