PROGNOSTIC AND PREDICTIVE PERFORMANCE OF R-ISS WITH SKY92 IN THE ELDERLY MULTIPLE MYELOMA

HOVON-87/NMSG-18 TRIAL

Supplemental Documents

Sample collection and workup for gene expression

Aspiration of bone marrow - obtained for the purpose of biobanking as part of the HOVON-87/NMSG-18 (HO87) protocol (www.hovon.nl) - was performed at the local hospitals using a punction needle on the iliac crest. [1]

For 180 patients out of 636 included patients (Figure S1 and S2), the following criteria were met: i) percentage of plasma cells after CD138 based plasma cell enrichment equal to or higher than 80%; enrichments were performed within 24-48 hours after puncture using the Human CD138 Positive Selection Kit magnetic beads based enrichment system (Stemcell technologies); ii) sufficient RNA yield and quality for the MMprofiler™ and iii) consent for the gene expression side study was given.

RNA was isolated using the AllPrep (Qiagen); RNA quantity and quality were assessed using Nanodrop (ThermoScientific) and BioAnalyzer (Agilent), respectively. The MMprofiler™ CE IVD assay (SkylineDx, Rotterdam, the Netherlands) was used to obtain SKY92 scores, classifying a patient as high-risk or standard-risk. Bone marrow aspirates were processed at the Erasmus MC laboratory. RNA isolations were performed, and RNA samples were shipped to the SkylineDx reference lab (Rotterdam, The Netherlands) where subsequent workup was performed. Hybridization cocktails were made using Affymetrix Inc. (Santa Clara, CA) GPR reagent for samples obtained in 2014 or earlier (n = 88), and Affymetrix 3' IVT PLUS reagents for samples obtained after 2014 (n = 92).

All samples passed the following wet lab QC acceptance criteria. RNA purification yield ≥100ng; RNA concentration: ≥20ng/µl; rRNA integrity: 28S/18S ratio ≥0.9; total RNA quality: two distinct bands just below 4000 nucleotides (nt) and 2000 nt after QC; cRNA quality: “Smear” between 200 nt and 4000 nt; fragmented cRNA quality: a relatively thick band between 50 nt and 150 nt.

Data QC acceptance criteria were comparable to the reference set (i.e. IVT %present ≥ 26.6; GPR %present ≥ 20.0). Samples were hybridized to the MMprofiler™ Affymetrix HG-U133 plus2 chips. Subsequently, the SKY92 scores were collected and used in the analyses presented here. The gene expression profiles have been submitted to GEO under accession GSE87900.

Detection of numerical changes for 1q21 and 13q14 and 17p13 was performed using locus specific FISH probes (for 1q21: 1q21/SRD (Kreatech Diagnostics, Amsterdam, The Netherlands) or CKS1B/CDKN2C (Cytocell Ltd, Cambridge, UK), for 13q14 LSI probe 13 or D13S319 (Vysis, Abbott Molecular, Downers Grove, Illinois, USA), and for 17p13 LSI TP53 (17p13.1). Translocations t(11;14)(q13;q32) and t(4;14)(p16;q32) and t(14;16)(q32;q23) were determined using dual color dual fusion probesets. Translocations t(11;14)(q13;q32) was determined by the use of LSI IGH/CCND1 (Vysis, Abbott Molecular). The t(4;14)(p16;q32) was determined using the LSI IGH/FGFR3 probeset (Vysis; Abbott Molecular) or the FGFR3/IGH t(4;14) Poseidon Probeset (Kreatech), and (14;16)(q32;q23) was determined using the LSI IGH/MAF probes (Vysis; Abbott Molecular). For a small number of samples, a SNP array was used to call copy number changes.

Application of SKY92 to CoMMpass

The CoMMpass data (version IA13) was obtained from the MMRF portal (https://research.themmrf.org/). The gene expression data is available as preprocessed data (Salmon TPM values). The preprocessing performed by CoMMpass comprised mapping of 2x83bp reads by SAILFISH v0.6.3 against the GRCh37 reference genome using Ensembl v74 gene annotations. Per sample, at least 60 million read-pairs were generated and the 5’ bias ratio and the 5’/3’ bias ratio had to be >0.4 and >0.65 respectively.

The SKY92 score – developed in the HOVON65/GMMG-HD4 (HO65) trial data - is a weighted summation of the expression given by 92, MAS5 normalized Affymetrix probe-sets.[2,3] Because the CoMMpass data is annotated in terms of Ensembl gene IDs, a translation is required. The HO65 gene expressions were renormalized using MAS5 based on the Brainarray HGU133Plus2_Hs_ENSG CDF file (version 22).[4] This allows for a direct remodeling between the two representations (see section ’Bridging Affymetrix probe-set IDs to Ensembl gene IDs). Only Ensembl gene IDs with an average log2 expression >8 were used for the remodeling.

In total, 19570 Ensembl Gene IDs overlap between CoMMpass and the HO65 expression data. Based on these, the CoMMpass TPM values were normalized per sample towards the HO65/GMMG-HD4 expression data

using robust spline normalization in the Lumi R package (version 2.38) after which the data was log2 transformed.[5] Applying the Ensembl Gene vs Affy probe-set mapping found in the HO65 data, to the CoMMpass data, enables the determination of the SKY92 scores.

A patient is classified as either high- or standard-risk depending on whether the SKY92-score exceeds a predetermined decision threshold. The bridged decision threshold is determined in a 5-fold cross-validation in the HO65, as the value at which a similar proportion of high-risk patients is classified using the remapped data as seen with the original SKY92.

Bridging Affymetrix probe-set IDs to Ensembl gene IDs

The following bridging approach, provides a way to translate expressions in terms of Ensembl gene IDs to expressions in terms of Affymetrix probe-set IDs. Therefore, two paired expression sets are required. Let $\boldsymbol{y}_b \in \boldsymbol{Y}$ be a vector of Affymetrix probe-set expressions indexed by b, and let $\boldsymbol{x}_d \in \boldsymbol{X}$ be a vector of Ensembl gene ID expressions indexed by d, both of length n corresponding to the number of paired subjects. The expressions of all probe-sets and gene IDs are centered and scaled to unit variance.

To describe $\boldsymbol{y}_b$ in terms of $\boldsymbol{X}$, start by modeling the relation $\boldsymbol{y}_b \sim \beta_1 \boldsymbol{x}_d + \varepsilon$ in a linear regression for all d. Then the first index set, contains all indices d for which $\beta_1 \neq 0$ with p<1x10^{-5}. In a principal component analysis, a projection matrix $\boldsymbol{Z_b} = \boldsymbol{W_b R_b}$ containing vectors $\boldsymbol{z_{bf}} \in \boldsymbol{Z}_b$ is obtained based on submatrix $\boldsymbol{W_b} = \boldsymbol{X}(;\boldsymbol{S1_b})$ which is rotated by matrix $\boldsymbol{R_b}$, such that $Var(\boldsymbol{z_{bf}}) > Var(\boldsymbol{z_{b[f+1]}})$, i.e. $\boldsymbol{z_{b1}}$is the first principal component.

The second index set $\boldsymbol{S2}_b$ of top principal components, contains all indices $f \in [1..x]$ for $x = \min\left(\left\lceil\frac{n}{4}\right\rceil, |\boldsymbol{S}_b|\right)$ or is empty if $x = 0$. Then a linear regression is applied to the model $\boldsymbol{y}_b = \boldsymbol{W_b}(;\boldsymbol{S2}_b)\boldsymbol{R_b}(\boldsymbol{S2}_b;)\boldsymbol{\gamma} + \boldsymbol{\varepsilon}$ in order to estimate the corresponding values of $\boldsymbol{\gamma}$.

Performance of SKY92-ISS in HO87 and CoMMpass

Previously, we reported the combination of SKY92 and ISS after systematic analysis of combinations of several prognostic markers.[6] SKY92-ISS is defined to distinguish 4 risk groups: high-risk if SKY92 HR; intermediate high-risk if SKY92 SR and ISS III; intermediate low-risk if SKY92 SR and ISS II; low-risk if SKY92 SR and ISS I. In the HO87, 177 patients with SKY92 and ISS status were classified into high-risk (SKY92 HR; n=23, 13%), intermediate high-risk (SKY92 SR + ISS-III; n=38, 21%), intermediate low-risk (SKY92 SR + ISS-II; n=79, 45%) or low-risk (SKY92 SR + ISS-I; n=37, 21%). The 3 year PFS rates were 10% (95% CI: 3-37%), 29% (95% CI: 18-48%), 29% (95% CI: 21-41%) and 43% (95% CI: 8-30%) for highest to the lowest risk group (p < 0.01; Figure S3E). The 3-year OS rates were 29% (95% CI: 15-56%), 55% (95% CI: 42-74%), 71% (95% CI: 62-82%) and 81% (95% CI: 69-95%) respectively (p < 0.001; Figure S3F).The hazard ratios associated with these groups relative to low-risk were 3.0 (95% CI: 1.6-5.5), 1.4 (95% CI: 0.8-2.5) and 1.4 (95% CI: 0.8-2.3) for PFS, and 4.1 (95% CI: 2.1-7.9), 2.1 (95% CI: 1.1-3.8) and 1.6 (95% CI: 0.9-2.8) for OS, for high-risk, intermediate-high, and intermediate-low classified patients respectively.

In the elderly subset of the CoMMpass data, the proportion of patients classified from SKY92-ISS highest to lowest risk are 26%, 20%, 26% and 28%. %). The 3 year PFS rates were 10% (23% CI: 11-50%), 24% (95% CI: 9-61%), 55% (95% CI: 38-80%) and 62% (95% CI: 44-88%) for highest to the lowest risk group (p < 0.01; Figure S4E). The 3-year OS rates were 44% (95% CI: 27-70%), 58% (95% CI: 40-85%), 83% (95% CI: 69-100%) and 96% (95% CI: 89-100%) respectively (p < 0.001; Figure S4F).The hazard ratios associated with these groups relative to low-risk were 3.9 (95% CI: 1.8-8.6), 2.6 (95% CI: 1.1-6.0) and 1.6 (95% CI: 0.7-3.5), and 21.8 (95% CI: 2.9-165), 11.8 (95% CI: 1.5-95) and 4.2 (95% CI: 0.5-38) for OS, for high-risk, intermediate-high, and intermediate-low classified patients respectively

Table S1. Concordance table. A high concordance is observed between translocations as determined by FISH vs those based on gene expression. The bone marrow samples used for FISH and GEP where not always obtained from the same aspirate. This may explain discrepancies.

		GEP					
		no t(4;14)	**t(4;14)**	**unknown**		**no t(11;14)**	**t(11;14)**	**unknown**
FISH	**no t(4;14)**	142	4	0	no t(11;14)	109	8	0
	t(4;14)	2	12	0	t(11;14)	3	14	0
	Not done	19	1	0	Not done	37	9	0

Table S2. Contrast analysis. (**A**) Six possible subgroups with the absolute number of patients in the HO87 set in brackets, when combining SKY92 with R-ISS. Only the candidate risk groups composed of subgroups with increasing risks are considered, such that all combinations that meet these constraints (**B**) are valid. (**C**) The candidate risk groups composed from the subgroups, into low, intermediate and high given the constraints and an indication whether the smallest risk group contains at least 10 patients in bold. (**D**) Contrast analysis between the high- and low-risk candidate groups. Bold indicates the best option.

A.

		SKY92	
		SR	HR
R-ISS	I	A (26)	D (4)
	II	B (112)	E (14)
	III	C (8)	F (4)

B.

Constraints
A ≤ D
B ≤ E
C ≤ F
A ≤ B
B ≤ C
D ≤ E
E ≤ F

C.

Candidate Risk Group			
Low	Intermediate	High	N≥10
A	B	CDEF	Yes
A	D	BCEF	No
A	**BC**	**DEF**	**Yes**
A	**BD**	**CEF**	**Yes**
A	**BCD**	**EF**	**Yes**
A	**BDE**	**CF**	**Yes**
A	BCDE	F	No
AB	C	DEF	No
AB	D	CEF	No
AB	**CD**	**EF**	**Yes**
AB	**DE**	**CF**	**Yes**
AB	CDE	F	No
AD	**B**	**CEF**	**Yes**
AD	**BC**	**EF**	**Yes**
AD	**BE**	**CF**	**Yes**
AD	BEC	F	No
ABC	D	EF	No
ABC	DE	F	No
ABD	C	EF	No
ABD	**E**	**CF**	**Yes**
ABD	CE	F	No
ABCD	E	F	No
ABDE	C	F	No

D.

Contrast (LR vs IR vs HR)	**Hazard ratio**	**[95% CI]**	**P(holm)**
A vs B vs CDEF	6.3	2.0-19	<0.0001
A vs BC vs DEF	7.6	2.3-25	<0.0001
A vs BD vs CEF	7.4	2.4-23	<0.0001
A vs BCD vs EF	**10.7**	**3.1-37**	**<0.0001**
A vs BDE vs CF	8.2	2.2-30	<0.0001
AB vs CD vs EF	7.2	2.8-19	<0.0001
AB vs DE vs CF	5.4	1.9-16	<0.0001
AD vs B vs CEF	3.8	1.2-12	0.0021
AD vs BC vs EF	5.5	1.6-19	0.0005
AD vs BE vs CF	4.2	1.1-16	0.0021
ABD vs E vs CF	4.0	1.3-12	0.0015

Table S3. Cox regression analysis. Models are shown for PFS and OS, including treatment interaction to detect a differential treatment effect**.** Shown are the two main effects (SKY-RISS and treatment) complemented with their interaction, indicating that there is a significantly different OS between the two treatment arms given the SKY-RISS risk groups.

			PFS			OS		
		n	Hazard ratio	[95%CI]	p	Hazard ratio	[95%CI]	p
SKY-RISS	I	26	1.0			1.0		
	II	124	3.7	[1.3 – 10.3]	<0.0001	4.4	[1.4 – 14.3]	<0.0001
	III	18	13.2	[4.1 – 42.7]		28.7	[7.6 – 109]	
Treatment	MPT-T	87	1.0		0.33	1.0		0.22
	MPR-R	81	2.4	[0.72 – 7.9]		1.7	[0.65 – 9.7]	
Interaction	MPR-R + SKY-RISS I	13	1.0			1.0		
	MPR-R + SKY-RISS II	61	0.34	[0.10 – 1.2]	0.09	0.34	[0.08 – 1.4]	0.02
	MPR-R + SKY-RISS III	7	0.18	[0.04 - 0.88]		0.10	[0.02 - 0.54]	
		n = 168	118 events = 118; 5 df: $p= 1x10^{-5}$			107 events = 118; 5 df: $p=3x10^{-10}$		

Table S4. Differences between data sets relative to the HO87

		CoMMpass			MRC-IX			HO87	
		N	%	p	N	%	p	N	%
Age				<.0001(a)			.24(a)		
	Median	70			74			72	
	Q1	67			70			69	
	Q3	74			77			76	
	Range	65 - 90			62 - 89			60 - 84	
	Number	93			102			180	
Sex				.07(b)			0.71(b)		
	Male	59	63%		59	54%		93	52%
	Female	34	37%		50	46%		87	48%
ISS stage				.04(b)			<.0001(b)		
	ISS I	31	33%		10	10%		44	25%
	ISS II	31	33%		37	39%		87	49%
	ISS III	31	33%		49	51%		46	26%
gain1q							0.41(b)		
	No	-			55	57%		82	62%
	Yes	-			42	43%		50	38%
del(13q)				0.002(b)			0.90(b)		
	No	42	79%		52	53%		85	54%
	Yes	11	21%		46	47%		72	46%
del(17p)				0.0003(b)			1.0(b)		
	No	45	68%		88	90%		136	89%
	Yes	21	32%		10	10%		16	11%
t(4;14)				0.22(b)			0.03(b)		
	No	52	85%		82	82%		146	91%
	Yes	9	15%		18	18%		14	9%
t(11;14)				0.07(b)			0.70(b)		
	No	38	76%		85	85%		117	87%
	Yes	12	24%		15	15%		17	13%

(a) Kruskal-Wallis test; (b) Fisher exact test

Tables S5 and S6 are provided in a separate Excel document.

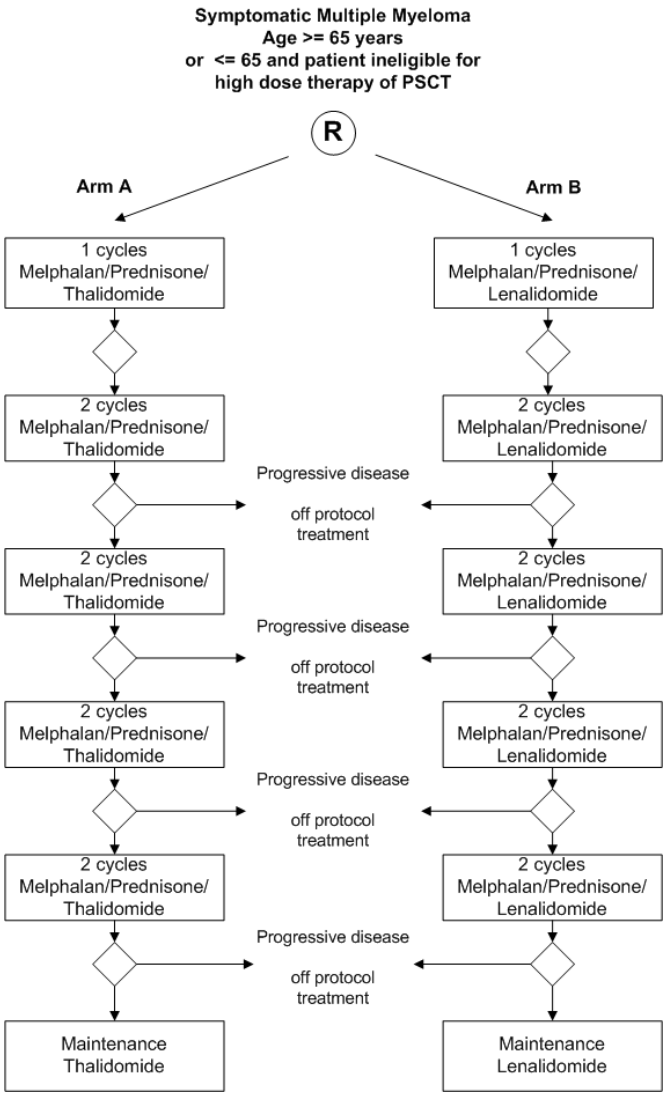


Figure S1: Flow chart for the HOVON-87/NMSG-18 randomized phase III trial in elderly patients with previously untreated symptomatic multiple myeloma comparing Melphalan, Prednisone, Thalidomide followed by Thalidomide maintenance (MPT-T) versus Melphalan, Prednisone, Lenalidomide followed by maintenance with Lenalidomide (MPR-R). This Figure originates from the online available study protocol by HOVON. The following data were adjusted from Zweegman et al (2016)[1]: A total of 668 patients were included and randomized in the study from March 12, 2009 until October 19, 2012, of whom 31 were found not to be eligible. Of the 636 eligible patients, 317 patients were randomly assigned to MPT-T and 319 patients MPR-R. The characteristics at baseline were well balanced. Therapy cycles were given every 4 weeks. The protocol contains 9 cycles of Melphalan 0.18 mg/kg per day for 4 days, prednisone 2 mg/kg per day for 4 days. In arm A, Thalidomide 200 mg was given from day 1 until 4 weeks after the last cycle of MPT, irrespective existing pancytopenia. Maintenance treatment with Thalidomide 100 mg was started 4 weeks after start of the last cycle of MPT. In arm B, Lenalidomide 10 mg was given on day 1-21 followed by a 1 week interval, irrespective of e-xisting pancytopenia at the start of treatment. Maintenance treatment with Lenalidomide was started 4 weeks after the start of the last MPR cycle, at a dose of 10 mg days 1-21. After cycle 1, 3, 5, 7 and 9 evaluation took place and in case of progressive disease after cycle 3, 5, 7 or 9, patients were taken off study.

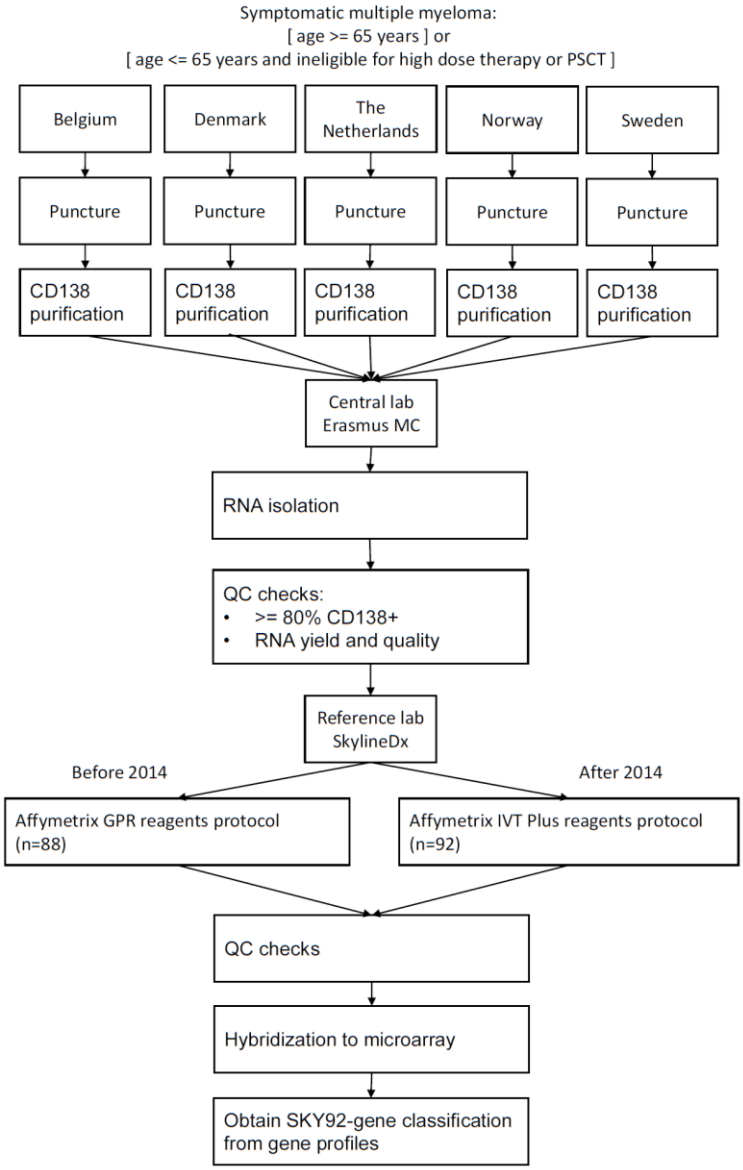


Figure S2: Sample collection and workup flow diagram for the HOVON-87/NMSG-18 MMprofiler™ samples. Patients included came from biobanked samples from the Erasmus MC/HOVON (Belgium and the Netherlands) and Nordic (Norway, Sweden and Denmark). Out of 636 trial patients in total, 537 had informed consent at the time of analysis. Of these, 135 were within the Nordic biobank area, and 402 within the biobank of the Erasmus MC/HOVON.

Of the 135 Nordic samples, 34 were suitable for gene expression profiling. For the remaining 101 samples, bone marrow sample was not taken or quality control criteria were not met (e.g. purity not reached, poor RNA quality and incomplete data).

Of the 402 samples within the Erasmus MC/HOVON biobanking area, 146 high quality samples were used for gene expression profiling. Of the remainder, 122 bone marrow aspirates were received at the Erasmus MC biobank. For the remaining 256 samples, no bone marrow was sampled or received (n=134), the sample did not reach sufficient purity (n=100) or other failed additional quality control (n=22), including RNA quality and incomplete data.

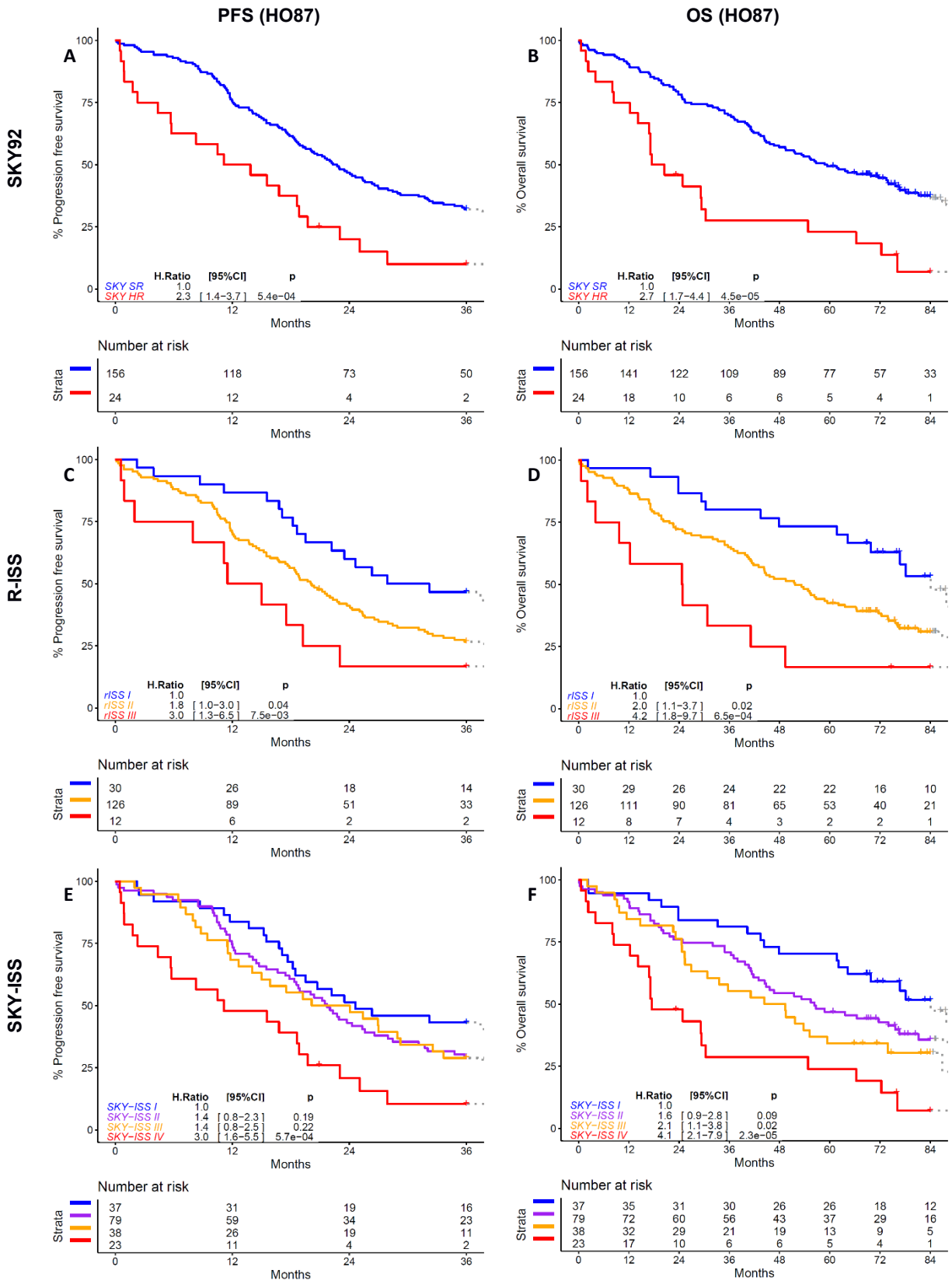


Figure S3. Survival in the HO87. SKY92 PFS and OS (**A, B**), R-ISS PFS and OS (**C, D**), SKY-ISS PFS and OS (**E, F**)

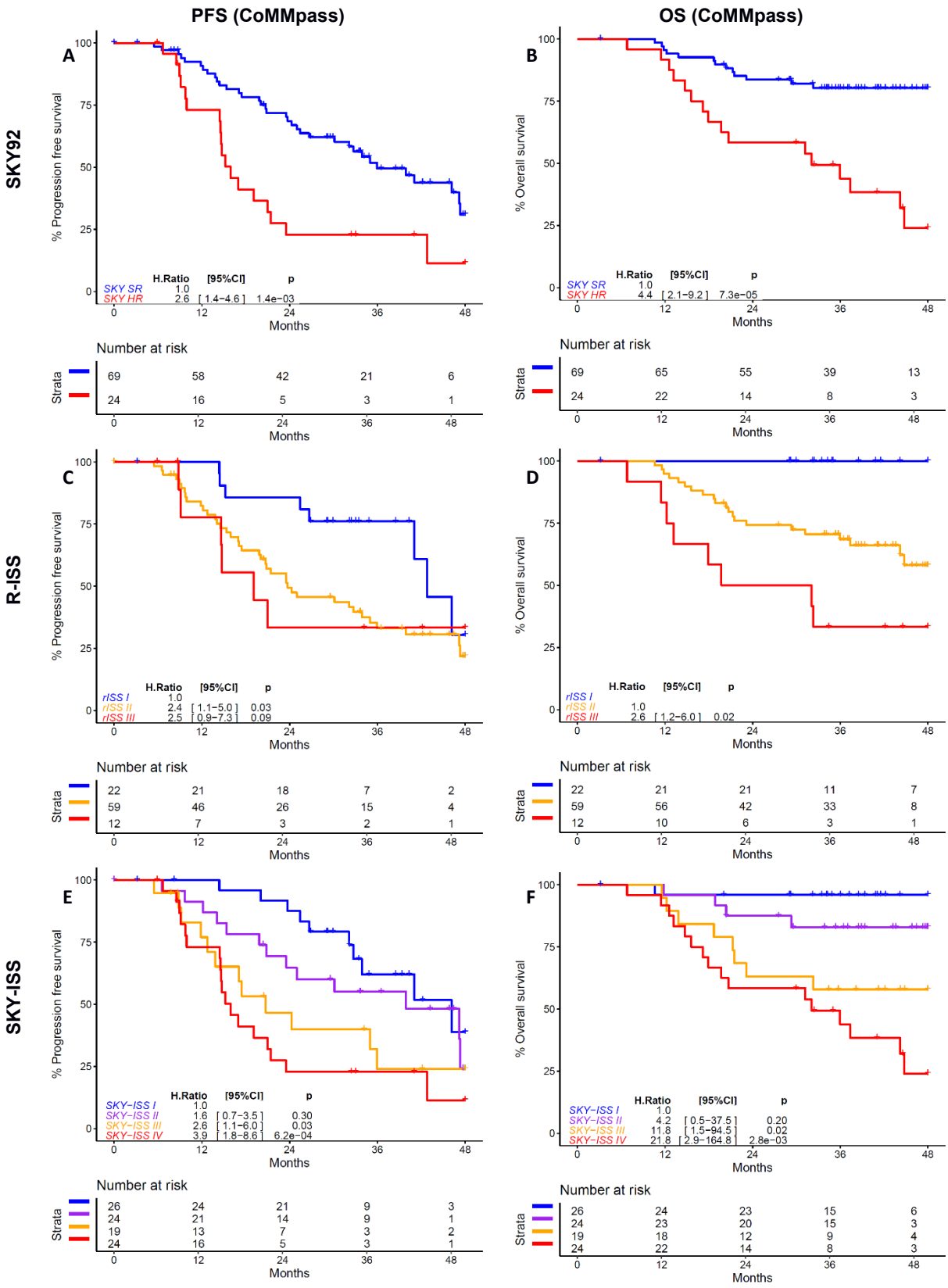
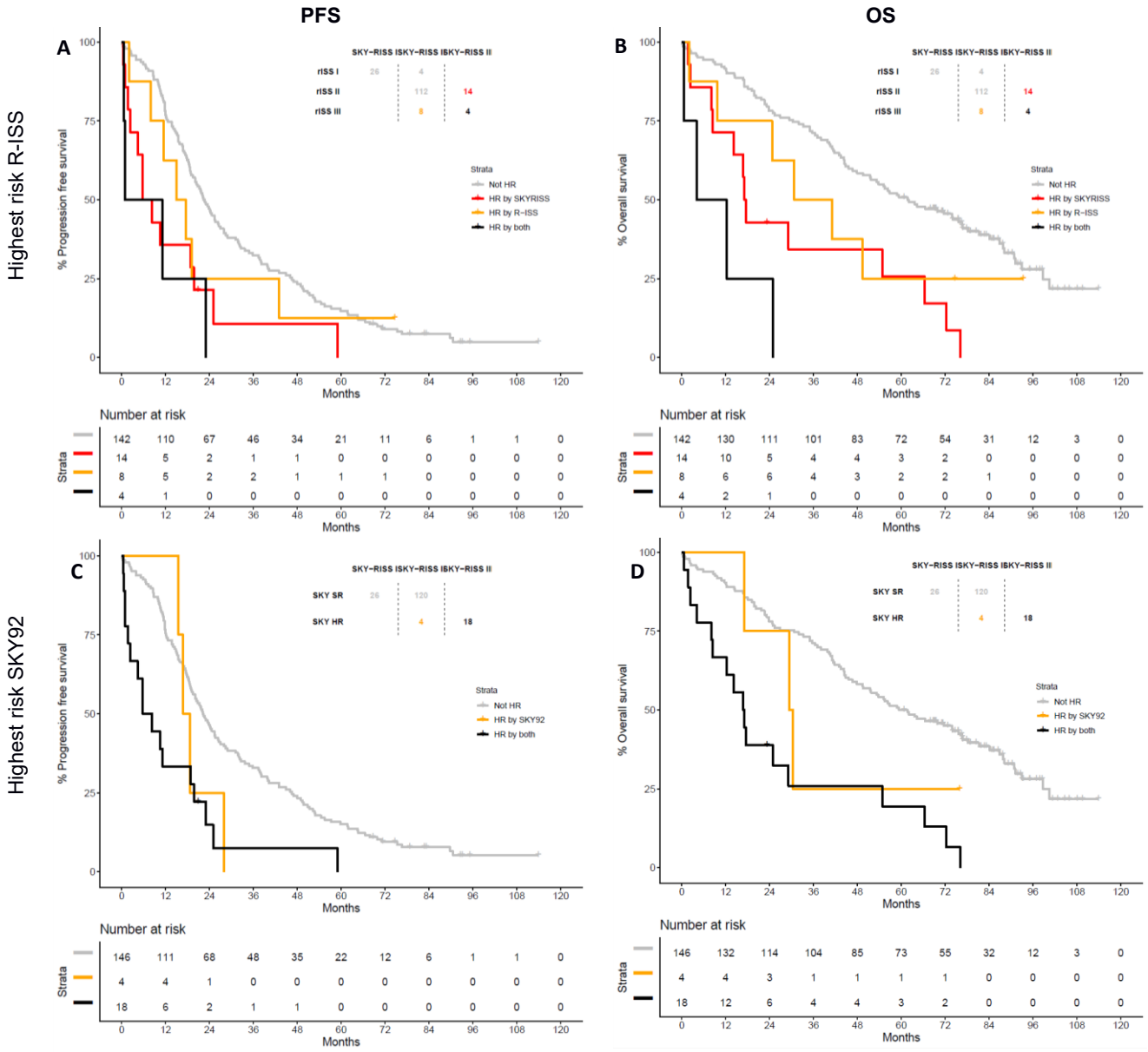

PFS (CoMMpass)
OS (CoMMpass)
SKY92
R-ISS
SKY-ISS
A
B
C
D
E
F
% Progression free survival
% Overall survival
Months
Number at risk
Strata
H.Ratio
[95%CI]
p
SKY SR 1.0
SKY HR 2.6 [1.4−4.6] 1.4e−03
SKY HR 4.4 [2.1−9.2] 7.3e−05
69 58 42 21 6
24 16 5 3 1
69 65 55 39 13
24 22 14 8 3
rISS I 1.0
rISS II 2.4 [1.1−5.0] 0.03
rISS III 2.5 [0.9−7.3] 0.09
rISS III 2.6 [1.2−6.0] 0.02
22 21 18 7 2
59 46 26 15 4
12 7 3 2 1
22 21 21 11 7
59 56 42 33 8
12 10 6 3 1
SKY−ISS I 1.0
SKY−ISS II 1.6 [0.7−3.5] 0.30
SKY−ISS III 2.6 [1.1−6.0] 0.03
SKY−ISS IV 3.9 [1.8−8.6] 6.2e−04
SKY−ISS II 4.2 [0.5−37.5] 0.20
SKY−ISS III 11.8 [1.5−94.5] 0.02
SKY−ISS IV 21.8 [2.9−164.8] 2.8e−03
26 24 21 9 3
24 21 14 9 1
19 13 7 3 2
24 16 5 3 1
26 24 23 15 6
24 23 20 15 3
19 18 12 9 4
24 22 14 8 3

Figure S4. Survival in the CoMMpass. SKY92 PFS and OS (**A, B**), R-ISS PFS and OS (**C, D**), SKY-ISS PFS and OS (**E, F**)



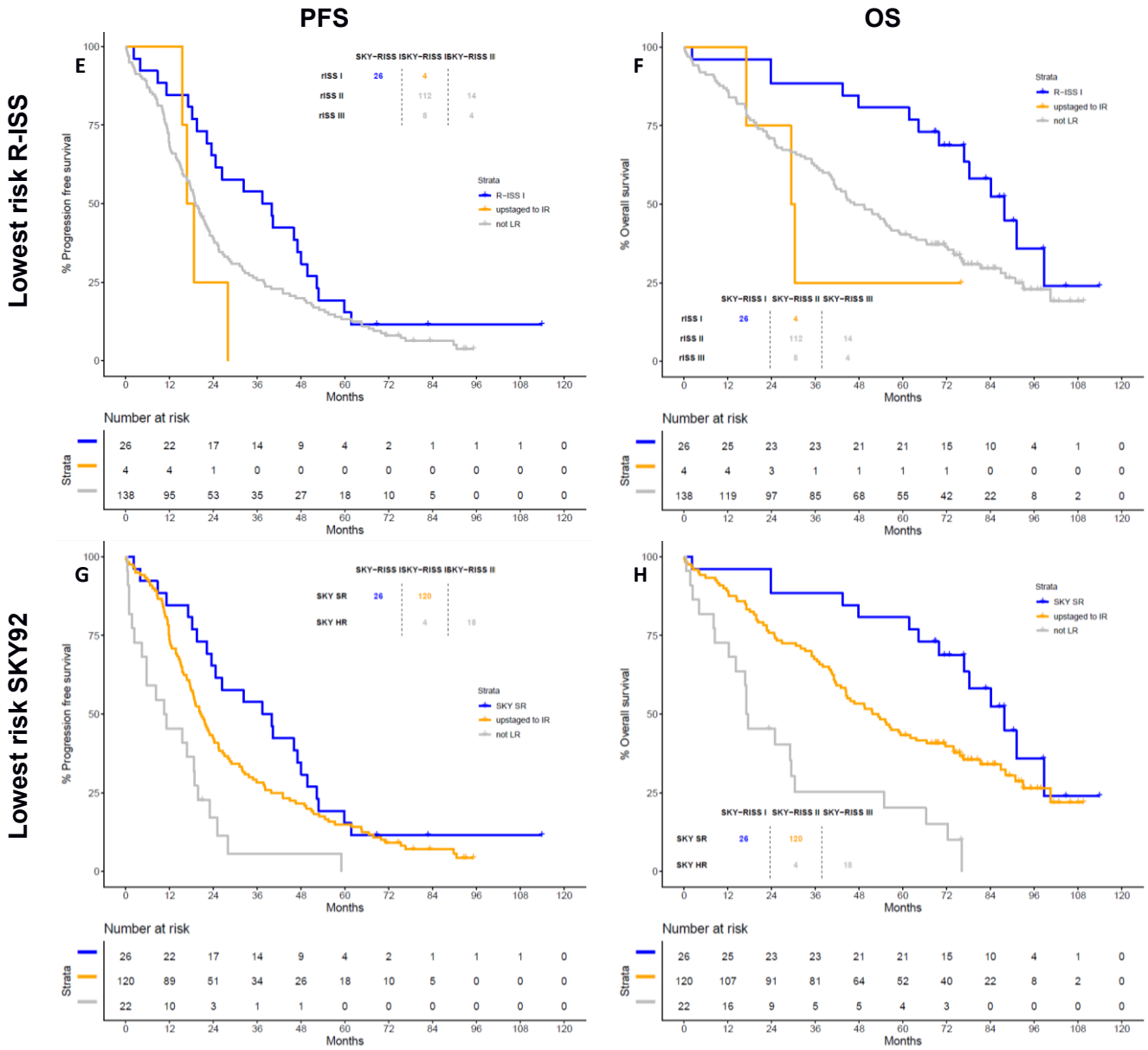


Figure S5. Benefit in terms of PFS (left) and OS (right) in the discovery data (HO87). Illustrated is the better correlation with survival relative to the SKY92 and R-ISS on their own. Highest-risk R-ISS (**A, B**). Highest-risk SKY92 (**C**, **D**). Lowest-risk R-ISS (**E**, **F**). Lowest-risk SKY92 (**G**, **H**).

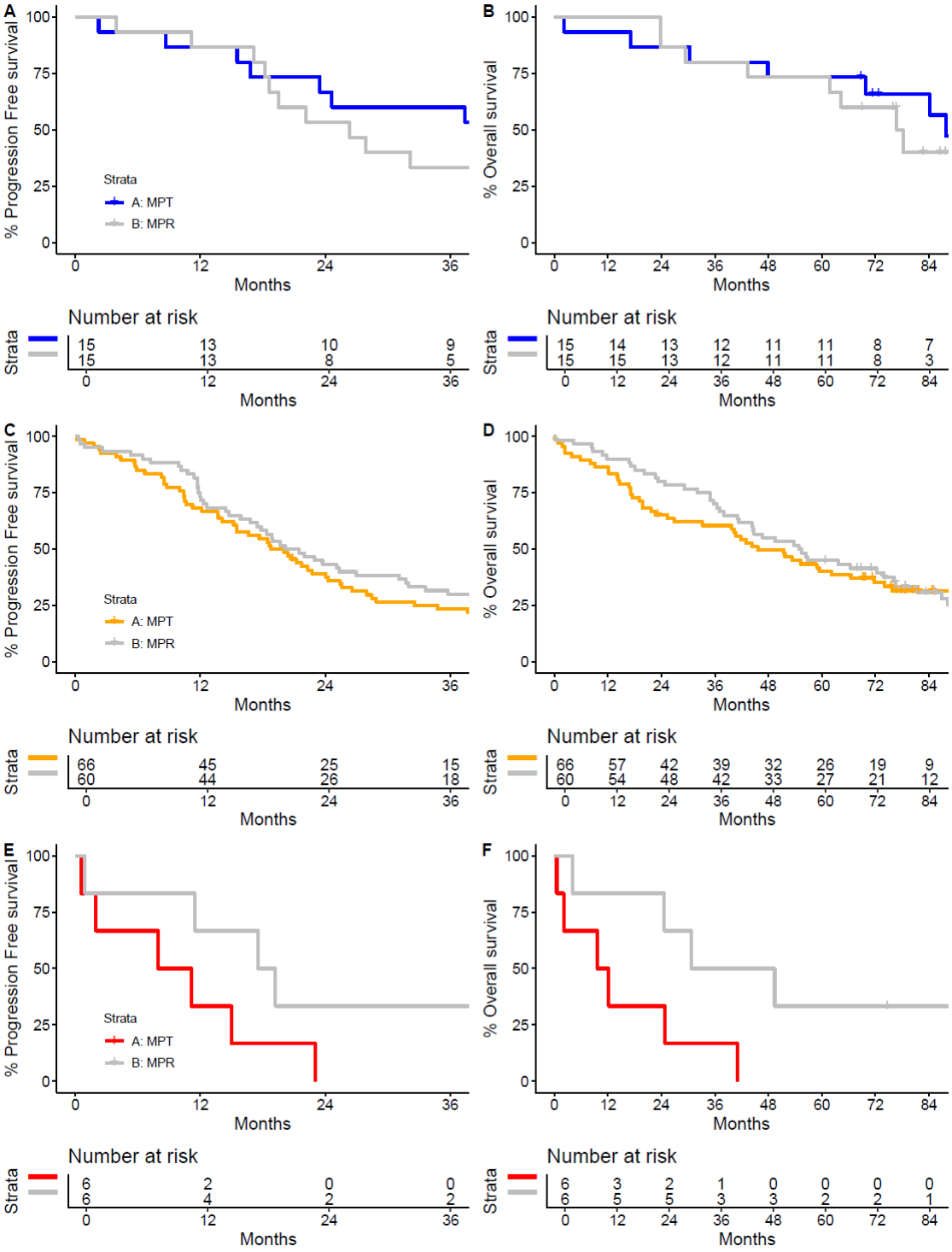


Figure S6. Survival differences in treatment arms per R-ISS risk group. Shown are the PFS (**A, C,** and **E**) and OS (**B, D,** and **F**) per R-ISS risk group: RISS I (**A** and **B**), RISS II (**C** and **D**) and RISS III (**E** and **F**)**.** The colored lines indicate the MPT-T survival, the gray lines indicate the MPR-R survival

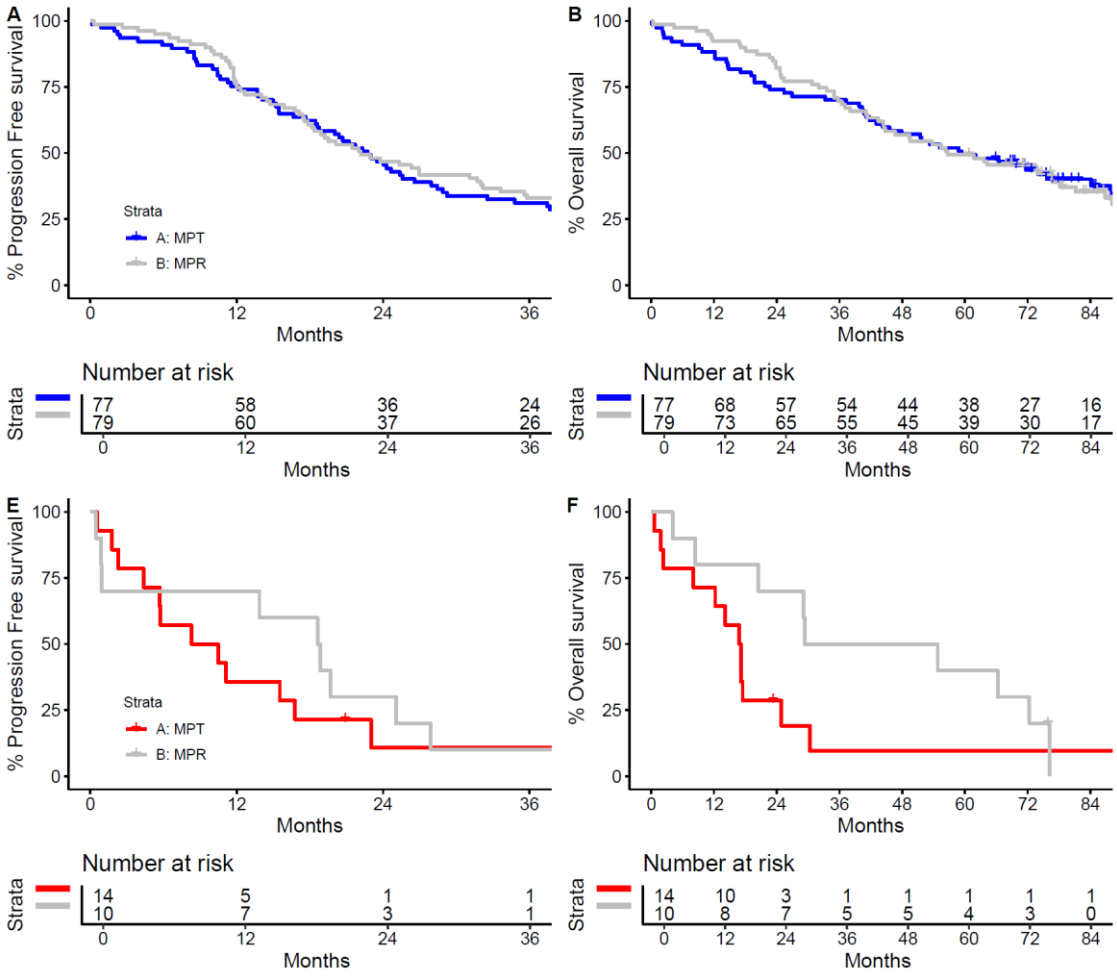


Figure S7. Survival differences in treatment arms per SKY92 risk group. Shown are the PFS (**A** and **C**) and OS (**B** and **D**). SKY92 SR (**A** and **B**), SKY92 HR (**C** and **D**). The colored lines indicate the MPT-T survival, the gray lines indicate the MPR-R survival

References

1. Zweegman S, van der Holt B, Mellqvist UH, et al. Melphalan, prednisone, and lenalidomide versus melphalan, prednisone, and thalidomide in untreated multiple myeloma. *Blood.* 2016;127(9):1109-1116.
2. Pepper SD, Saunders EK, Edwards LE, Wilson CL, Miller CJ. The utility of MAS5 expression summary and detection call algorithms. *BMC Bioinformatics.* 2007;8:273.
3. Kuiper R, Broyl A, de Knegt Y, et al. A gene expression signature for high-risk multiple myeloma. *Leukemia.* 2012;26:2406.
4. Dai M, Wang P, Boyd AD, et al. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res.* 2005;33(20):e175.
5. Du P, Kibbe WA, Lin SM. lumi: a pipeline for processing Illumina microarray. *Bioinformatics.* 2008;24(13):1547-1548.
6. Kuiper R, van Duin M, van Vliet MH, et al. Prediction of high- and low-risk multiple myeloma based on gene expression and the International Staging System. *Blood.* 2015;126(17):1996-2004