

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Development and Validation of a Simplified Score to Predict Early Relapse in Newly Diagnosed Multiple Myeloma in a Pooled Dataset of 2,190 Patients

This is the author's manuscript

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1791004> since 2022-06-14T10:52:15Z

Published version:

DOI:10.1158/1078-0432.CCR-21-0134

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

This is the author's final version of the contribution published as:

Zaccaria GM, Bertamini L, Petrucci MT, Offidani M, Corradini P, Capra A, Romano A, Liberati AM, Mannina D, de Fabritiis P, Cascavilla N, Ruggeri M, Mina R, Patriarca F, Benevolo G, Belotti A, Gaidano G, Nagler A, Hájek R, Spencer A, Sonneveld P, Musto P, Boccadoro M, Gay F. Development and Validation of a Simplified Score to Predict Early Relapse in Newly Diagnosed Multiple Myeloma in a Pooled Dataset of 2,190 Patients. Clin Cancer Res. 2021 Jul 1;27(13):3695-3703. doi: 10.1158/1078-0432.CCR-21-0134. Epub 2021 Apr 29. PMID: 33926915.

© 2021 American Association for Cancer Research.

The publisher's version is available at:

<https://aacrjournals.org/clincancerres/article-abstract/27/13/3695/671504/Development-and-Validation-of-a-Simplified-Score?redirectedFrom=fulltext>

<https://doi.org/10.1158/1078-0432.ccr-21-0134>

When citing, please refer to the published version.

Link to this full text:

<http://hdl.handle.net/2318/1791004>

This full text was downloaded from iris-AperTO: <https://iris.unito.it/>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56

Research article

Developing and Validating a Simplified Score to Predict Early Relapse in Newly Diagnosed Multiple Myeloma (S-ERMM): Analysis from a Pooled Dataset of 2190 Patients

Running title: S-ERMM: A Simplified Early Relapse in Multiple Myeloma Score

Gian Maria Zaccaria^{1,2*}; Luca Bertamini^{1*}; Maria Teresa Petrucci³; Massimo Offidani⁴; Paolo Corradini⁵; Andrea Capra¹; Alessandra Romano⁶; Anna Marina Liberati⁷; Donato Mannina⁸; Paolo de Fabritiis⁹; Nicola Cascavilla¹⁰; Marina Ruggeri¹; Roberto Mina¹; Francesca Patriarca¹¹; Giulia Benevolo¹²; Angelo Belotti¹³; Gianluca Gaidano¹⁴; Arnon Nagler¹⁵; Roman Hájek^{16,17}; Andrew Spencer¹⁸; Pieter Sonneveld¹⁹; Pellegrino Musto^{20,21}; Mario Boccadoro¹; Francesca Gay^{1**}

1. Myeloma Unit, Division of Hematology, University of Torino, Azienda Ospedaliero-Universitaria Città della Salute e della Scienza di Torino, Torino, Italy
2. Hematology and Cell Therapy Unit, IRCCS Istituto Tumori 'Giovanni Paolo II', Bari, Italy
3. Hematology, Department of Translational and Precision Medicine, Azienda Ospedaliera Policlinico Umberto I, Sapienza University of Rome, Rome, Italy
4. Clinica di Ematologia, AOU Ospedali Riuniti di Ancona, Ancona, Italy
5. Divisione di Ematologia, Fondazione IRCCS Istituto Nazionale dei Tumori di Milano; Università degli Studi di Milano, Milano, Italy
6. Department of General Surgery and Medical-Surgical Specialties, Haematology Section, University of Catania, Catania, Italy.
7. Università degli Studi di Perugia, Azienda Ospedaliera Santa Maria, Terni, Italy
8. Division of Hematology, Azienda Ospedaliera Papardo, Messina, Italy
9. Hematology, St. Eugenio Hospital ASL Roma 2, Tor Vergata University, Rome, Italy
10. Ematologia, Ospedale "Casa Sollievo della Sofferenza" IRCCS, San Giovanni Rotondo, Italy
11. Clinica Ematologica e Unità di Terapie Cellulari, Azienda Sanitaria Universitaria Friuli Centrale (ASUFC), Dipartimento di Area Medica (DAME), Università di Udine, Udine, Italy
12. SC Hematology, AO Città della Salute e della Scienza, Turin, Italy
13. Hematology Division, ASST Spedali Civili Brescia, Brescia, Italy
14. Division of Hematology, Department of Translational Medicine, University of Eastern Piedmont, 28100 Novara, Italy
15. Hematology Division, Chaim Sheba Medical Center, Tel Hashomer, Israel
16. Department of Haematooncology, University Hospital Ostrava, Ostrava, Czech Republic
17. Faculty of Medicine, University of Ostrava, Ostrava, Czech Republic
18. Alfred Health-Monash University, Melbourne, Australia
19. Department of Hematology, Erasmus MC Cancer Institute, Rotterdam, the Netherlands
20. Department of Emergency and Organ Transplantation, "Aldo Moro" University School of Medicine, Bari, Italy
21. Unit of Hematology and Stem Cell Transplantation, AOUC Policlinico, Bari, Italy

**These authors equally contributed to this manuscript and share first authorship.*

****Correspondence to:** Dr. Francesca Gay, Myeloma Unit, Division of Hematology, University of Torino, Azienda Ospedaliero-Universitaria Città della Salute e della Scienza di Torino, via Genova 3 - 10126 Torino, Italy; tel: +39 0116334279; fax: +39 0116334187 (E-mail: fgay@cittadellasalute.to.it).
ORCID ID: 0000-0002-8619-412X

Keywords: multiple myeloma; early relapse; score; prognostic features; clinical trials

57 **Original article**
58 **Text word count:** 3873 words
59 **Abstract word count:** 250 words
60 **Number of references:** 33
61 **Number of tables and figures:** 2 tables and 3 figures
62 **Supplementary data:** 1 Supplementary Appendix and 6 supplementary figures (S1-S6)
63 **Supplementary documents:** TRIPOD Checklist - Prediction Model Development
64
65

66 **Authorship contributions**

67 Conceptualization: GMZ, LB, MB, FG
68 Resources: all authors
69 Data curation: all authors
70 Software: GMZ, AC
71 Formal analysis: GMZ, LB, AC, FG
72 Supervision: MB, FG
73 Funding acquisition: none
74 Validation: all authors
75 Investigation: all authors
76 Visualization: GMZ, LB, AC
77 Methodology: GMZ, LB, FG
78 Writing-original draft: GMZ, LB, AC, RM, FG
79 Project administration: none
80 Writing-review and editing: all authors

81 All the authors approved the final version of the manuscript and agreed to be accountable for all aspects of
82 the work in ensuring that questions related to the accuracy or integrity of any part of the work are
83 appropriately investigated and resolved.
84

85
86 **Competing interests**

87 MTP has received honoraria from and has served on the advisory boards for Celgene, Janssen-Cilag, Amgen,
88 Bristol-Myers Squibb, Takeda, Sanofi, and GSK.
89 MO has received honoraria from and has served on the advisory boards for Amgen, Bristol-Myers Squibb,
90 Celgene, GSK, Janssen, Sanofi, and Takeda.
91 PC has participated as lecturer and/or has served on the advisory boards for AbbVie, ADC Therapeutics,
92 Amgen, Celgene, Daiichi Sankyo, Gilead, Incyte, Janssen, Jazz Pharmaceuticals, Kite, Kyowa Kirin, Novartis,
93 Roche, Sanofi, Servier, and Takeda.
94 AML has received personal fees from Incyte; has received research funding from Novartis, Janssen, AbbVie,
95 Roche, Celgene, Amgen, Bristol-Myers Squibb, Takeda, Incyte, Pfizer, Beigene, Oncopeptides, Verastem,
96 Karyopharm, Archigen, Biopharma, Debiopharm, Morphosys, Fibrogen, and Onconova.
97 RM has received honoraria from Sanofi, Celgene, Takeda, and Janssen; has served on the advisory boards for
98 Sanofi, Takeda, and Janssen; has received consultancy fees from Janssen.
99 FP has served on the advisory boards for Celgene and Janssen.
100 GB has received honoraria from Novartis, Celgene, Amgen, and Takeda.
101 AB has served on the advisory boards for Janssen, Celgene, and Amgen.
102 GG has served on the advisory boards for AbbVie, Janssen, and AstraZeneca; has served on speaker's bureaus
103 for AbbVie and Janssen.
104 RH has had a consultant or advisory relationship with Janssen, Amgen, Celgene, AbbVie, Bristol-Myers Squibb,
105 Novartis, PharmaMar, and Takeda; has received honoraria from Janssen, Amgen, Celgene, Bristol-Myers
106 Squibb, PharmaMar, and Takeda; has received research funding from Janssen, Amgen, Celgene, Bristol-Myers
107 Squibb, Novartis, and Takeda.
108 AS has received honoraria from Celgene, Janssen, Amgen, Takeda, Secura Bio, Specialised Therapeutics
109 Australia, AbbVie, Servier, Haemalogix, Sanofi, Roche, Pfizer, and Amgen; has received consultancy fees from
110 Celgene, Janssen, Secura Bio, Specialised Therapeutics Australia, AbbVie, Servier, Haemalogix, Sanofi, Roche,
111 Pfizer, and Amgen; has served on the speakers' bureaus for Celgene, Janssen, Takeda, and Amgen; has
112 received grant/research support from Celgene, Janssen, Amgen, Takeda, Servier, and Haemalogix.
113 PS has served on the advisory boards for Amgen, Celgene, Genenta, Janssen, Seattle Genetics, Takeda, and
114 Karyopharm.
115 PM has received honoraria from and/or served on the scientific advisory boards for Celgene, Janssen, Takeda,
116 Bristol-Myers Squibb, Amgen, Novartis, Gilead, Jazz, Sanofi, AbbVie, and GlaxoSmithKline.

117 MB has received honoraria from Sanofi, Celgene, Amgen, Janssen, Novartis, Bristol-Myers Squibb, and AbbVie;
118 has served on the advisory boards for Janssen and GSK; has received research funding from Sanofi, Celgene,
119 Amgen, Janssen, Novartis, Bristol-Myers Squibb, and Mundipharma.
120 FG has received honoraria from Amgen, Celgene, Janssen, Takeda, Bristol-Myers Squibb, AbbVie, and GSK; has
121 served on the advisory boards for Amgen, Celgene, Janssen, Takeda, Bristol-Myers Squibb, AbbVie, GSK,
122 Roche, Adaptive Biotechnologies, Oncopeptides, Bluebird, and Secura Bio.
123 The remaining authors declare no competing financial interests.
124

125 **Data Sharing**

126 After the publication of this article, data collected for this analysis and related documents will be made
127 available to others upon reasonably justified request, which has to be written and addressed to the attention
128 of the corresponding author Dr. Francesca Gay at the following e-mail address: fgay[at]cittadellasalute.to.it.
129 The corresponding author Dr. Francesca Gay is responsible to evaluate and eventually accept or refuse every
130 request to disclose data and their related documents, in compliance with the ethical approval conditions, in
131 compliance with applicable laws and regulations, and in conformance with the agreements in place with the
132 involved subjects, the participating institutions, and all the other parties directly or indirectly involved in the
133 participation, conduct, development, management and evaluation of this analysis.
134

135 **Financial Support**

136 No funding was provided for this contribution.
137

138 **Acknowledgements**

139 GMZ acknowledges the support of European Myeloma Network (EMN) Research Italy. The authors wish to
140 thank all the study participants and referring clinicians for their valuable contributions.

141 **Statement of translational relevance**

142

143 Despite the huge amount of literature, there is a lack of consensus on how to better predict
144 early relapse (ER) in patients with multiple myeloma (MM). We pooled data from 7
145 European clinical trials enrolling 2190 patients with newly diagnosed MM from October
146 2003 to March 2017 to develop the Simplified Early Relapse in Multiple Myeloma (S-
147 ERMM) score. This analysis provided further evidence of the critical role of predicting ER in
148 MM patients, which is strongly associated with poor outcome. The S-ERMM predicted ER
149 by using simple and widely available baseline features. After external validation, the future
150 development of this prognostic index may consider its combination with other static-risk
151 features (genomic abnormalities, circulating tumor cells) and dynamic risk evaluation
152 (response to the therapy such as minimal residual disease) for ER prediction. The
153 identification of high-risk patients with dismal prognosis is the first step towards a better
154 design of therapeutic approaches for this patient subgroup.

155

156

157 **Abstract**

158

159 **Background.** Despite the improvement of therapeutic regimens, several multiple myeloma
160 (MM) patients still experience early relapse (ER). This subset of patients currently
161 represents an unmet medical need.

162 **Methods.** We pooled data from 7 European multicenter phase II/III clinical trials enrolling
163 2190 newly diagnosed (ND)MM patients from 2003 to 2017. Baseline patient evaluation
164 included 14 clinically relevant features. Patients with complete data (n=1218) were split
165 into training (n=844) and validation sets (n=374). In the training set, a univariate (UV)
166 analysis and a multivariate (MV) logistic regression model on ER within 18 months (ER18)
167 were made. The most accurate model was selected on the validation set. We also developed
168 a dynamic version of the score by including response to treatment.

169 **Results.** The Simplified Early Relapse in MM (S-ERMM) score was modeled on 6 features
170 weighted by a score: 5 points for high lactate dehydrogenase or t(4;14); 3 for del17p,
171 abnormal albumin or bone marrow plasma cells >60%; and 2 for λ free-light chain. The S-
172 ERMM identified 3 patient groups with different risks of ER18: Intermediate (Int) vs. Low
173 (OR=2.39, p<0.001) and High vs. Low (OR=5.59, p<0.001). S-ERMM High/Int patients had
174 significantly shorter OS (High vs. Low: HR=3.24, p<0.001; Int vs. Low: HR=1.86, p<0.001)
175 and PFS2 (High vs. Low: HR=2.89, p<0.001; Int vs. Low: HR=1.76, p<0.001) than S-ERMM
176 Low. The Dynamic (D)S-ERMM modulated the prognostic power of the S-ERMM.

177 **Conclusion.** Based on simple, widely available baseline features, the S-ERMM and DS-
178 ERMM properly identified patients with different risks of ER and survival outcomes.

179

180

181 **1. Introduction**

182 In the past few years, the prognosis of patients with multiple myeloma (MM) has been
183 markedly improved by the introduction of new drugs and better therapeutic strategies
184 both at diagnosis and at relapse (1–4). Traditionally, the maximal benefit in terms of
185 duration of remission has been observed with first-line therapies. With the use of high-
186 dose chemotherapy and autologous stem-cell transplantation (ASCT) combined with novel
187 agents or the adoption of multi-targeted agents including immunomodulatory (IMiD)
188 agents, proteasome inhibitors (PIs) and monoclonal antibodies, the current median
189 progression-free survival (PFS) of newly-diagnosed (ND)MM patients ranged between 41
190 and 50 months (4,5). Despite this remarkable improvement, still a significant proportion of
191 patients experiences an early relapse (ER), which has been associated with a dismal
192 prognosis. Several studies reported the association of baseline clinical features with ER (6–
193 12), but there is no clear consensus on what are the most important determinants; as a
194 matter of fact, even patients without well-known high-risk features at baseline may relapse
195 early (7,13). Data published so far mainly come from registries or from retrospective
196 analyses that do not systematically consider updated standard-of-care risk assessments
197 (e.g. cytogenetics, Revised-International Staging System [R-ISS]) (12).

198 Unfortunately, a consensus on the appropriate definition of ER is also lacking. So far, it has
199 been defined as relapse within either 12/18 months from the start of induction treatment
200 (10,12) or 12/24 months from transplantation (6–9,11,14,15).

201 Indeed, patients with ER have an inferior prognosis, as compared to patients who relapse
202 later and, as such, represent a high-risk group and an unmet medical need (12,14).

203 The correct identification of patient risk at baseline is the first step toward a risk-adaptive
204 therapeutic approach. The aim of our analysis was to develop and validate the Simplified
205 Early Relapse in Multiple Myeloma (S-ERMM), a score to predict the risk of ER based on
206 widely available clinical and biological features. Thereafter, the S-ERMM score was re-
207 modulated during the patient clinical course by integrating response to therapy. We also
208 aimed to correlate the S-ERMM with the long-term outcomes overall survival (OS) and
209 PFS2.
210

211 **2 Methods**

212

213 **2.1 Source of data and participants**

214 Individual patient data from 2190 NDMM patients enrolled in seven multicenter European,
215 open-label, phase II/III clinical trials evaluating novel agent-based therapies from October
216 2003 to March 2017 were pooled together and analyzed: NCT01093196, NCT01346787,
217 NCT01857115, NCT01190787, NCT00551928, NCT01091831, NCT02203643 (2,3,16–20).

218 Each study was approved by ethics committees or institutional review boards at the
219 respective study sites and was conducted in accordance with the Declaration of Helsinki; all
220 patients provided written informed consent. All patients received new drugs (IMiD agents
221 and/or PIs) as upfront treatment, with or without transplantation. Trial details, treatment
222 schedules and eligibility criteria are reported in Supplementary Tables S1A-B.
223

224 **2.2 Prognostic factors and outcomes**

225 Data were retrieved from electronic case-report forms (eCRFs). All the available individual
226 baseline features were analyzed. Age, creatinine levels, albumin, β 2-microglobulin (β 2m)
227 and monoclonal plasma cells in the bone marrow (BMPCs) were evaluated as continuous
228 features. According to the International Myeloma Working Group recommendations, the

229 percentage of BMPCs considered was the highest in case of discrepancy between BM biopsy
230 and BM aspirate (21).
231 Free light chain (FLC, λ vs. κ), M-component subtype (IgA vs. others), lactate
232 dehydrogenase (LDH) levels \geq upper limit of normal (ULN), presence vs. absence of
233 plasmacytomas, presence vs. absence of chromosomal abnormalities (CAs) detected by
234 interphase fluorescence *in situ* hybridization [iFISH; del17p, t(4;14), t(14;16), t(11;14)]
235 were evaluated as categorical values. iFISH analysis was centralized in one laboratory (see
236 the Supplementary Methods). High-risk CAs were defined as the presence of del(17p)
237 and/or t(4;14) and/or t(14;16). (22) The cut-offs for del(17p) and IgH translocation were
238 10% and 15%, respectively. Baseline R-ISS stage (II/III vs. I) was also included in the
239 prognostic factor evaluation (23). Patients with complete data were then split into training
240 and validation sets. In the validation set, patients treated with more innovative and
241 effective therapies were included. These patients also had a shorter median follow-up.
242 Based on the available literature, two cut-offs for ER were evaluated: 18 (ER18) and 24
243 (ER24) months from diagnosis. In the ER18 analysis, patients who died for reasons other
244 than progressive disease (PD) or who withdrew consent within 18 months were excluded
245 from the analysis because they were not at risk of progression for the entire first 18
246 months. Patients experiencing PD within 18 months from diagnosis were included in the
247 ER18 population; those not experiencing PD within 18 months were included in the
248 reference population. The reference population was then divided into 2 groups: patients
249 experiencing PD after 18 months from diagnosis at the time of their last follow-up (Late
250 relapse group) and patients who were free from progression at the time of their last follow-
251 up (No PD group).
252 Methods for the ER24 analysis were similar, but they included a cut-off of 24 months after
253 diagnosis (see the Supplementary Appendix). The results regarding the best cut-off are
254 reported in the main text of this contribution. For the sake of completeness, the other
255 analyses are included in the Supplementary Appendix.
256 OS was calculated from the start of treatment until the date of death or the date the patient
257 was last known to be alive. PFS2 was calculated from the start of treatment until the date of
258 PD after the second line of treatment (second PD) or death (regardless of the cause of
259 death), whichever came first. Other clinical endpoints are detailed in the Supplementary
260 Methods.

261

262 **2.3 Statistical analysis**

263 From the training set, a univariate analysis (UV) on ER18 as outcome was performed
264 according to chi-square and Kruskal-Wallis tests, as appropriate. Features with $p < 0.1$ were
265 then tested in a multivariate (MV) logistic regression model. We compared 2 MV analyses,
266 one including the R-ISS and the other including individual features defining the R-ISS (LDH,
267 albumin, $\beta 2m$ and CAs). In order to account for potential confounders, each MV analysis
268 was adjusted for age. Subsequently, each MV analysis was identified through a backward
269 selection based on the minimization of the Akaike Information Criterion to identify
270 independent prognostic factors. Continuous parameters were not categorized a priori
271 because this would have negatively affected the power of the analysis. After selecting the
272 best MV model, the optimal cut-offs for the most significant continuous features were re-
273 evaluated by spline function. MV models were used to estimate odds ratio (OR) for ER18
274 risk, 95% confidence intervals (CIs) and p-values.

275 Each model was tested on the validation set by assessing the area under the curve (AUC), in
276 order to select the most accurate model including individual features or features
277 aggregated into the R-ISS.

278 Once the most accurate model was selected, three prognostic groups of patients with Low,
279 Intermediate (Int) and High risk of ER were defined by categorizing the linear predictors of
280 the final MV logistic model. Hence, two optimal cut-points were found maximizing the ORs
281 defined by the MV in the training set. A scalar score was thus proportionally assigned to
282 each predictor according to the coefficients of the final MV model. As the linear score, two
283 optimal cut-points were found maximizing the ORs defined by the MV in both the training
284 and validation sets. Thus, we developed the S-ERMM score, which identified three different
285 groups of patients with Low, Int, and High risks of ER18. Other statistical survival analyses
286 are detailed in the Supplementary Methods.

287 In order to integrate baseline prognostic evaluation and response to treatment, we
288 developed the Dynamic (D)S-ERMM, a logistic model that included S-ERMM score and
289 achievement of at least a very good partial response (\geq VGPR). Since this score included
290 response, it should not be assessed at baseline, but at a subsequent timepoint after
291 treatment, in order to re-modulate patient risk during therapy (dynamic risk score). We
292 therefore analyzed data from a landmark point, which was set at the median time to
293 achieve \geq VGPR and included only patients who did not relapse before the landmark point.
294 We assessed the role of \geq VGPR and S-ERMM in a MV logistic regression model to predict
295 ER18. The DS-ERMM was modeled on the proportional coefficients obtained from the MV
296 model. To measure the prognostic performances on this sub-cohort, we compared the
297 concordance (C)-index assessed in both models (24).

298 Statistical analysis was performed using R (v.3.5.2). We used the Transparent Reporting of
299 a Multivariable Prediction Model for Individual Prognosis Or Diagnosis (TRIPOD) criteria
300 to validate our methods (25,26).

301

302 **3 Results**

303

304 **3.1 Patient characteristics**

305 Data from 2190 patients were available; 3 patients were excluded because of screening
306 failure.

307 In the ER18 analysis, 50 patients died for reasons other than PD and 51 withdrew their
308 consent within 18 months and were excluded; patients eligible for the analyses were 2086.
309 Patients with complete data (n=1218) were then split into training (n=844) and validation
310 (n=374) sets and included in the logistic regression analysis.

311 Training set: in the overall population (median follow-up 70 months, interquartile range
312 [IQR]=48-81 months), the median age was 66 years, 73% of patients presented with R-ISS
313 stage II/III, 10% with LDH>ULN; 14% with del(17p), 14% with t(4;14). Patients with
314 BMPCs>60% were 29% and 36% had λ FLC. A total of 312/844 (37%) patients
315 experienced ER18. Patients in the ER18 vs. the reference population were significantly
316 older (p=0.026), had higher β 2m (p<0.001) and lower albumin (p<0.001) levels; a higher
317 proportion of patients had LDH>ULN (p=0.001), t(4;14) (p<0.001), R-ISS stage II/III
318 (p<0.001), del17p (p=0.005) and BMPCs>60% (p=0.001; Table 1).

319 Validation set: in the overall population (median follow-up 35 months, IQR 29-41), the
320 median age was 57 years, which was significantly lower (p<0.001) than that in the training
321 set. Patients who experienced ER18 were 61/374 (16%). The distribution of baseline
322 features between ER18 and the reference population was similar to that in the training set,
323 except for the absence of significant difference in the proportion of patients with
324 BMPCs>60% and del(17p), although this may be related to the smaller sample size (Table
325 1).

326 The median time to \geq VGPR was 9 months in the training set and 3 months in the validation
327 set. \geq VGPR at 9 months was achieved in 40% and 81% of patients in the training and
328 validation sets, respectively.

329
330

331 **3.2 Best model of ER**

332 Based on the UV analysis of patients who experienced ER18, 10/14 features were included
333 in the MV analysis: age, FLC, BMPCs, del17p, t(4;14), t(14;16), albumin, β 2m, LDH, and R-
334 ISS stage. In the MV analysis incorporating the R-ISS, age, R-ISS II/III vs. I and increased
335 BMPCs increased the risk of ER18. When the MV analysis was performed including single
336 features defining the R-ISS, increased BMPCs, λ FLC, LDH>ULN, presence of del17p, and
337 t(4;14) increased the probability of ER18 (Table 2).

338 Each MV model was then tested on the validation set. The AUC was 0.62 (95% CI=0.55-
339 0.69) for the ER18 model including the R-ISS and 0.66 (95% CI=0.58-0.73) for the ER18
340 model incorporating individual features. The ER18 model incorporating individual features
341 resulted in the highest AUC (0.66) and was therefore selected to develop the S-ERMM
342 score.

343 UV and MV ER24 analyses and the AUC in the validation set are reported in the
344 Supplementary Results and in Table S3.

345
346

347 **3.3 S-ERMM score**

348 The ER18 linear index was calculated as $0.047 \times \text{BMPCs } \%/5 + 0.589 \times \text{LDH}/\text{ULN}$ (IF
349 LDH>ULN) + $0.459 \times \text{del17p}$ (IF present) + $0.705 \times \text{t}(4;14)$ (IF present) + $0.293 \times \text{FLC}$ (IF λ)
350 - $0.284 \times \text{albumin}$.

351 In the training set, the linear score significantly discriminated three patient groups (High,
352 intermediate [Int] and Low risk) with significantly different risks of ER18 (Figure 1).

353 BMPC and albumin levels were dichotomized according to the optimal cut-off: high BMPC
354 level if $>60\%$ and abnormal albumin level if ≤ 3.5 or ≥ 5 (Figure S1).

355 The S-ERMM score (<https://sermm.emnitaly.org/>) was mathematically consistent with the
356 linear index and was defined including 6 features identified in the MV analysis: 5 points for
357 LDH>ULN or the presence of t(4;14); 3 points for the presence of del17p, abnormal
358 albumin and BMPCs $>60\%$; and 2 points for the presence of λ FLC (Figure 1).

359 The Low-risk group included patients with a total score ≤ 5 (68% of patients in the training
360 set, 29% of whom experienced an ER18); the Int-risk group patients with a total score
361 between 6 and 10 (25% of patients in the training set, 50% of whom with ER18); and the
362 High-risk group patients with a total score ≥ 11 (7% of patients in the training set, 70% of
363 whom with ER18). In the training set, the S-ERMM significantly discriminated three groups
364 of patients with different risks of ER18: Int vs. Low (OR=2.39, 95% CI=1.73-3.30, $p<0.001$)
365 and High vs. Low (OR=5.59, 95% CI=3.08-10.16, $p<0.001$). The S-ERMM was confirmed in
366 the validation set: Int vs. Low (OR=2.27, 95% CI=1.23-4.17, $p=0.008$) and High vs. Low
367 (OR=4.87, 95% CI=2.01-11.76, $p=0.001$). The impact of the S-ERMM on the ER18 risk was
368 higher than that of each single feature.

369 In the DS-ERMM analyses, the training population (n=673) included patients evaluable for
370 response at 9 months (median time to \geq VGPR), 162 (24%) of whom experienced ER18. In
371 this population, both the S-ERMM and the achievement of \geq VGPR were statistically
372 independent predictors of ER in the MV analysis (Figure S2).

373 The DS-ERMM score was defined as the S-ERMM score obtained at baseline minus 4 points
374 in case of achievement of \geq VGPR. Patients who reached the 9-month cut-off (which was not

375 reached by 171 patients, 150 of whom relapsed/died before), were thus reclassified in
376 three groups: the Low-risk group included 250 patients (37%) with a total score ≤ 0 (only
377 12% of whom experienced an ER18); the Int-risk group included 271 patients (40%) with a
378 total score between 1 and 5 (only 24% of whom with ER18); and the High-risk group
379 included 152 patients (23%) with a total score ≥ 6 (45% of whom with ER18). These three
380 groups had different risks of ER18 (Figure S3): Int vs. Low (OR=2.36, 95% CI=1.46-3.80,
381 $p < 0.001$) and High vs. Low (OR=6.34, 95% CI=3.84-10.46, $p < 0.001$).

382 In the validation set, there were no significant differences in terms of risk of ER18 between
383 DS-ERMM Int vs. Low, while a trend towards a higher risk was observed in the High vs. Low
384 comparison (OR=2.40; $p = 0.09$; see the Supplementary Results).

385 Following the application of the S-ERMM score at baseline and the re-modulation of patient
386 risk at 9 months according to DS-ERMM (for those patients who did not relapse during the
387 first 9 months), 20% of patients in the total population of the training set were classified as
388 High-risk patients, 39% as Int-risk patients, and 41% as Low-risk patients.

389

390 **3.4 Survival analysis**

391 A landmark analysis with landmark point at 18 months was performed. OS and PFS2 were
392 significantly shorter in the ER18 population than in the reference population and the Late
393 relapse and No relapse populations (Figures S4A-D). Similarly, ER18 patients showed an
394 inferior outcome after relapse (Supplementary Results and Figure S5).

395 The median OS was 31.5 months in patients with S-ERMM High, 59.5 with S-ERMM Int and
396 not reached (NR) with S-ERMM Low. Median PFS2 was 19.8 months in patients with S-
397 ERMM High, 40.0 months with S-ERMM Int and 62.3 months with S-ERMM Low. OS and
398 PFS2 were significantly shorter in S-ERMM Int vs. S-ERMM Low patients (OS, HR=1.86,
399 95% CI=1.48-2.33; PFS2, HR=1.76, 95% CI=1.45-2.14; both $p < 0.001$) and in S-ERMM High
400 vs. S-ERMM Int patients (OS, HR=1.74, 95% CI=1.22-2.50, $p = 0.002$; PFS2, HR=1.64, 95%
401 CI=1.18-2.28; $p = 0.003$; Figure 2). The median PFS was 31.6 months in S-ERMM Low, 17.3
402 months in S-ERMM Int, and 13.2 months in S-ERMM High patients.

403 Subgroup analyses for OS according to first-line treatment confirmed the prognostic role of
404 S-ERMM in ASCT-ineligible patients (Int vs. Low, HR=1.75, 95% CI=1.30-2.35, $p < 0.001$;
405 High vs. Int, HR=1.85, 95% CI=1.10-3.11, $p = 0.020$) and in ASCT-eligible patients (Int vs.
406 Low, HR=1.94, 95% CI=1.36-2.78, $p < 0.001$; High vs. Int, HR=1.81, 95% CI=1.09-3.01,
407 $p = 0.022$) (Figure S6); in PI-treated patients (Int vs. Low, HR=1.73, 95% CI=0.93-3.20,
408 $p = 0.083$; High vs. Int, HR=3.13, 95% CI=1.21-8.08, $p = 0.018$); and in patients treated with
409 IMiD agents (Int vs. Low, HR=1.85, 95% CI=1.45-2.37; $p < 0.001$; High vs. Int, HR=1.64, 95%
410 CI=1.11-2.43, $p = 0.013$).

411 According to DS-ERMM, OS and PFS2 were significantly shorter in DS-ERMM Int vs. DS-
412 ERMM Low patients (OS, HR=1.96, 95% CI=1.44-2.66; PFS2, HR=1.86, 95% CI=1.44-2.38;
413 both $p < 0.001$) and in DS-ERMM High vs. DS-ERMM Low patients (OS, HR=3.28, 95%
414 CI=2.37-4.54, $p < 0.001$; PFS2, HR=2.91, 95% CI=2.22-3.82; $p < 0.001$; Figure 3).

415

416 **4 Discussion**

417 Several studies reported dismal survival outcomes in MM patients experiencing an ER;
418 however, the definition of ER varies from study to study, and consensus is still lacking. Also,
419 the impact of well-known disease-related risk factors (e.g., albumin, $\beta 2m$, CAs by iFISH and
420 LDH) on the risk of ER has not been thoroughly assessed in NDMM patients. The correct
421 evaluation of baseline ER risk thus remains an unmet medical need.

422 We confirmed ER with an 18-month cut-off as a strong predictor of the long-term
423 outcomes' OS and PFS2 in the context of novel-treatment approaches (e.g., PIs, IMiD agents
424 and ASCT). Our decision of adopting the 18-month cut-off for the definition of ER was also
425 supported by the available literature, in which most of the studies defined ER as both 18
426 months from diagnosis and 12 months from ASCT (6–10,12,27).

427 We developed the S-ERMM score by identifying and integrating 6 features that predicted
428 the risk of ER (presence of t(4;14), del17p, LDH>ULN, BMPCs>60%, abnormal albumin and
429 λ FLC). S-ERMM is a simple tool enabling the identification of 3 patient groups with
430 significantly different risks of ER (Figure 2) and significantly different PFS2 and OS. In
431 particular, S-ERMM High patients had a median OS of 31 months, significantly shorter than
432 that of S-ERMM Int (median 60 months) and S-ERMM Low patients (median NR after 6
433 years of follow-up).

434 Although several studies tried to correlate ER risk with clinical and biological features (6–
435 9,11,12,14), most of them did not cover all of the recognized MM prognostic factors (23)
436 [e.g. extensive CA analysis (8,9,14) and LDH assessment (10)]. In our series, we included
437 widely available and well-recognized baseline features. The S-ERMM score included
438 albumin levels (which reflect the inflammatory state at diagnosis) (28), high-risk CAs
439 (del17p and t(4;14), associated with a biologically aggressive disease), and high LDH (29)
440 and BMPC levels (associated with tumor burden) (30).

441 The majority of analyses published so far on ER in MM are single-center or retrospective
442 studies. To the best of our knowledge, only Bygrave et al. analyzed young ASCT-eligible
443 patients enrolled in a single clinical trial (7). Indeed, in our analysis, data from clinical trials
444 underwent a systematic data assessment, with baseline features assessment, centralized
445 laboratory analyses and uniform evaluation of response and clinical outcomes (31). In this
446 light, the development and validation of the S-ERMM score in a population consisting of
447 both young (transplant-eligible) and elderly (>65 years) patients enrolled in 4 phase II/III
448 clinical trials treated with novel agents from different drug classes with or without ASCT
449 support the application of this score to NDMM patients. On the other hand, this is a selected
450 population of European clinical trials that indeed needs validation in real-life settings.

451 Response to therapy is a strong predictor of better OS and PFS2 (15), and the achievement
452 of a deep response (minimal residual disease [MRD] negativity) may abrogate the poor
453 prognosis conferred by high-risk FISH at diagnosis. Therefore, the importance of
454 integrating static (baseline) and dynamic (response) prognostic features led to the
455 incorporation of response to treatment (\geq VGPR) into the S-ERMM score. The assessment of
456 the S-ERMM score at the time of diagnosis and the re-modulation of patient risk at 9
457 months (for those patients who did not relapse during the first 9 months) improved our
458 ability to detect ER patients. In fact, in the initial population of our analysis, only 7% of
459 patients were included in the high-risk group (S-ERMM High), while 68% of patients were
460 included in the S-ERMM Low group (29% of whom had an ER18) and 25% in the S-ERMM
461 Int group (50% of whom had an ER18). Of note, in the DS-ERMM analysis, the Low-risk
462 group included only 37% of patients (with only 12% who had an ER18) and the Int-risk
463 group 40% of patients (with only 24% who had an ER18). Ultimately, using sequentially
464 these two scores in the overall population, 20% of patients were determined to be at High
465 risk, 39% at Int risk, and 41% at Low risk. This improvement in the evaluation of patient
466 risk of ER highlighted the role of the dynamic modulation of patient risk at baseline.

467 Unfortunately, in the validation set, there were no significant differences in terms of risk of
468 ER18 between DS-ERMM Int vs. Low, while trends towards a higher risk were observed in
469 the High vs. Low comparison.

470 Of note, the optimal response (degree and timing) to be incorporated as a dynamic factor
471 should consider the type of patient population and the availability of treatment options:

472 these two factors determine the choice of a specific therapy, with different degrees of
473 efficacy and time to best response. In the validation set, the rate of \geq VGPR was definitely
474 higher than in the training set and the median time to response was lower. We presume
475 that the assessment of a deeper response, such as the achievement of MRD negativity, could
476 better discriminate patients in the context of novel, highly effective therapies.
477 Unfortunately, MRD evaluation was not available in most of the trials included in the
478 training set and could not be used as optimal response to recalculate the risk of ER after
479 therapy. Still, our main aim was to identify patients at risk of ER using risk assessment at
480 diagnosis, and the S-ERMM score was prognostic in the context of both older (training set)
481 and more recent (validation set) drug regimens.

482 Our analysis has some limitations. First, the risk classification based on the S-ERMM score
483 was designed to better identify patients at high risk of ER and, as a consequence, was
484 unbalanced, with only a small proportion of patients in the S-ERMM High group.
485 Nevertheless, the risk group stratification improved after the re-modulation of risk
486 assessment by using the DS-ERMM score.

487 Another limitation was the low number of patients treated upfront with a combination of
488 PIs and IMiD agents in the training set, since this currently represents a standard of care
489 for both young and elderly patients. Nevertheless, our results were validated in a
490 population who received intensive and effective induction and consolidation therapies
491 including the second-generation PI carfilzomib with or without IMiD agents and ASCT
492 intensification. In this context, the S-ERMM maintained its prognostic role, but the
493 percentage of patients experiencing ER was definitely lower than that reported in the
494 training set.

495 In conclusion, we were able to correctly classify a good proportion of patients who
496 experienced early relapse by assessing the S-ERMM score at baseline and re-modulating
497 patient risk at 9 months with the DS-ERMM score. An external validation of the S-ERMM
498 and DS-ERMM scores is warranted, especially in patients treated with combinations of PIs,
499 IMiD agents, and anti-CD38 monoclonal antibodies. Our ability to predict ER could also be
500 improved by the inclusion of other risk features at baseline with known prognostic impact,
501 such as amp(1q21), TP53 mutational status, and circulating plasma cells (27,32,33).
502 Unfortunately, these data were not available for this analysis.

503 The correct identification of patient risk at diagnosis and during therapy is an essential
504 step towards a risk-adapted approach, the cure of patients, and the prevention of over- and
505 under-treatment.
506

507 5 References

508

- 509 1. Kumar SK, Rajkumar SV, Dispenzieri A, Lacy MQ, Hayman SR, Buadi FK, et al. Improved
510 survival in multiple myeloma and the impact of novel therapies. *Blood*. 2008;111:2516–
511 20.
- 512 2. Palumbo A, Cavallo F, Gay F, Di Raimondo F, Yehuda DB, Petrucci MT, et al. Autologous
513 transplantation and maintenance therapy in multiple myeloma. *N Engl J Med*.
514 2014;371:895–905.
- 515 3. Gay F, Oliva S, Petrucci MT, Conticello C, Catalano L, Corradini P, et al. Chemotherapy
516 plus lenalidomide versus autologous transplantation, followed by lenalidomide plus
517 prednisone versus lenalidomide maintenance, in patients with multiple myeloma: A
518 randomised, multicentre, phase 3 trial. *Lancet Oncol*. 2015;16:1617–29.
- 519 4. Attal M, Lauwers-cances V, Hulin C, Leleu X, Caillot D, Escoffre M, et al. Lenalidomide,
520 bortezomib, and dexamethasone with transplantation for myeloma. *N Engl J Med*.
521 2017;376:1311–20.
- 522 5. Durie BGM, Hoering A, Sexton R, Abidi MH, Epstein J, Rajkumar SV, et al. Longer term
523 follow-up of the randomized phase III trial SWOG S0777: bortezomib, lenalidomide and
524 dexamethasone vs. lenalidomide and dexamethasone in patients (Pts) with previously
525 untreated multiple myeloma without an intent for immediate autologous stem . *Blood*
526 *Cancer J*. Springer US; 2020;10.
- 527 6. Kumar S, Mahmood ST, Lacy MQ, Dispenzieri A, Hayman SR, Buadi FK, et al. Impact of
528 early relapse after auto-SCT for multiple myeloma. *Bone Marrow Transplant*.
529 2008;42:413–20.
- 530 7. Bygrave C, Pawlyn C, Davies F, Craig Z, Cairns D, Hockaday A, et al. Early relapse after
531 high-dose melphalan autologous stem cell transplant predicts inferior survival and is
532 associated with high disease burden and genetically high-risk disease in multiple
533 myeloma. *Br J Haematol*. 2020;Epub ahead of print.
- 534 8. Jimenez-Zepeda VH, Reece DE, Trudel S, Chen C, Tiedemann R, Kukreti V. Early relapse
535 after single auto-SCT for multiple myeloma is a major predictor of survival in the era of
536 novel agents. *Bone Marrow Transplant*. 2015;50:204–8.
- 537 9. Ong SY, De Mel S, Chen YX, Ooi MG, Surendran S, Lin A, et al. Early relapse post
538 autologous transplant is a stronger predictor of survival compared with pretreatment
539 patient factors in the novel agent era: Analysis of the Singapore Multiple Myeloma
540 Working Group. *Bone Marrow Transplant*. 2016;51:933–7.
- 541 10. Corre J, Montes L, Martin E, Perrot A, Caillot D, Leleu X, et al. Early relapse after
542 autologous transplant for myeloma is associated with poor survival regardless of
543 cytogenetic risk. *Haematologica*. 2019;105:E480–3.
- 544 11. Helm-Petersen S, Sørrig R, Klausen TW, Preiss B, Frølund UC, Helleberg C, et al. Early
545 relapsed disease of multiple myeloma following up-front HDM-ASCT: a study based on
546 the Danish Multiple Myeloma Registry in the period 2005 to 2014. *Leukemia*.
547 2018;32:2054–7.
- 548 12. Majithia N, Rajkumar S V, Lacy MQ, Buadi FK, Dispenzieri A, Gertz MA, et al. Early relapse
549 following initial therapy for multiple myeloma predicts poor outcomes in the era of novel
550 agents. *Leukemia*. 2016;30:2208–13.
- 551 13. Durie BGM, Jacobson J, Barlogie B, Crowley J. Magnitude of response with myeloma
552 frontline therapy does not predict outcome: Importance of time to progression in
553 Southwest Oncology Group chemotherapy trials. *J Clin Oncol*. 2004;22:1857–63.
- 554 14. Kumar SK, Dispenzieri A, Fraser R, Mingwei F, Akpek G, Cornell R, et al. Early relapse after

- 555 autologous hematopoietic cell transplantation remains a poor prognostic factor in
556 multiple myeloma but outcomes have improved over time. *Leukemia*. 2018;32:986–95.
- 557 15. Gopalakrishnan S, D’Souza A, Scott E, Fraser R, Davila O, Shah N, et al. Revised
558 International Staging System Is Predictive and Prognostic for Early Relapse (<24 months)
559 after Autologous Transplantation for Newly Diagnosed Multiple Myeloma. *Biol Blood
560 Marrow Transplant*. Elsevier Inc.; 2019;25:683–8.
- 561 16. Magarotto V, Bringhen S, Offidani M, Benevolo G, Patriarca F, Mina R, et al. Triplet vs
562 doublet lenalidomide-containing regimens for the treatment of elderly patients with
563 newly diagnosed multiple myeloma. *Blood*. 2016;127:1102–8.
- 564 17. Bringhen S, Petrucci MT, Larocca A, Conticello C, Rossi D, Magarotto V, et al. Carfilzomib,
565 cyclophosphamide, and dexamethasone in patients with newly diagnosed multiple
566 myeloma: A multicenter, phase 2 study. *Blood*. 2014;124:63–9.
- 567 18. Larocca A, Bringhen S, Petrucci MT, Oliva S, Falcone AP, Caravita T, et al. A phase 2 study
568 of three low-dose intensity subcutaneous bortezomib regimens in elderly frail patients
569 with untreated multiple myeloma. *Leukemia*. 2016;30:1320–6.
- 570 19. Palumbo A, Bringhen S, Rossi D, Cavalli M, Larocca A, Ria R, et al. Bortezomib-Melphalan-
571 Prednisone-Thalidomide Followed by Maintenance With Bortezomib-Thalidomide
572 Compared With Bortezomib-Melphalan-Prednisone for Initial Treatment of Multiple
573 Myeloma: A Randomized Controlled Trial. *J Clin Oncol*. 2010;28:5101–9.
- 574 20. Gay F, Bertamini L, Zaccaria GM, Aquino S, Spadano A, Siniscalchi A, et al. PREDICTORS
575 OF EARLY RELAPSE IN YOUNG MULTIPLE MYELOMA PATIENTS: INTEGRATED ANALYSIS OF
576 BASELINE FEATURES AND ACHIEVEMENT OF MINIMAL RESIDUAL DISEASE NEGATIVITY.
577 *HemaSphere*. 2020;4:457 [Abstract #EP999, EHA 2020 25th Congress].
- 578 21. Caers J, Garderet L, Kortüm KM, O’Dwyer ME, van de Donk NWCJ, Binder M, et al.
579 European Myeloma Network recommendations on tools for the diagnosis and
580 monitoring of multiple myeloma: what to use and when. *Haematologica*.
581 2018;103:1772–84.
- 582 22. Sonneveld P, Avet-Loiseau H, Lonial S, Usmani S, Siegel D, Anderson KC, et al. Treatment
583 of multiple myeloma with high-risk cytogenetics: A consensus of the International
584 Myeloma Working Group. *Blood*. 2016;127:2955–62.
- 585 23. Palumbo A, Avet-Loiseau H, Oliva S, Lokhorst HM, Goldschmidt H, Rosinol L, et al. Revised
586 international staging system for multiple myeloma: A report from international myeloma
587 working group. *J Clin Oncol*. 2015;33:2863–9.
- 588 24. Therneau TM, Watson DA. The concordance statistic and the Cox model. Technical
589 Report #85. Rochester (US-MN): Department of Health Sciences, Mayo Clinic; 2017.
- 590 25. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable
591 prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement.
592 *Ann Intern Med*. American College of Physicians; 2015;162:55–63.
- 593 26. Moons KGM, Altman DG, Reitsma JB, Ioannidis JPA, Macaskill P, Steyerberg EW, et al.
594 Transparent reporting of a multivariable prediction model for individual prognosis or
595 diagnosis (TRIPOD): Explanation and elaboration. *Ann Intern Med*. American College of
596 Physicians; 2015;162:W1–73.
- 597 27. D’Agostino M, Zaccaria GM, Ziccheddu B, Rustad EH, Genuardi E, Capra A, et al. Early
598 Relapse Risk in Patients with Newly Diagnosed Multiple Myeloma Characterized by Next-
599 generation Sequencing. *Clin Cancer Res*. *Clin Cancer Res*; 2020;26:4832–41.
- 600 28. Greipp PR, San Miguel J, Durie BGM, Crowley JJ, Barlogie B, Bladé J, et al. International
601 staging system for multiple myeloma. *J Clin Oncol*. 2005;23:3412–20.
- 602 29. Dimopoulos MA, Barlogie B, Smith TL, Alexanian R. High serum lactate dehydrogenase

- 603 level as a marker for drug resistance and short survival in multiple myeloma. *Ann Intern*
604 *Med.* 1991;115:931–5.
- 605 30. Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos M-V, et al.
606 International Myeloma Working Group updated criteria for the diagnosis of multiple
607 myeloma. *Lancet Oncol.* 2014;15:e538-48.
- 608 31. Zaccaria GM, Ferrero S, Rosati S, Ghislieri M, Genuardi E, Evangelista A, et al. Applying
609 data warehousing to a phase III clinical trial from the Fondazione Italiana Linfomi (FIL)
610 ensures superior data quality and improved assessment of clinical outcomes. *JCO Clin*
611 *Cancer Informatics.* 2019;1–15.
- 612 32. Schmidt TM, Barwick BG, Joseph N, Heffner LT, Hofmeister CC, Bernal L, et al. Gain of
613 Chromosome 1q is associated with early progression in multiple myeloma patients
614 treated with lenalidomide, bortezomib, and dexamethasone. *Blood Cancer J. Springer*
615 *US;* 2019;9:Article number 94.
- 616 33. Bertamini L, Grasso M, D’Agostino M, Pascarella A, Tosi P, Monaco F, et al. Poor
617 Prognosis of Multiple Myeloma Predicted By High Levels of Circulating Plasma Cells Is
618 Independent from Other High-Risk Features but Is Modulated By the Achievement of
619 Minimal Residual Disease Negativity. *Blood. American Society of Hematology;*
620 *2020;136:12-13 [Abstract #720, ASH 2020 62nd Meeting].*
621
622

6 Tables

Table 1. Patient characteristics of the overall population, stratified according to the ER18 outcome as training set and validation set

| | POOLED-SET | TRAINING SET* | | | | VALIDATION SET* | | | |
|--|--------------------|--------------------|-------------------|----------------------|--------|--------------------|-------------------|----------------------|--------|
| | Overall population | Overall population | ER18 population | Reference population | P | Overall population | ER18 population | Reference population | P |
| N of patients (%) | 2190 | 844 | 312 (37) | 532 (63) | - | 374 | 61 (16) | 313 (84) | - |
| Age, y Median [IQR] | 63.0 [56.0, 72.0] | 66.0, [57.0-73.0] | 68.0 [58.0, 75.0] | 65.0 [57.0, 73.0] | 0.026 | 57.0 [51.0, 62.0] | 56.0 [48.0, 62.0] | 58.0 [52.0, 62.0] | 0.173 |
| Missing N (%) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | | 0 (0) | 0 (0) | 0 (0) | |
| Albumin, g/dL: Median [IQR] | 3.8 [3.4, 4.2] | 3.8, [3.4-4.2] | 3.7 [3.2, 4.1] | 3.9 [3.5, 4.2] | <0.001 | 3.9 [3.5, 4.3] | 3.7 [3.4, 4.1] | 3.9 [3.5, 4.3] | 0.046 |
| Missing N (%) | 10 (0.5) | 0 (0) | 0 (0) | 0 (0) | | 0 (0) | 0 (0) | 0 (0) | |
| β2m, mg/dL: Median [IQR] | 3.4 [2.4, 5.1] | 3.6, [2.5-5.4] | 4.1 [2.8, 6.1] | 3.3 [2.4, 5.0] | <0.001 | 2.9 [2.0, 4.2] | 3.7 [2.2, 5.8] | 2.8 [2, 4] | 0.015 |
| Missing N (%) | 8 (0.4) | 0 (0) | 0 (0) | 0 (0) | | 0 (0) | 0 (0) | 0 (0) | |
| LDH>ULN, N (%) | 211 (9.6) | 83 (10) | 45 (14) | 38 (7) | 0.001 | 54 (14) | 17 (28) | 37 (12) | 0.002 |
| Missing N (%) | 267 (12.2) | 0 (0) | 0 (0) | 0 (0) | | 0 (0) | 0 (0) | 0 (0) | |
| del17p: N (%) | 246 (11.2) | 121 (14) | 59 (19) | 62 (12) | 0.005 | 52 (14) | 11 (18) | 41 (13) | 0.414 |
| Missing N (%) | 477 (21.8) | 0 (0) | 0 (0) | 0 (0) | | 0 (0) | 0 (0) | 0 (0) | |
| t(4;14): N (%) | 228 (10.4) | 117 (14) | 63 (20) | 54 (10) | <0.001 | 57 (15) | 17 (28) | 40 (13) | 0.005 |
| Missing N (%) | 482 (22.0) | 0 (0) | 0 (0) | 0 (0) | | 0 (0) | 0 (0) | 0 (0) | |
| t(11;14): N (%) | 341 (15.5) | 156 (18) | 50 (16) | 106 (20) | 0.188 | 87 (23) | 14 (23) | 73 (23) | 1 |
| Missing N (%) | 522 (23.8) | 0 (0) | 0 (0) | 0 (0) | | 0 (0) | 0 (0) | 0 (0) | |
| t(14;16): N (%) | 78 (3.6) | 32 (4) | 16 (5) | 16 (3) | 0.171 | 19 (5) | 4 (7) | 15 (5) | 0.798 |
| Missing N (%) | 506 (23.1) | 0 (0) | 0 (0) | 0 (0) | | 0 (0) | 0 (0) | 0 (0) | |
| R-ISS, II/III: N (%) | 1388 (63.5) | 613 (73) | 255 (82) | 358 (67) | <0.001 | 250 (67) | 53 (87) | 197 (63) | <0.001 |
| Missing N (%) | 388 (17.7) | 0 (0) | 0 (0) | 0 (0) | | 0 (0) | 0 (0) | 0 (0) | |

| | | | | | | | | | |
|--|----------------|------------------|----------------|----------------|-------|----------------|----------------|----------------|-------|
| Creatinine, mg/dL: Median [IQR] | 0.9 [0.7, 1.1] | 0.9 [0.8-1.2] | 1.0 [0.8, 1.2] | 0.9 [0.8, 1.1] | 0.136 | 0.8 [0.7, 1.0] | 0.8 [0.7, 1.2] | 0.8 [0.7, 1.0] | 0.427 |
| Missing N (%) | 67 (3.1) | 0 (0) | 0 (0) | 0 (0) | | 0 (0) | 0 (0) | 0 (0) | |
| BMPCs, >60% N (%) | 612 (28) | 243 (29) | 112 (36) | 131 (25) | 0.001 | 139 (37) | 27 (44) | 112 (36) | 0.267 |
| Missing N (%) | 121 (5.5) | 0 (0) | 0 (0) | 0 (0) | | 0 (0) | 0 (0) | 0 (0) | |
| FLC, λ: N (%) | 727 (36) | 301 (36) | 123 (39) | 178 (33) | 0.095 | 143 (38) | 18 (30) | 125 (40) | 0.165 |
| Missing N (%) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | | 0 (0) | 0 (0) | 0 (0) | |
| M component, IgA: N (%) | 451 (21) | 186 (22) | 65 (21) | 121 (23) | 0.575 | 57 (15) | 10 (16) | 47 (15) | 0.937 |
| Missing N (%) | 3 (0.1) | 0 (0) | 0 (0) | 0 (0) | | 0 (0) | 0 (0) | 0 (0) | |
| Plasmacytomas, N (%) | 266 (12) | 78 (9) | 27 (9) | 51 (10) | 0.743 | 49 (13) | 10 (16) | 39 (12) | 0.532 |
| Missing N (%) | 0 (0) | 0 (0) | 0 (0) | 0 (0) | | 0 (0) | 0 (0) | 0 (0) | |

**Patients with complete data only.*

Abbreviations. N, number; y, years; ER, early relapse within 18 months from diagnosis; IQR, interquartile range; β2m, β 2-microglobulin; LDH, lactate dehydrogenase; ULN, upper limit of normal; del17p, deletion 17p; t, translocation; R-ISS, Revised International Staging System; BMPCs, bone marrow plasma cells; FLC, free light chains; M, monoclonal.

Table 2. Univariate (UV) and multivariate (MV) analyses of the baseline features to predict ER18

| | ER18* | | | | | |
|---------------------------------|--------------------------|--------------------|--------|------------------------------------|--------------------|--------|
| | Analysis including R-ISS | | | Analysis including single features | | |
| | UV Analysis | MV Analysis | | UV Analysis | MV Analysis | |
| | p | OR (95% CI) | p | p | OR (95% CI) | p |
| Age, (increased by 1 y) | 0.026 | 1.01 (1.00 - 1.02) | 0.113 | 0.026 | 1.01 (1.00 - 1.03) | 0.094 |
| Albumin, (increased by 1 mg/dL) | | | | <0.001 | 0.75 (0.60 - 0.95) | 0.015 |
| β2m, (increased by 1 mg/dL) | | | | <0.001 | - | - |
| LDH (> vs. ≤ULN) | | | | 0.001 | 2.03 (1.27 - 3.26) | 0.003 |
| del17p (presence vs. no) | | | | 0.005 | 1.65 (1.10 - 2.47) | 0.016 |
| t(4;14) (presence vs. no) | | | | <0.001 | 2.12 (1.40 - 3.19) | <0.001 |
| R-ISS (II/III vs. I) | <0.001 | 1.91 (1.35 - 2.71) | <0.001 | | | |
| BMPCs % (increased by 5%) | <0.001 | 1.05 (1.02 - 1.08) | <0.001 | <0.001 | 1.06 (1.03 - 1.09) | <0.001 |
| FLC (λ vs. κ) | 0.095 | - | - | 0.095 | 1.31 (0.97 - 1.78) | 0.076 |

* Only significant features ($p < 0.1$) in UV analysis and age were included.

- Excluded before the MV analysis by the Akaike information criterion (AIC).

Abbreviations. ER18, early relapse within 18 months from diagnosis, R-ISS Revised International Staging System; UV, univariate; MV, multivariate; OR, odds ratio; CI, confidence interval; β2m, β 2-microglobulin; LDH, lactate dehydrogenase; ULN, upper limit of normal; del17p, deletion 17p; t, translocation; R-ISS, Revised International Staging System; BMPCs, bone marrow plasma cells; FLC, free light chains; M, monoclonal.

7 Figures: titles and legends

Figure 1. Flowchart of the S-ERMM score construction

Abbreviations. ER, early relapse; features, features; BMPCs, bone marrow plasma cells; LDH, lactate dehydrogenase; ULN, upper limit of normal; del17p, deletion 17p; t, translocation; FLC, free light chains; alb, albumin; Low, low; Int, intermediate; High, high; coeff., coefficient; OR, odds ratio; CI, confidence interval; p, p-value; S-ERMM, Simplified Early Relapse in Multiple Myeloma score.

Figure 2. OS (A) and PFS2 (B) stratified by S-ERMM score

A) OS: S-ERMM Int vs. S-ERMM Low, S-ERMM High vs. S-ERMM Low and S-ERMM High vs. S-ERMM Int.

B) PFS2: S-ERMM Int vs. S-ERMM Low, S-ERMM High vs. S-ERMM Low and S-ERMM High vs. S-ERMM Int.

Abbreviations. OS, overall survival; PFS2, progression free survival-2; S-ERMM, Simplified Early Relapse in Multiple Myeloma score; Int, intermediate; HR, hazard ratio; CI, confidence interval; p, p-value.

Figure 3. OS (A) and PFS2 (B) according to the DS-ERMM score

Landmark analysis at 9 months, corresponding to the median time to achieve \geq VGPR.

A) OS: DS-ERMM Int vs. DS-ERMM Low, DS-ERMM High vs. DS-ERMM Low, and DS-ERMM High vs. DS-ERMM Int.

B) PFS2: DS-ERMM Int vs. DS-ERMM Low, DS-ERMM High vs. DS-ERMM Low and DS-ERMM High vs. DS-ERMM Int.

Abbreviations. OS, overall survival; PFS2, progression-free survival-2; DS-ERMM score, Dynamic Simplified Early Relapse in Multiple Myeloma score; VGPR, very good partial response; Int, intermediate; HR, hazard ratio; CI, confidence interval; p, p-value.

Figure 1

Training Set

Model selected for the definition of the score: ER18 involving 6/14 fts: **BMPCs , LDH>ULN, del17p, t(4;14), FLC, albumin**

The ER score was calculated as
 $0.047 \times \text{BMPCs \%}/5 + 0.589 \times \text{LDH (IF >ULN)} + 0.459 \times \text{del17p (IF present)} + 0.705 \times \text{t(4;14) (IF present)} + 0.293 \times \text{FLC (IF FLC}=\lambda) - 0.284 \times \text{albumin}$

Categorization of continuous variables BMPCs (>60%) and albumin (≥ 3.5 and ≤ 5)

Definition of the S-ERMM score:
BMPCs → Beta Coeff. proportionality: 1.7 → score: 3
albumin → Beta Coeff. proportionality: 1.6 → score: 3
del17p → Beta Coeff. proportionality: 1.8 → score: 3
t(4;14) → Beta Coeff. proportionality: 2.8 → score: 5
LDH>ULN → Beta Coeff. proportionality: 2.7 → score: 5
FLC → Beta Coeff. proportionality: 1.0 → score: 2

Int vs. Low → OR=2.51, 95% CI=1.81–3.48, p<0.001
High vs. Low → OR=4.59, 95% CI=2.45–8.61, p<0.001

Int vs. Low → OR=2.39, 95% CI=1.73–3.30, p<0.001
High vs Low → OR=5.59, 95% CI=3.08–10.16, p<0.001

Validation Set

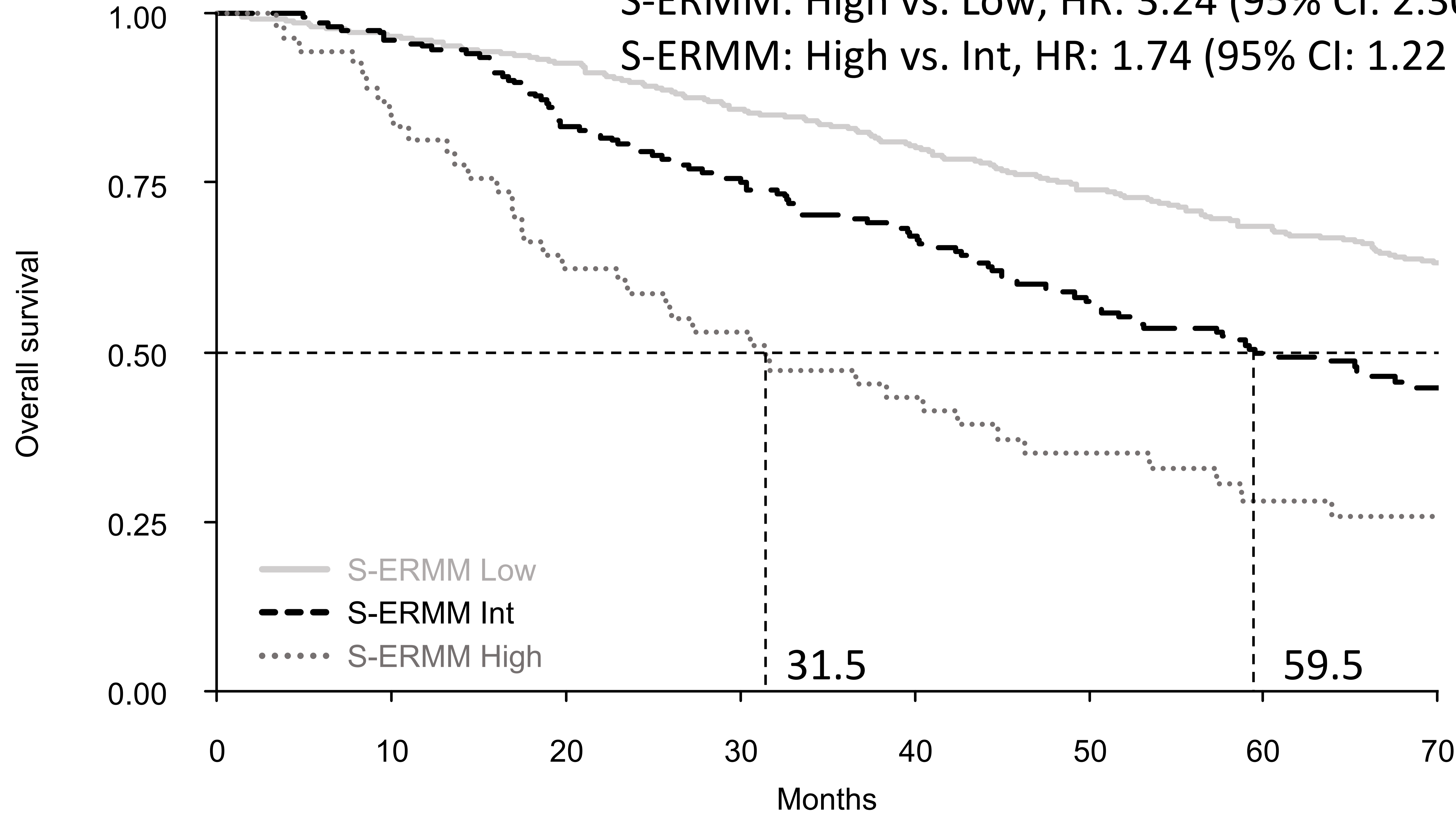
Int. vs. Low → OR=2.52, 95% CI=1.36–4.68, p=0.003
High vs. Low → OR=5.55, 95% CI=2.58–14.22, p<0.001

Int vs Low → OR=2.27, 95% CI=1.23–4.17, p=0.008
High vs Low → OR=4.87, 95% CI=2.01–11.76, p=0.001

Figure 2

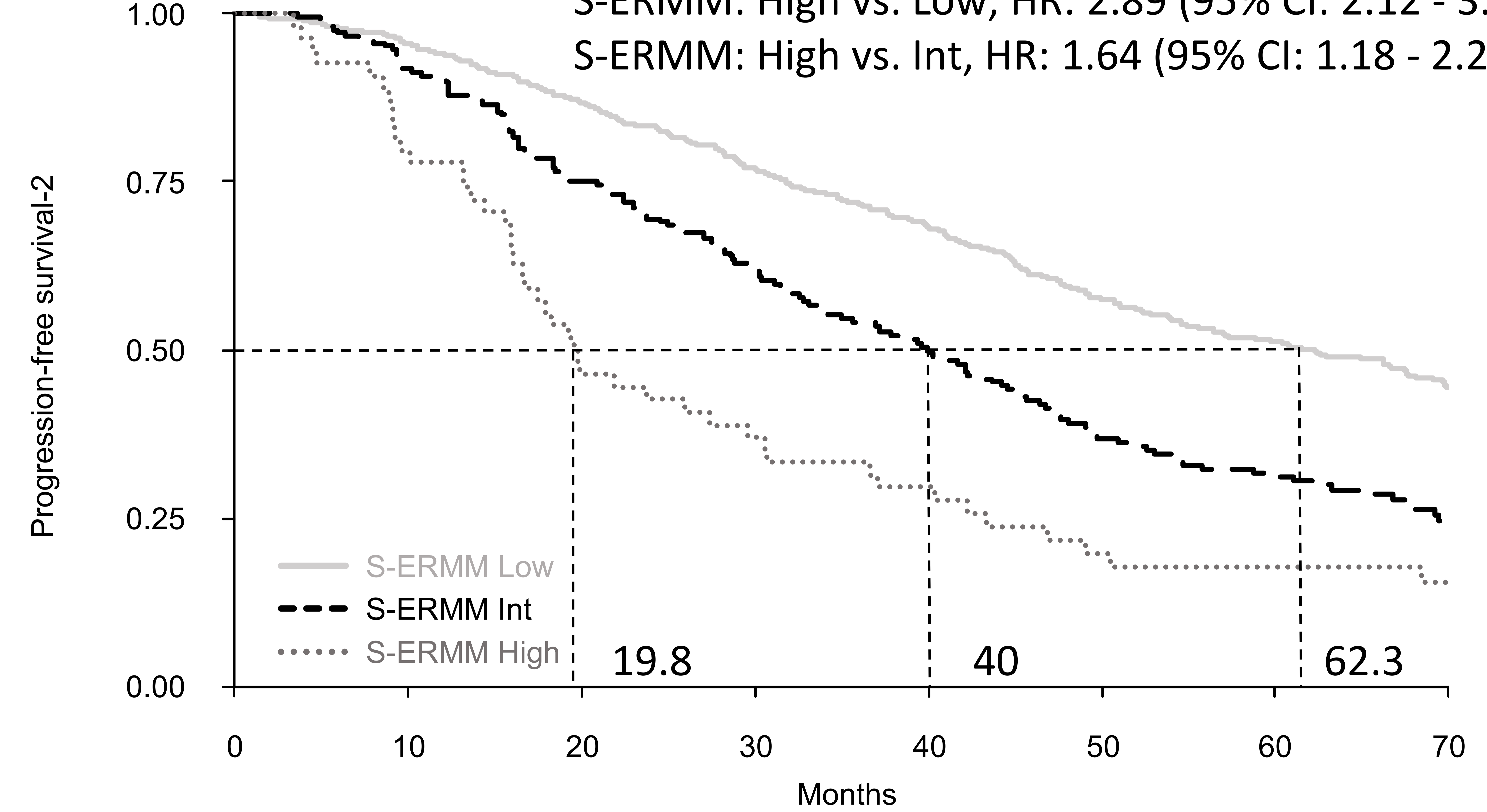
A. OS

S-ERMM: Int vs. Low, HR: 1.86 (95% CI: 1.48 - 2.33, p<0.001)
 S-ERMM: High vs. Low, HR: 3.24 (95% CI: 2.30 - 4.54, p<0.001)
 S-ERMM: High vs. Int, HR: 1.74 (95% CI: 1.22 - 2.50, p=0.002)



B. PFS2

S-ERMM: Int vs. Low, HR: 1.76 (95% CI: 1.45 - 2.14, p<0.001)
 S-ERMM: High vs. Low, HR: 2.89 (95% CI: 2.12 - 3.93, p<0.001)
 S-ERMM: High vs. Int, HR: 1.64 (95% CI: 1.18 - 2.28, p=0.003)



| | | | | | | | | |
|-------------|---------|----------|----------|----------|----------|----------|-----------|-----------|
| S-ERMM Low | 574 (0) | 539 (16) | 501 (33) | 448 (49) | 406 (63) | 351 (87) | 290 (123) | 215 (178) |
| S-ERMM Int | 214 (0) | 198 (8) | 166 (14) | 147 (18) | 124 (25) | 99 (33) | 81 (38) | 53 (59) |
| S-ERMM High | 56 (0) | 45 (3) | 33 (3) | 28 (3) | 22 (4) | 16 (6) | 12 (7) | 11 (7) |

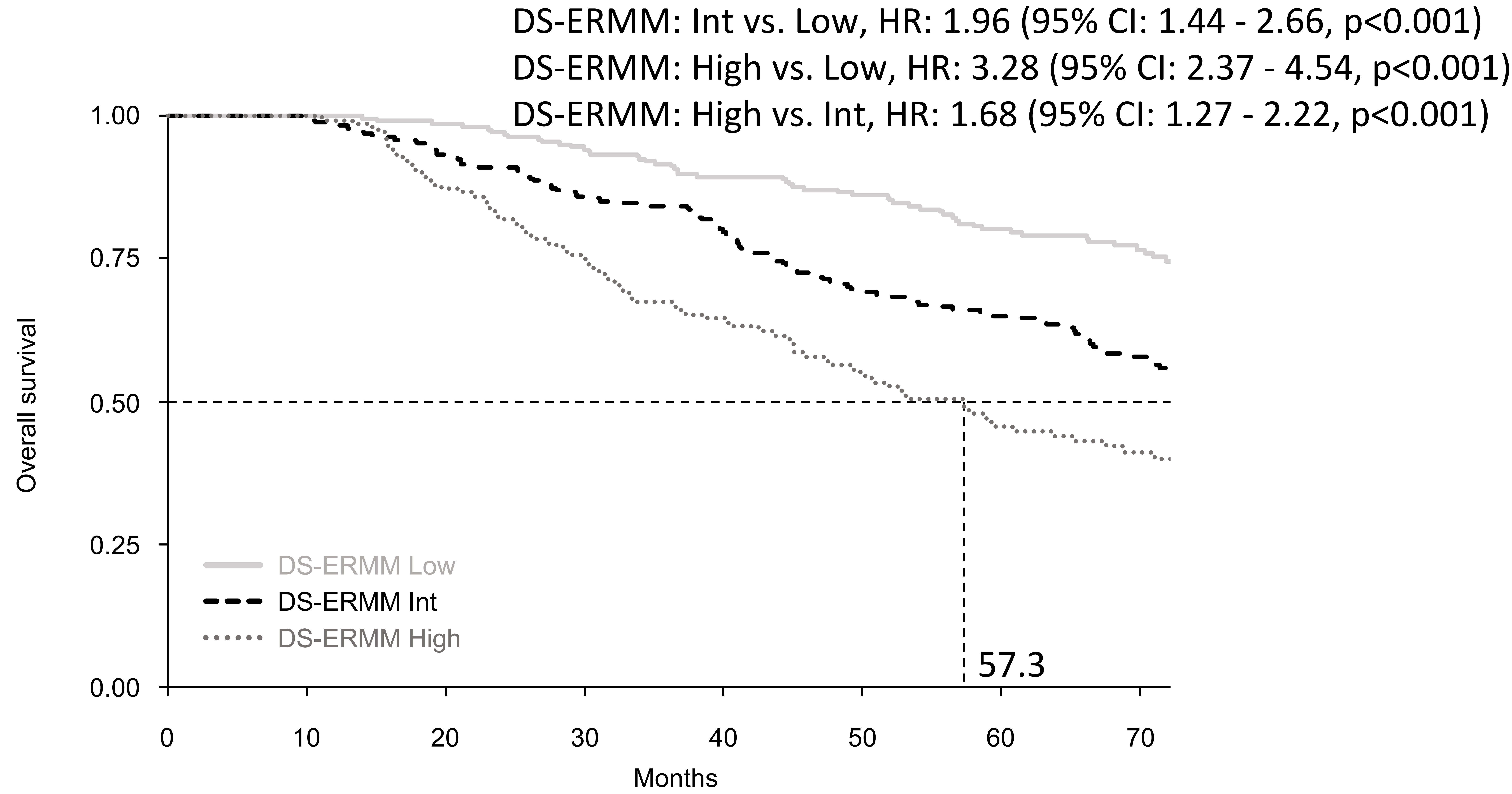
Number at risk (censored)

| | | | | | | | | |
|-------------|---------|----------|----------|----------|----------|----------|----------|-----------|
| S-ERMM Low | 574 (0) | 533 (16) | 475 (25) | 410 (39) | 350 (52) | 280 (68) | 223 (96) | 154 (139) |
| S-ERMM Int | 214 (0) | 189 (8) | 151 (12) | 123 (15) | 95 (19) | 65 (25) | 54 (27) | 29 (42) |
| S-ERMM High | 56 (0) | 43 (2) | 25 (2) | 20 (2) | 16 (2) | 10 (3) | 8 (4) | 7 (4) |

Number at risk (censored)

Figure 3

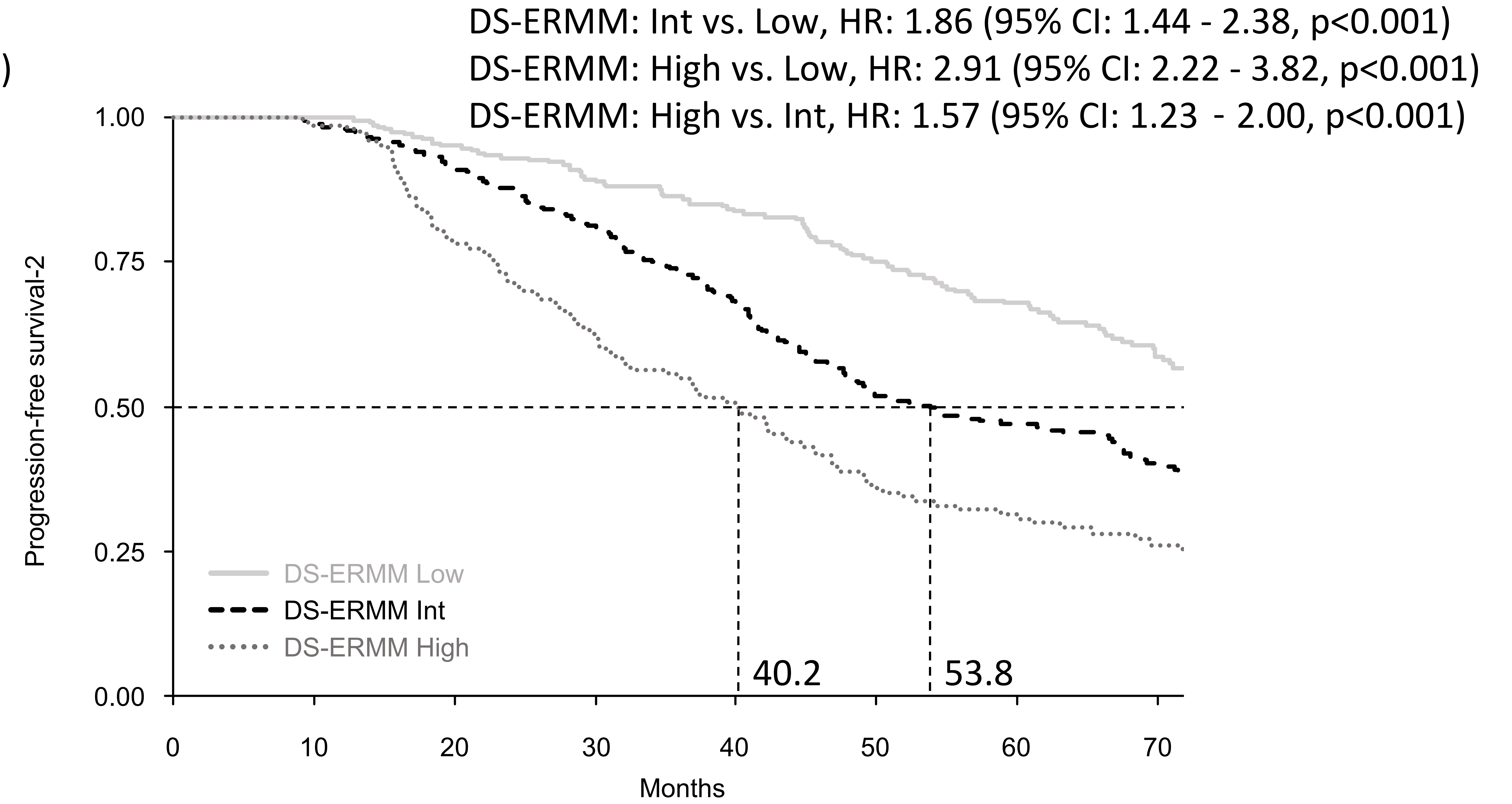
A. OS



| | | | | | | | | |
|--------------|---------|---------|---------|----------|----------|----------|----------|----------|
| DS-ERMM Low | 250 (0) | 250 (0) | 242 (5) | 221 (15) | 201 (24) | 181 (37) | 152 (54) | 117 (83) |
| DS-ERMM Int | 271 (0) | 268 (1) | 245 (8) | 221 (13) | 200 (19) | 159 (34) | 132 (52) | 94 (77) |
| DS-ERMM High | 152 (0) | 151 (1) | 129 (4) | 110 (6) | 90 (10) | 72 (15) | 58 (17) | 39 (31) |

Number at risk (censored)

B. PFS2



| | | | | | | | | |
|--------------|---------|---------|---------|----------|----------|----------|----------|---------|
| DS-ERMM Low | 250 (0) | 250 (0) | 235 (3) | 212 (12) | 191 (20) | 162 (30) | 132 (45) | 91 (70) |
| DS-ERMM Int | 271 (0) | 267 (1) | 240 (7) | 210 (12) | 169 (18) | 121 (28) | 96 (42) | 66 (59) |
| DS-ERMM High | 152 (0) | 150 (0) | 117 (2) | 92 (4) | 72 (6) | 49 (9) | 41 (11) | 24 (22) |

Number at risk (censored)