

# Heterogeneity in long-term outcomes for patients with Revised International Staging System stage II, newly diagnosed multiple myeloma

Anaïs Schavgoulidze,<sup>1</sup> Valérie Lauwers-Cances,<sup>2</sup> Aurore Perrot,<sup>1</sup> Titouan Cazaubiel,<sup>3</sup> Marie-Lorraine Chretien,<sup>4</sup> Philippe Moreau,<sup>5</sup> Thierry Facon,<sup>6</sup> Xavier Leleu,<sup>7</sup> Lionel Karlin,<sup>8</sup> Anne-Marie Stoppa,<sup>9</sup> Olivier Decaux,<sup>10</sup> Karim Belhadj,<sup>11</sup> Bertrand Arnulf,<sup>12</sup> Mohamad Mohty,<sup>13</sup> Clara Mariette,<sup>14</sup> Cécile Fohrer-Sonntag,<sup>15</sup> Pascal Lenain,<sup>16</sup> Jean-Pierre Marolleau,<sup>17</sup> Mourad Tiab,<sup>18</sup> Carla Araujo,<sup>19</sup> Frédérique Orsini-Piocelle,<sup>20</sup> Arnaud Jaccard,<sup>21</sup> Murielle Roussel,<sup>1</sup> Lotfi Benboubker,<sup>22</sup> Jean-Richard Eveillard,<sup>23</sup> Mamoun Dib,<sup>24</sup> Marion Divoux,<sup>25</sup> Michel Attal,<sup>1</sup> Hervé Avet-Loiseau<sup>1</sup> and Jill Corre<sup>1</sup>

<sup>1</sup>Institut Universitaire du Cancer de Toulouse-Oncopole and Centre de Recherches en Cancérologie de Toulouse Institut National de la Santé et de la Recherche Médicale, Toulouse; <sup>2</sup>Centre Hospitalier Universitaire Toulouse, Toulouse; <sup>3</sup>Centre Hospitalier Universitaire Bordeaux, Bordeaux; <sup>4</sup>Centre Hospitalier Universitaire Dijon, Dijon; <sup>5</sup>Centre Hospitalier Universitaire Nantes, Nantes; <sup>6</sup>Centre Hospitalier Universitaire Lille, Lille; <sup>7</sup>Centre Hospitalier Universitaire Poitiers, Poitiers; <sup>8</sup>Centre Hospitalier Universitaire Lyon, Lyon; <sup>9</sup>Institut Paoli Calmettes, Marseille; <sup>10</sup>Centre Hospitalier Universitaire Rennes, Rennes; <sup>11</sup>Centre Hospitalier Universitaire Créteil, Créteil; <sup>12</sup>Centre Hospitalier Universitaire, Hopital Saint Louis, Paris; <sup>13</sup>Centre Hospitalier Universitaire, Hopital Saint-Antoine, Paris; <sup>14</sup>Centre Hospitalier Universitaire Grenoble Alpes, Grenoble; <sup>15</sup>Centre Hospitalier Universitaire Hautepierre, Strasbourg; <sup>16</sup>Centre de Lutte Contre le Cancer - Centre Henri Becquerel, Rouen; <sup>17</sup>Centre Hospitalier Universitaire Amiens, Amiens; <sup>18</sup>Centre Hospitalier Départemental Vendée, La Roche-sur-Yon; <sup>19</sup>Centre Hospitalier de la Côte Basque, Bayonne; <sup>20</sup>Centre Hospitalier Annecy Genevois, Metz-Tessy; <sup>21</sup>Centre Hospitalier Universitaire Limoges, Limoges; <sup>22</sup>Centre Hospitalier Régional Universitaire Tours, Tours; <sup>23</sup>Centre Hospitalier Universitaire de Brest, Brest; <sup>24</sup>Centre Hospitalier Universitaire Angers, Angers and <sup>25</sup>Centre Hospitalier Régional Universitaire Nancy Vandœuvre les Nancy, Nancy, France

**Correspondence:** Jill CORRE  
corre.jill@iuct-oncopole.fr


**Received:** December 22, 2021.

**Accepted:** May 5, 2022.

**Early view:** September 29, 2022.

<https://doi.org/10.3324/haematol.2021.280566>

©2023 Ferrata Storti Foundation

Published under a CC BY-NC license 

## Abstract

In the era of personalized treatment in multiple myeloma, high-risk patients must be accurately identified. The International Myeloma Working Group recommends using the Revised International Staging System (R-ISS) to pick out high-risk patients. The main purpose of our work was to explore the heterogeneity of outcome among R-ISS stage II patients assessing the impact of International Staging System (ISS) stage, chromosomal abnormalities and lactate dehydrogenase level in this subgroup. Data were collected from 1,343 patients up to 65 years old with newly diagnosed myeloma, enrolled in three clinical trials implemented by the Intergroupe Francophone du Myélome. All patients were eligible for intensive treatment. Patients in R-ISS stage II but ISS stage I had 1.6 times higher risk of death than patients in R-ISS stage I (adjusted hazard ratio=1.6; 95% confidence interval: 1.1-2.2;  $P=0.01$ ) and patients in R-ISS stage II but with ISS stage III had a better overall survival than patients in R-ISS stage III (adjusted hazard ratio=0.7; 95% confidence interval: 0.4-0.9,  $P=0.02$ ). However, among patients classified in R-ISS II, ISS stage and chromosomal abnormalities (del[17p] and t[4;14]) were still relevant prognostic factors for death. Dividing R-ISS stage II into three subgroups: ISS I with standard-risk chromosomal abnormalities, ISS II or III with standard-risk chromosomal abnormalities and patients with high-risk chromosomal abnormalities, median overall survival times were, respectively, not reached, 112 months and 71 months ( $P<0.001$ ). In conclusion, stratification of patients in the R-ISS stage II group can be improved by taking into account chromosomal abnormalities and ISS. However, this does not improve predictive performance of survival models.

## Introduction

Multiple myeloma (MM) is characterized by clonal accumulation of plasma cells in the bone marrow. Despite consider-

able progress in patients' treatment, there is still wide heterogeneity in outcomes. This can be explained at least partially by the large molecular heterogeneity of MM, with a subgroup of high-risk patients who, although benefiting from

all therapeutic improvements, do not compensate for their poor prognosis at diagnosis. This is the case for patients harboring high-risk chromosomal abnormalities (CA).<sup>1-7</sup>

Currently, the International Myeloma Working Group (IMWG) recommends using the Revised International Staging System (R-ISS) to identify high-risk patients. This score combines the International Staging System (ISS) evaluation (based on serum  $\beta_2$ -microglobulin and albumin levels), abnormal serum lactate dehydrogenase (LDH) level and three high-risk CA: del(17p), t(4;14) and t(14;16).<sup>8</sup> Combining predictive factors in a score easy to understand and calculate has the advantage of a fast scoring system and the disadvantages of an oversimplified one. Especially when the parameters constituting the score are independent prognostic factors, their predictive overlap is low, and the risk they confer is additive. Indeed, condensing factors with independent predictive significance into a super-score is rarely successful at improving patients' classification. Moreover, patients whose tumor harbors a t(4;14) translocation should not be considered as having the same risk as those harboring deletion 17p.<sup>9,10</sup> Furthermore, the choice of t(14;16) translocation as a CA of interest is still debated as no study has been able to demonstrate its independent impact on the prognosis of patients with MM.<sup>11-13</sup> Specifically, for patients to be considered as high-risk using the R-ISS, they must have at least ISS stage III and at least one of the specified CA or a LDH level higher than the upper limit of the normal range (ULN). This means that a patient with deletion 17p can be classified as middle risk (stage II of R-ISS) if he or she does not have ISS stage III at diagnosis and that a patient without any CA could be classified as high risk just because of his or her level of LDH, a biochemical parameter well-known for its lack of specificity.<sup>14</sup> These potential misclassifications of patients led us to question the clinical utility of using the R-ISS instead of the ISS, CA and LDH level separately.

In the era of personalized treatment in MM, high-risk patients must, more than ever, be accurately identified, especially for the construction of clinical trials. R-ISS has been assessed by some studies,<sup>15-23</sup> but to our knowledge, none of these studies has assessed the performance measures for survival prediction of this combined score compared to the use of the different factors separately. The main purpose of our work was to explore the heterogeneity of outcome among R-ISS stage II patients assessing the impact of ISS, CA and LDH level in this subgroup and to assess the predictive accuracy of the R-ISS and ISS on transplant-eligible patients with newly diagnosed MM (NDMM).

## Methods

### Patients and methods

Data were collected from NDMM patients up to 65 years old, enrolled in three clinical trials implemented by the

Intergroupe Francophone du Myélome (IFM): (i) IFM 2005-02 a phase III multicenter randomized, double-blind study comparing maintenance therapy using lenalidomide to placebo after autologous stem cell transplantation (NCT00430365); (ii) IFM/DFCI 2009 a phase III multicenter randomized, open-label study comparing a conventional-dose combination using lenalidomide, bortezomib and dexamethasone to high-dose treatment with autologous stem cell transplantation (NCT01191060); and (iii) IFM 2014-02 a phase III multicenter randomized, open-label study comparing the efficacy of combined high-dose chemotherapy using melphalan and bortezomib *versus* melphalan alone followed by stem cell transplant in front-line MM patients who were not progressing after induction therapy (NCT02197221). All patients gave written informed consent before entering the source trials. The three studies were approved by the institutional ethics committee of the different coordinating centers (Centre Hospitalier Universitaire Purpan, Toulouse, France for IFM 2005-02 and IFM 2014-02; University Hospital, Nantes, France for IFM/DFCI 2009).

All patients were eligible for high-dose treatment with autologous stem cell transplantation. All data needed for revised staging calculations were included in the data collection plan of these studies. High-risk CA were those defined by the R-ISS: del(17p), t(4;14) and t(14;16). R-ISS stage I includes patients with ISS stage I, no high risk CA and LDH level lower than the ULN. R-ISS stage III includes patients with ISS stage III and either high-risk CA or LDH level higher than the ULN. R-ISS stage II includes all other possible combinations.

### Statistical analysis

Categorical data were presented as counts and percentages and compared using the  $\chi^2$  test or Fisher exact test. Continuous variables were described by mean  $\pm$  standard deviation or median and interquartile range and compared using analysis of variance or the Kruskal-Wallis test. Follow-up duration was estimated using the reverse Kaplan-Meier method.<sup>24</sup> Overall survival (OS) was defined as the time from randomization to death or the last date the patient was known alive. Progression-free survival was defined as the time from randomization to the first documentation of progressive disease, or death due to any cause and patients without progression were censored at the last date of clinical evaluation. All surviving patients were censored after 10 years of follow-up. As it was not possible to adjust correctly for treatment received, because each trial was set up at a different time and had different experimental and control arms, all survival analyses were stratified on treatment arms assuming that the effect of prognostic factors would be similar across strata. This assumption was tested for ISS and R-ISS by fitting proportional hazards models with inter-

actions between the treatment arm and ISS and R-ISS. Hence, overall and progression-free survival curves, estimated using the Kaplan-Meier method, were compared using the stratified log-rank test. The hazard ratio (HR) along with 95% confidence interval (95% CI) for progression and for death were estimated by fitting multivariate stratified Cox proportional hazard models adjusted for age and sex, included as covariates in all models. Time-dependent receiver operating characteristic (ROC) curves were estimated at 10 years to assess predictive power. Discrimination was assessed by the Harrell concordance index (C-index) which estimates the proportion of all pairs of patients in whom prediction and outcome are concordant and takes values from 0.5 (no discrimination) to 1.0 (perfect discrimination).<sup>25</sup>  $R^2$  was estimated as a measure of the proportion of the survival time explained by the model.<sup>26</sup> To compare the different predictive values of R-ISS or ISS, sequential models were built in the same data set which included all patients in whom ISS, t(4;14), del(17p), t(14;16) and LDH were simultaneously available. Tests were two-sided, and  $P$  values  $<0.05$  were considered statistically significant. All analyses were performed using Stata version 14.2 (StataCorp).

## Results

The three pooled clinical trials had included 1,614 NDMM patients, 1,343 (83%) of whom were assessable by the R-ISS at diagnosis and constituted our database (486 patients from the IFM 2005-02 study, 623 patients from the IFM/DFCI 2009 and 234 from the IFM 2014-02 study). The median age of the patients was 58 years and 59% were male. Thirty-six percent of patients had ISS stage I, 47% ISS stage II and 17% ISS stage III; 18% had at least one high-risk CA and 19% had a LDH level higher than the ULN. The median duration of follow-up was 95 months (interquartile range, 81 months - not reached) and the estimated 5-year and 10-year probabilities of overall survival were 73% (95% CI: 70%-75%) and 46% (95% CI: 42%-49%), respectively. Due to high-risk CA or high LDH level, 138 of 487 (28%) ISS stage I patients were classified as R-ISS stage II and, conversely, the absence of high-risk CA and a normal LDH level had classified 116 of 222 (52%) ISS stage III patients as R-ISS stage II. Overall, 26%, 66% and 8% of patients had R-ISS stage I, II and III, respectively. The revised staging system results in a substantial increase in stage II category for R-ISS compared to ISS. The patients' and disease characteristics at diagnosis according to R-ISS disease stages are presented in Table 1. As predictable, increasing R-ISS was associated with a worse Eastern Cooperative Oncology Group performance status, a higher serum creatinine level and a lower hemoglobin level at diagnosis. Survival analyses by R-ISS categories showed that median overall survival

was not reached for the R-ISS stage I patients, and was 108 and 62 months, respectively, for R-ISS stage II and III whereas median progression-free survival times were 41, 35 and 27 months for R-ISS stage I, II and III patients, respectively (Figure 1). In the same cohort of patients, median overall survival was not reached for ISS stage I patients, and was 104 and 82 months for ISS stage II and III. The median progression-free survival times were, 41, 34 and 29 months for ISS stage I, II and III patients, respectively. To assess whether or not improvement in discrimination was obtained between ISS and R-ISS, we compared overall survival and progression-free survival as predicted by the two staging systems between patients in whom the stage was not modified by the revised classification and patients in whom the stage was changed. Patients with R-ISS stage II but ISS stage I had a 1.6 times higher risk of death than patients with R-ISS stage I (adjusted HR=1.6; 95% CI: 1.1-2.2;  $P=0.01$ ) whereas no statistically significant improvement was observed for progression-free survival (adjusted HR=1.1; 95% CI: 0.8-1.4;  $P=0.68$ ). Moreover, patients with R-ISS stage II but ISS stage III had a better overall survival than patients with R-ISS stage III (adjusted HR=0.6; 95% CI: 0.4-0.9;  $P=0.01$ ) whereas no statistically significant improvement was observed for progression-free survival (adjusted HR=0.8; 95% CI: 0.6-1.0;  $P=0.09$ ) (Figure 2).

Focusing on R-ISS stage II, a large and heterogeneous category which encompasses patients from ISS stage I to III, we checked whether ISS stages, CA and high LDH level were still relevant risk factors for death or progression in this subgroup (Table 2). In a multivariate Cox proportional analysis, we observed that LDH level was not an independent risk factor for death, that ISS stage II and ISS stage III patients had similar and higher risk of death than ISS stage I patients and also that patients with high-risk CA had a worse overall survival than standard-risk patients. These results were similar whatever the CA studied: del(17p) vs. no del(17p) (adjusted HR=2.7; 95% CI: 1.8-4.2), t(4;14) vs. no t(4;14) (adjusted HR=2.3; 95% CI: 1.6-3.3), and t(14;16) vs no t(14;16) (adjusted HR=2.1; 95% CI: 0.8-5.8) (Figure 3). For progression-free survival analyses, adverse cytogenetics was still a relevant risk factor in this subgroup of patients. In the light of this observation we divided R-ISS stage II into three subgroups (hereafter referred to as "modified R-ISS"): ISS stage I with standard-risk CA, ISS stage II or III with standard-risk CA, and high-risk CA patients. In these subgroups, median overall survival times were not reached, 112 months and 71 months, respectively ( $P<0.001$ ).

Adjusted for age and sex, the performance of the R-ISS, modified R-ISS and ISS at predicting survival was similar. Time-dependent ROC curves that assess the predictive ability of a marker, C-index, which estimates the proportion of all pairs of patients in whom prediction and outcome are concordant, and  $R^2$ , the explained variation of survival times by the model, showed no notable improvement in predic-



**Table 1.** Patient and disease characteristics at diagnosis according to Revised International Staging System disease stages.

	R-ISS			P-value
	Stage I N 349 (26.0%)	Stage II N 888 (66.1%)	Stage III N 106 (7.9%)	
Age, years				
Median	57.3	58.0	59.1	0.287
IQR	50.2-62.0	52.3-62.0	52.8-62.0	
Sex, N (%)				
Male	205 (58.7)	522 (58.8)	61 (57.5)	0.970
Female	144 (41.3)	366 (41.2)	45 (42.5)	
Performance status, N (%)				
0	177 (55.1)	382 (46.2)	30 (31.6)	<0.001
1	126 (39.3)	350 (42.4)	48 (50.5)	
2	18 (5.6)	94 (11.4)	17 (17.9)	
Monoclonal isotype, N (%)				
IgG	206 (59.0)	547 (61.6)	55 (51.9)	0.186
IgA	73 (20.9)	204 (23.0)	29 (27.4)	
Light-chain	62 (17.8)	121 (13.6)	18 (17.0)	
Other	8 (2.3)	16 (1.8)	4 (3.8)	
Hemoglobin, g/dL				
Mean (SD)	12.3 (1.7)	10.8 (1.7)	9.4 (1.4)	<0.001
Creatinine, µmol/L				
Median	75.0	80.0	95.0	<0.001
IQR	64.6-86.9	68.0-98.7	76.1-123.9	
ISS disease stage, N (%)				
Stage I	349 (100.0)	138 (15.5)	0 (0.0)	<0.001
Stage II	0 (0.0)	634 (71.4)	0 (0.0)	
Stage III	0 (0.0)	116 (13.1)	106 (100.0)	
High-risk CA, N (%)				
Standard risk	349 (100.0)	568 (78.3)	36 (39.1)	<.001
High-risk	0 (0.0)	157 (21.7)	56 (60.9)	
t(4;14), N (%)				
No	349 (100.0)	636 (87.1)	54 (58.7)	<.001
Yes	0 (0.0)	94 (12.9)	38 (41.3)	
del(17p), N (%)				
No	349 (100.0)	673 (91.8)	72 (78.3)	<0.001
Yes	0 (0.0)	60 (8.2)	20 (21.7)	
t(14;16), N (%)				
No	349 (100.0)	752 (98.8)	87 (95.6)	0.002
Yes	0 (0.0)	9 (1.2)	4 (4.4)	
LDH level, N (%)				
Higher than ULN	0 (0.0)	179 (22.0)	61 (64.2)	<0.001
Lower than ULN	349 (100.0)	634 (78.0)	34 (35.8)	

R-ISS: Revised International Staging System; IQR: interquartile range; SD: standard deviation; ISS: International Staging System; LDH: lactate dehydrogenase; CA: chromosomal abnormalities; ULN upper limit of normal range.

tion. The same results were observed for progression-free survival (Table 3).

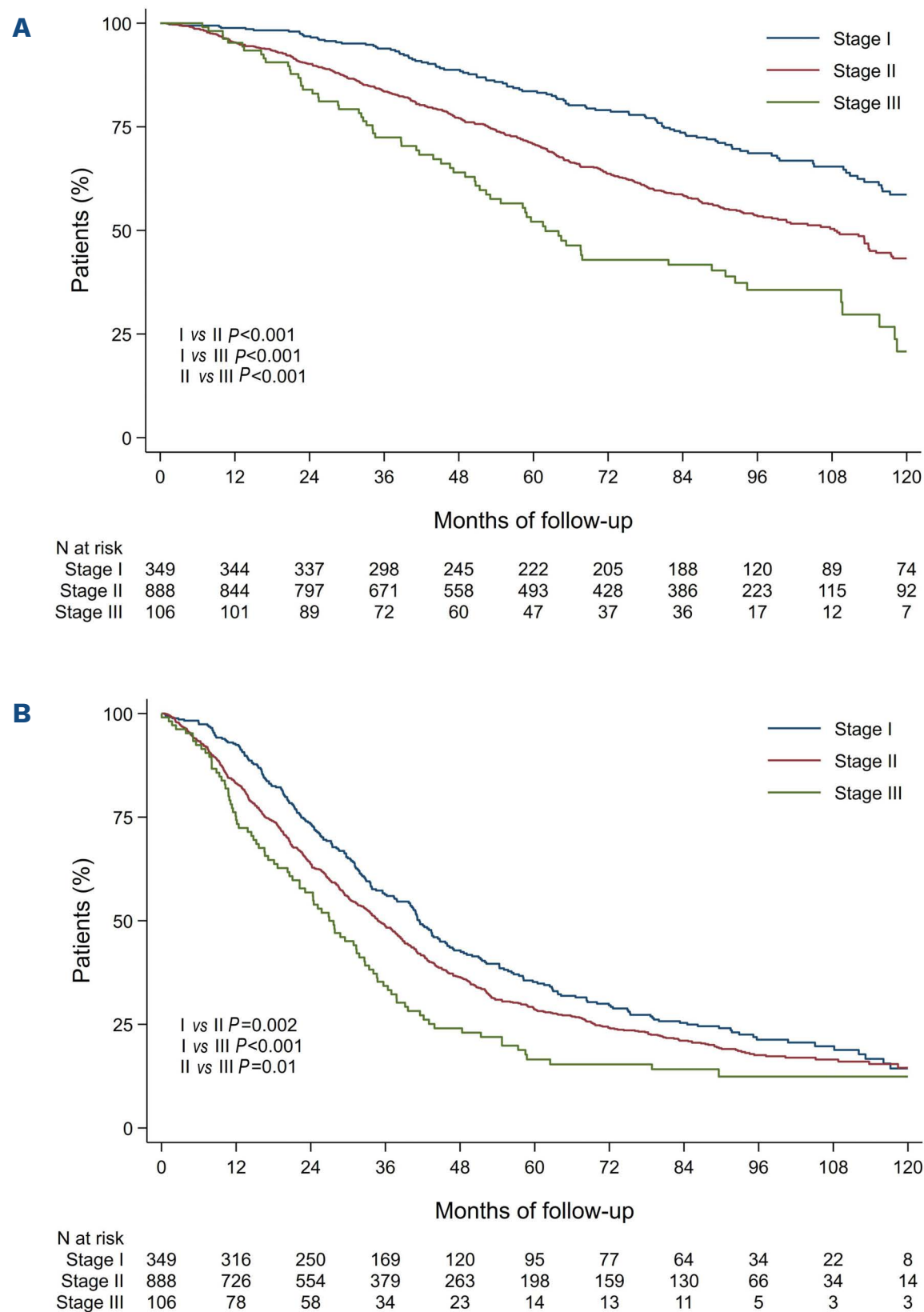
## Discussion

Accurate survival prediction is important as prognostic factors may soon influence treatment choice. The indisputable prognostic factors for MM already used to stratify patients in therapeutic trials are ISS disease stage and CA, in particular t(4;14) and del(17p), while the value of t(14;16)

is still being debated. Subsequently, a revised version of the ISS was proposed and validated in further studies.<sup>8,15-23</sup> This revised staging system was presented as simple and powerful because it allows three different prognostic groups to be identified based on three known prognostic factors: ISS, high-risk CA and LDH level higher than normal.<sup>8</sup> We cannot deny the simplification, since from three factors, only one was constructed, and the results of our study had effectively shown an improved prediction for ISS stages I and III by adding the presence of high-risk CA or high LDH level. This improvement in prediction in the

low-risk and high-risk classes of the R-ISS was achieved at the expense of the creation of an intermediate-risk class grouping the majority of patients (66% in our study) in which we have shown that the ISS and CA remain independent prognostic factors. We have also shown that the subdivision of the R-ISS stage II group according to ISS and CA allows a better stratification of patients but without improving the performance measurement of the model. The performance of prognostic models is rarely reported in clinical research, although this is strongly recommended in guidelines for transparent reporting.<sup>27</sup> Performance measures were not reported for either the ISS or the R-ISS at the time of their first publication and

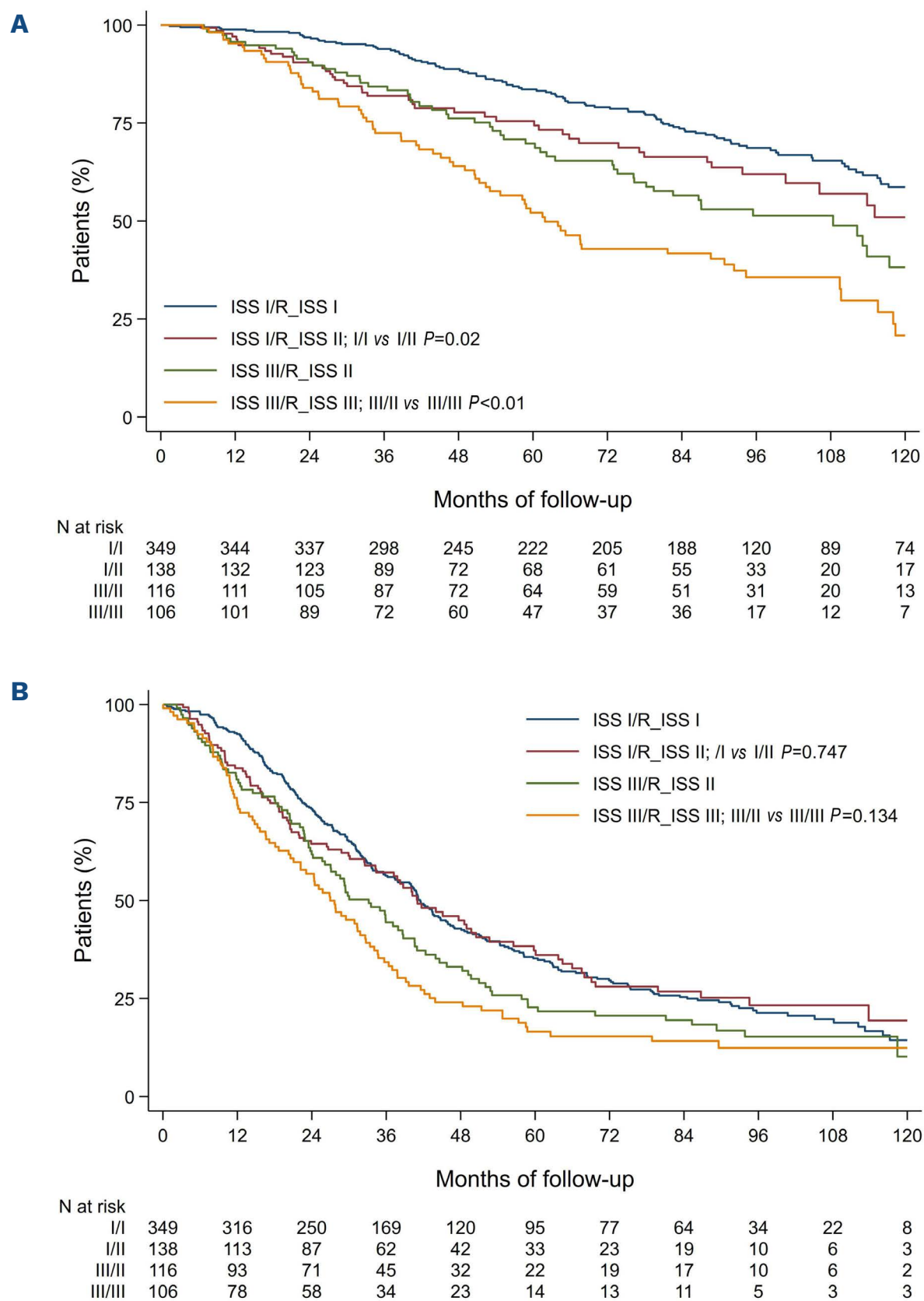
few studies have tested their performance. Abe *et al.*, Cheng *et al.* and Zhang *et al.* reported similar areas under the ROC curves or C-indexes between the ISS and R-ISS.<sup>19,28-30</sup> We have shown that whatever the performance measure, there was no significant improvement in discrimination and classification of patients using the ISS, R-ISS or its modification by subdivision of the stage II category, suggesting that a better performance measure could only be reached using other biomarkers than those used by the R-ISS. The better stratification of the R-ISS stage II subgroup according to the presence of high-risk CA was also described by the Mayo clinic but an ISS effect was not reported or perhaps was not investigated in this



**Figure 1. Kaplan-Meier survival curves according to Revised International Staging System disease stages.** (A) Overall survival. (B) Progression-free survival.

study due to the small sample size of their retrospective cohort.<sup>16</sup> In the same vein, Walker *et al.* have suggested splitting the large R-ISS stage II subgroup according to cytogenetic signature.<sup>23</sup> In their cohort of NDMM patients treated with novel therapies, Cho *et al.* described a similar event rate between ISS stage III patients reclassified as R-ISS stage II and those reclassified as R-ISS stage III, showing that the ISS effect is still of importance, despite the absence of high-risk CA or high LDH level.<sup>21</sup> Independent risk factors are cumulative and the more independent risk factors to which a patient is exposed, the higher is his or her likelihood of having a worse outcome.<sup>31</sup> Hence,

a simplified prognostic factor with fewer categories than the combined categories of the independent risk factors that make it up is unlikely to offer a better classification. In addition, in the era of personalized medicine, we argue that the paradigm concerning predictive factors is no longer related to the question “how simple should a staging system be?” but rather “what system allows me to classify patients correctly enough to offer them the most appropriate treatment?” This is to be understood, regardless of the number of variables to be included, since health information available today is increasingly numerous and precise, and contributes to more accurate pre-



**Figure 2. Kaplan-Meier survival curves comparing International Staging System disease stages I and III patients after Revised International Staging System reclassification.** (A) Overall survival. (B) Progression-free survival. ISS: International Staging System; R\_ISS: Revised International Staging System.

**Table 2.** Multivariate Cox proportional hazards regressions models for overall survival and progression-free survival, stratified by treatment and adjusted for age and sex, of Revised International Staging System stage II patients.

	Overall survival			Progression-free survival		
	Adjusted hazard ratio	95% CI	P-value	Adjusted hazard ratio	95% CI	P-value
<b>ISS disease stage</b>						
ISS stage II / stage I	1.70	1.14-2.53	0.009	1.19	0.89-1.60	0.239
ISS stage III / stage I	1.97	1.19-3.27	0.008	1.37	0.94-2.00	0.101
High-risk CA	2.41	1.78-3.26	<0.001	1.56	1.22-1.98	<0.001
High LDH level	1.06	0.75-1.50	0.737	0.91	0.70-1.19	0.497

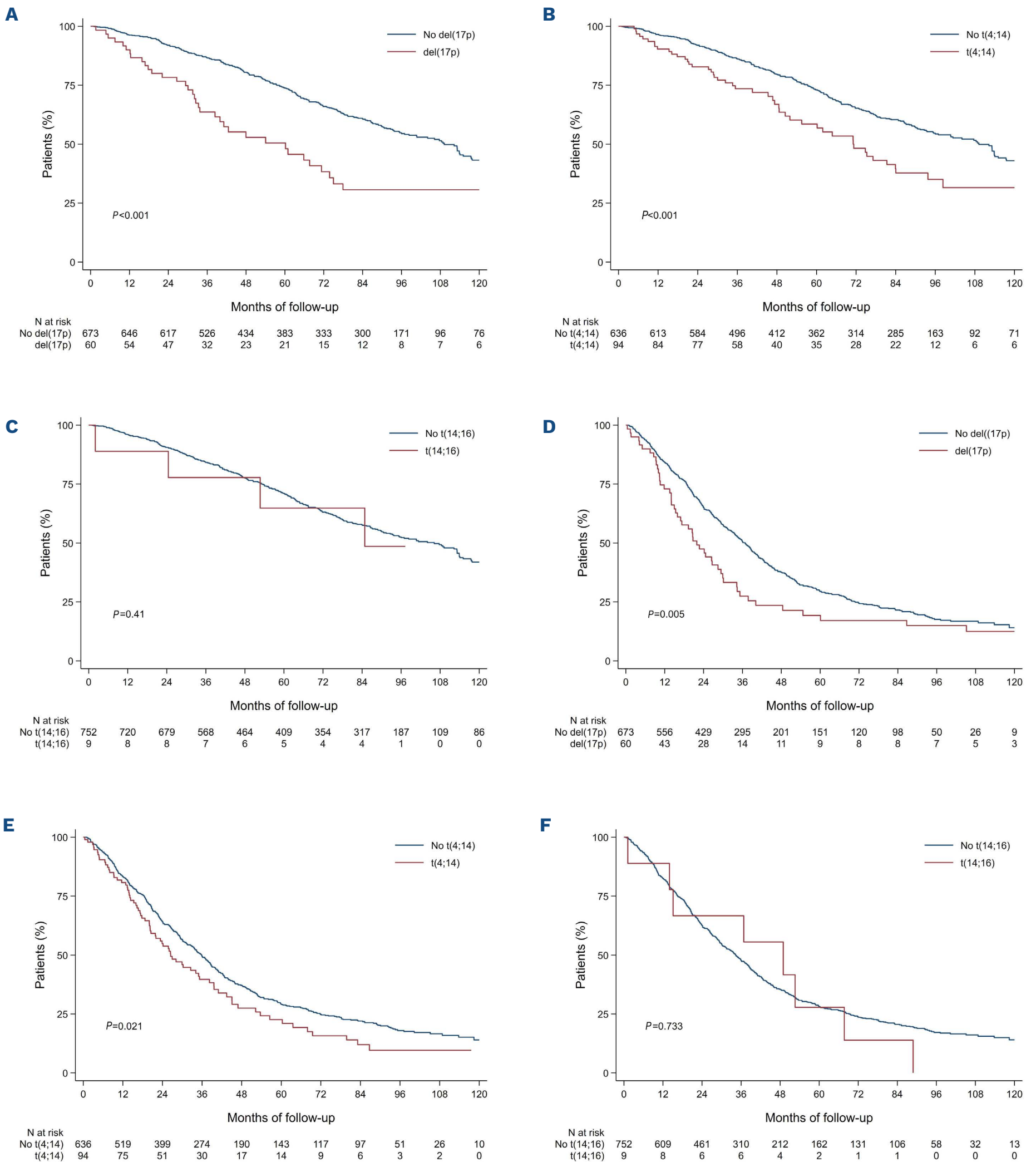
High-risk chromosomal abnormality is defined by any of del(17p) or t(4;14) or t(14;16). High lactate dehydrogenase level is defined by a measure higher than the upper limit of the normal range. 95% CI: 95% confidence interval; ISS: International Staging System; CA: chromosomal abnormalities; LDH: lactate dehydrogenase.

diction models, as biotechnological and informatics enable us to simplify the most complex information. To exemplify these comments, we have recently shown that a predictive index based on six cytogenetic abnormalities (the linear predictor score) outperformed the predictive ability of the current definition of a high-risk cytogenetic group, the ISS or R-ISS alone and also the separate information based on ISS and the presence of del(17p) or t(4;14).<sup>5</sup> For example, the C-index reached a value of 0.70 for the linear predictor score against only 0.55 for the R-ISS. Unfortunately, in the current study, we could not compare R-ISS to the linear predictor score as the necessary cytogenetic data, such as chromosome 1 abnormalities, were not available. Some other studies suggest improving the R-ISS classification by adding other criteria such as the detection of circulating plasma cells<sup>32</sup> or gene-expression data.<sup>33</sup> Some teams go further and suggest moving beyond traditional assays (cytogenetics) to favor whole genome sequencing or even whole exome sequencing to capture all the complexity of biological features.<sup>34</sup> Obviously, we know that a score cannot be perfect since prognosis is a moving target. It evolves with the development of new therapeutic strategies, and prognostic scores that are efficient today will be rapidly outdated in the future. At the individual level, prognosis also changes according to relapses and the clonal evolution of the disease. Moreover, it has been shown that an undetectable minimal residual disease could overcome the poor prognosis of patients with high-risk myeloma; a single estimate of prognosis at diagnosis is no longer sufficient and the depth of response has to be taken into account.<sup>35</sup> Finally, we should consider that both the ISS and R-ISS were developed to better balance adverse biology across study arms in randomized trials and were not designed for the

purpose of making treatment decisions for individual patients. Therefore, these scores probably do not have the precision needed to guide clinical decision-making for individual patients. However, as they are widely used by clinicians it is important to draw attention to the heterogeneity of patients with R-ISS stage II and to try to redefine this category.

The strength of this study lies in the use of more than one performance measure for survival models, time-dependent ROC analysis, the C-index and  $R^2$  to quantify the estimated predictive values of the ISS and R-ISS, but also in analyses performed using data from three large clinical trials, which ensured the quality of data with a sufficient follow-up of patients. However, this was also a limitation of the study as the assessment of the predictive ability in patients included in clinical trials limited our results to selected, newly diagnosed patients treated with high-dose chemotherapy. Nevertheless, we do not expect different results for older patients or for those who do not undergo transplantation, since ISS and high-risk CA are also independent prognostic factors in real-life settings.<sup>5,36-50</sup> In any case, our results must be confirmed in an older population. Another limitation of our study is the underrepresentation of the cytogenetic aberration t(14;16), which is part of the definition of the R-ISS, although its prognostic significance has not been confirmed by all studies.<sup>11,13,36,51</sup> The frequency of t(14;16) was low in our study (1.1%) whereas it can reach 3.5% in the real population. One explanation for this could be that information on this CA was specifically required in the IFM/DFCI 2009 study, whereas it was only recorded as "other CA" in the IFM 2005-02 and the IFM 2014 studies, as the IFM does not consider t(14;16) as an independent risk factor.<sup>11</sup> Moreover, LDH level was not reported as a significant prognostic factor in our study, contrary to the main





**Figure 3. Kaplan-Meier survival curves according to high-risk cytogenetic abnormalities among Revised International Staging System stage II patients.** (A) Overall survival according to del(17p) status. (B) Overall survival according to t(4;14) status. (C) Overall survival according to t(14;16) status. (D) Progression-free survival according to del(17p) status. (E) Progression-free survival according to t(4;14) status. (F) Progression-free survival according to t(14;16) status.



**Table 3.** R<sup>2</sup> values, C-index and time-dependent receiver operating characteristics, with 95% confidence intervals, for different Cox proportional hazards regression models for overall survival and progression-free survival of patients with multiple myeloma.

Models	Overall survival			Progression free survival		
	R <sup>2</sup>	C-index	Time-dependent ROC	R <sup>2</sup>	C-index	Time-dependent ROC
ISS disease stage, 95% CI	7.4 4.2-11.1	0.61 0.59-0.64	63.2 58.3-68.1	2.3 1.0-4.2	0.56 0.54-0.58	59.7 53.8-65.6
R-ISS disease stage, 95% CI	9.4 5.8-13.6	0.62 0.59-0.65	65.0 59.4-70.6	2.7 1.2-4.7	0.56 0.54-0.59	58.2 52.3-64.0
Modified R-ISS disease stage by dividing the stage II group, 95% CI	13.9 9.4-18.7	0.62 0.59-0.64	65.3 60.6-70.1	3.8 2.0-6.0	0.58 0.56-60.4	58.3 52.8-63.9

ROC: receiver operating characteristic; ISS: International Staging System; R-ISS: Revised International Staging System.

study that validated the R-ISS,<sup>8</sup> but it should be remembered that a high LDH level was not finally retained in the construction of the ISS either.<sup>38</sup> This discrepancy may also reflect the fact that increased LDH cannot be related to damage to a specific organ, thus making it an inconstant or non-specific biological risk factor. In our study, in order to ensure that each patient was properly classified according to his or her LDH level, we checked the normal ranges of LDH in 80 laboratories over the study period as normal levels of LDH in the blood can vary depending on the laboratory. Moreover, our population was younger than the population who served to establish the R-ISS and in better health than the whole myeloma population. To be included in our trials, patients had to be under 66 years of age and have normal or subnormal kidney, liver, heart and lung function to be eligible for a transplant. As LDH levels are elevated in the aging population and also in numerous clinical conditions, our selection of patients may have resulted in the creation of a less high-risk group of patients.

In conclusion, our data suggest that the R-ISS accurately classifies patients into stages I and III but that within the R-ISS stage II group, ISS and CA are still relevant prognostic factors. As the R-ISS has become the standard for risk stratification in clinical trials, our data suggest that the R-ISS stage II group could be divided according to the presence or not of high-risk CA and ISS to refine stratification. Finally, improving the performance of any model is

a challenge and it seems that models have to be made more complex to increase their performance. We believe that this can be achieved by incorporating cytogenetics, genomics and also longitudinally relevant biomarkers such as imaging information and residual disease status.

#### Disclosures

*No conflicts of interest to disclose.*

#### Contributions

*JC, VL-C and HA-L designed the research; JC, AS and VL-C analyzed data and wrote the manuscript, which was reviewed and edited by the other co-authors; VL-C performed the statistical analysis. All other authors provided study samples and clinical data.*

#### Acknowledgments

*We thank the Intergroupe Francophone du Myélome for providing patients' samples and clinical data, with special thanks to Sandrine Rollet.*

#### Data-sharing statement

*The data that support the findings of this study are available on request from the corresponding author, JC. Of note, patient-level data will be anonymized, and study documents will be redacted in order to protect the privacy of trial participants.*

## References

1. Avet-Loiseau H, Leleu X, Roussel M, et al. Bortezomib plus dexamethasone induction improves outcome of patients with t(4;14) myeloma but not outcome of patients with del(17p). *J Clin Oncol.* 2010;28(30):4630-4634.
2. Moreau P, Cavo M, Sonneveld P, et al. Combination of International Scoring System 3, high lactate dehydrogenase, and t(4;14) and/or del(17p) identifies patients with multiple myeloma (MM) treated with front-line autologous stem-cell transplantation at high risk of early MM progression-related death. *J Clin Oncol.* 2014;32(20):2173-2180.
3. Bock F, Lu G, Srour SA, et al. Outcome of patients with multiple myeloma and CKS1B gene amplification after autologous hematopoietic stem cell transplantation. *Biol Blood Marrow Transplant.* *J Am Soc Blood Marrow Transplant.*

- 2016;22(12):2159-2164.
4. Ziogas DC, Dimopoulos MA, Kastritis E. Prognostic factors for multiple myeloma in the era of novel therapies. *Expert Rev Hematol.* 2018;11(11):863-879.
  5. Perrot A, Lauwers-Cances V, Tournay E, et al. Development and validation of a cytogenetic prognostic index predicting survival in multiple myeloma. *J Clin Oncol.* 2019;37(19):1657-1665.
  6. Walker BA, Mavrommatis K, Wardell CP, et al. A high-risk, double-hit, group of newly diagnosed myeloma identified by genomic analysis. *Leukemia.* 2019;33(1):159-170.
  7. Corre J, Perrot A, Caillot D, et al. Del17p without TP53 mutation confers poor prognosis in intensively treated newly diagnosed multiple myeloma patients. *Blood.* 2021;137(9):1192-1195.
  8. Palumbo A, Avet-Loiseau H, Oliva S, et al. Revised International Staging System for multiple myeloma: a report from International Myeloma Working Group. *J Clin Oncol.* 2015;33(26):2863-2869.
  9. San Miguel JF, Schlag R, Khuageva NK, et al. Bortezomib plus melphalan and prednisone for initial treatment of multiple myeloma. *N Engl J Med.* 2008;359(9):906-917.
  10. Du C, Mao X, Xu Y, et al. 1q21 gain but not t(4;14) indicates inferior outcomes in multiple myeloma treated with bortezomib. *Leuk Lymphoma.* 2020;61(5):1201-1210.
  11. Avet-Loiseau H, Malard F, Campion L, et al. Translocation t(14;16) and multiple myeloma: is it really an independent prognostic factor? *Blood.* 2011;117(6):2009-2011.
  12. Fonseca R, Blood E, Rue M, et al. Clinical and biologic implications of recurrent genomic aberrations in myeloma. *Blood.* 2003;101(11):4569-4575.
  13. Goldman-Mazur S, Jurczynszyn A, Castillo JJ, et al. A multicenter retrospective study of 223 patients with t(14;16) in multiple myeloma. *Am J Hematol.* 2020;95(5):503-509.
  14. Jacobs EL, Haskell CM. Clinical use of tumor markers in oncology. *Curr Probl Cancer.* 1991;15(6):299-360.
  15. Kastritis E, Terpos E, Roussou M, et al. Evaluation of the Revised International Staging System in an independent cohort of unselected patients with multiple myeloma. *Haematologica.* 2017;102(3):593-599.
  16. González-Calle V, Slack A, Keane N, et al. Evaluation of Revised International Staging System (R-ISS) for transplant-eligible multiple myeloma patients. *Ann Hematol.* 2018;97(8):1453-1462.
  17. Ozaki S, Handa H, Saitoh T, et al. Evaluation of the Revised International Staging System (R-ISS) in Japanese patients with multiple myeloma. *Ann Hematol.* 2019;98(7):1703-1711.
  18. Chen H-M, Wei W, Peng R, et al. [Clinical application of R-ISS staging system in 412 newly diagnosed patients with multiple myeloma]. *Zhongguo Shi Yan Xue Ye Xue Za Zhi.* 2019;27(1):110-114.
  19. Abe Y, Sunami K, Yamashita T, et al. Improved survival outcomes and relative youthfulness of multiple myeloma patients with t(4;14) receiving novel agents are associated with poorer performance of the Revised International Staging System in a real aging society. *Oncotarget.* 2019;10(5):595-605.
  20. Jimenez-Zepeda VH, Duggan P, Neri P, et al. Revised International Staging System applied to real world multiple myeloma patients. *Clin Lymphoma Myeloma Leuk.* 2016;16(9):511-518.
  21. Cho H, Yoon DH, Lee JB, et al. Comprehensive evaluation of the Revised International Staging System in multiple myeloma patients treated with novel agents as a primary therapy. *Am J Hematol.* 2017;92(12):1280-1286.
  22. Tandon N, Rajkumar SV, LaPlant B, et al. Clinical utility of the Revised International Staging System in unselected patients with newly diagnosed and relapsed multiple myeloma. *Blood Cancer J.* 2017;7(2):e528-e528.
  23. Walker I, Coady A, Neat M, et al. Is the Revised International Staging System for myeloma valid in a real world population? *Br J Haematol.* 2018;180(3):451-454.
  24. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials.* 1996;17(4):343-346.
  25. Newson RB. Comparing the predictive powers of survival models using Harrell's C or Somers' D. *Stata J.* 2010;10(3):339-358.
  26. Royston P. Explained variation for survival models. *Stata J.* 2006;6(1):83-96.
  27. Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* 2015;162(1):W1-73.
  28. Zhang Y, Chen X-L, Chen W-M, Zhou H-B. Prognostic nomogram for the overall survival of patients with newly diagnosed multiple myeloma. *Biomed Res Int.* 2019;2019:5652935.
  29. Cheng Q, Zhao F, Zhang B, et al. Prognostic nomogram incorporating cytokines for overall survival in patients with newly diagnosed multiple myeloma. *Int Immunopharmacol.* 2021;99:108016.
  30. Cheng Q, Cai L, Zhang Y, et al. Circulating plasma cells as a biomarker to predict newly diagnosed multiple myeloma prognosis: developing nomogram prognostic models. *Front Oncol.* 2021;11:639528.
  31. Weinhold N, Salwender HJ, Cairns DA, et al. Chromosome 1q21 abnormalities refine outcome prediction in patients with multiple myeloma - a meta-analysis of 2,596 trial patients. *Haematologica.* 2021;106(10):2754-2758.
  32. Galieni P, Travaglini F, Vagnoni D, et al. The detection of circulating plasma cells may improve the Revised International Staging System (R-ISS) risk stratification of patients with newly diagnosed multiple myeloma. *Br J Haematol.* 2021;193(3):542-550.
  33. Kuiper R, Zweegman S, van Duin M, et al. Prognostic and predictive performance of R-ISS with SKY92 in older patients with multiple myeloma: the HOVON-87/NMSG-18 trial. *Blood Adv.* 2020;4(24):6298-6309.
  34. Rustad EH, Yellapantula VD, Glodzik D, et al. Revealing the impact of structural variants in multiple myeloma. *Blood Cancer Discov.* 2020;1(3):258-273.
  35. Goicoechea I, Puig N, Cedena M-T, et al. Deep MRD profiling defines outcome and unveils different modes of treatment resistance in standard- and high-risk myeloma. *Blood.* 2021;137(1):49-60.
  36. Fonseca R, Bergsagel PL, Drach J, et al. International Myeloma Working Group molecular classification of multiple myeloma: spotlight review. *Leukemia.* 2009;23(12):2210-2221.
  37. Keats JJ, Reiman T, Maxwell CA, et al. In multiple myeloma, t(4;14)(p16;q32) is an adverse prognostic factor irrespective of FGFR3 expression. *Blood.* 2003;101(4):1520-1529.
  38. Greipp PR, San Miguel J, Durie BGM, et al. International Staging System for multiple myeloma. *J Clin Oncol.* 2005;23(15):3412-3420.
  39. Avet-Loiseau H, Attal M, Moreau P, et al. Genetic abnormalities and survival in multiple myeloma: the experience of the Intergroupe Francophone du Myélome. *Blood.* 2007;109(8):3489-3495.
  40. Chang H, Qi X, Jiang A, et al. 1p21 deletions are strongly associated with 1q21 gains and are an independent adverse prognostic factor for the outcome of high-dose chemotherapy in patients with multiple myeloma. *Bone Marrow Transplant.* 2010;45(1):117-121.
  41. Waheed S, Shaughnessy JD, van Rhee F, et al. International Staging System and metaphase cytogenetic abnormalities in the

- era of gene expression profiling data in multiple myeloma treated with total therapy 2 and 3 protocols. *Cancer*. 2011;117(5):1001-1009.
42. Dimopoulos MA, Kastritis E, Michalis E, et al. The International Scoring System (ISS) for multiple myeloma remains a robust prognostic tool independently of patients' renal function. *Ann Oncol*. 2012;23(3):722-729.
43. Avet-Loiseau H, Durie BGM, Cavo M, et al. Combining fluorescent in situ hybridization data with ISS staging improves risk assessment in myeloma: an International Myeloma Working Group collaborative project. *Leukemia*. 2013;27(3):711-717.
44. Hebraud B, Leleu X, Lauwers-Cances V, et al. Deletion of the 1p32 region is a major independent prognostic factor in young patients with myeloma: the IFM experience on 1195 patients. *Leukemia*. 2014;28(3):675-679.
45. Jian Y, Chen X, Zhou H, et al. Prognostic impact of cytogenetic abnormalities in multiple myeloma: a retrospective analysis of 229 patients. *Medicine (Baltimore)*. 2016;95(19):e3521.
46. Sergentanis TN, Kastritis E, Terpos E, Dimopoulos MA, Psaltopoulou T. Cytogenetics and survival of multiple myeloma: isolated and combined effects. *Clin Lymphoma Myeloma Leuk*. 2016;16(6):335-340.
47. Pawlyn C, Morgan GJ. Evolutionary biology of high-risk multiple myeloma. *Nat Rev Cancer*. 2017;17(9):543-556.
48. Pawlyn C, Davies FE. Toward personalized treatment in multiple myeloma based on molecular characteristics. *Blood*. 2019;133(7):660-675.
49. Thakurta A, Ortiz M, Bleclua P, et al. High subclonal fraction of 17p deletion is associated with poor prognosis in multiple myeloma. *Blood*. 2019;133(11):1217-1221.
50. Sato S, Kamata W, Okada S, Tamai Y. Clinical and prognostic significance of t(4;14) translocation in multiple myeloma in the era of novel agents. *Int J Hematol*. 2021;113(2):207-213.
51. Mina R, Joseph NS, Gay F, et al. Clinical features and survival of multiple myeloma patients harboring t(14;16) in the era of novel agents. *Blood Cancer J*. 2020;10(4):40.