

# Performance analysis of AL amyloidosis cardiac biomarker staging systems with special focus on renal failure and atrial arrhythmia

Tobias Dittrich,<sup>1,2,3</sup> Axel Benner,<sup>4</sup> Christoph Kimmich,<sup>1,2</sup> Fabian aus dem Siepen,<sup>2,5</sup> Kaya Veelken,<sup>1,2</sup> Arnt V. Kristen,<sup>2,5</sup> Tilmann Bochtler,<sup>1,2,3</sup> Hugo A. Katus,<sup>5</sup> Carsten Müller-Tidow,<sup>1,2</sup> Ute Hegenbart<sup>1,2</sup> and Stefan O. Schönland<sup>1,2</sup>

<sup>1</sup>Department of Internal Medicine V, Division of Hematology/Oncology, Heidelberg University Hospital; <sup>2</sup>Amyloidosis Center, Heidelberg University Hospital; <sup>3</sup>Clinical Cooperation Unit Molecular Hematology/Oncology, German Cancer Research Center (DKFZ) and Department of Internal Medicine V, Heidelberg University Hospital; <sup>4</sup>Division of Biostatistics, German Cancer Research Center (DKFZ), Heidelberg and <sup>5</sup>Department of Internal Medicine III, Division of Cardiology, Heidelberg University Hospital, Germany



Haematologica 2019  
Volume 104(7):1451-1459

## ABSTRACT

Systemic light chain amyloidosis is a rare and life-threatening disorder, for which accurate risk stratification is crucial. Current cardiac staging systems (MAYO2004, MAYO3b, and MAYO2012) are mainly based on biomarkers, which have uncertain reliability in the context of atrial fibrillation, arrhythmia or pacemaker stimulation as well as renal insufficiency. We compared the performance of the established staging systems with particular regard to these comorbidities in 1,224 patients with systemic light chain amyloidosis diagnosed at our center from July 2002 until March 2017. We first characterized the subsets with an estimated glomerular filtration rate <50 mL/min/1.73 m<sup>2</sup> (415 patients) and any kind of atrial arrhythmia (183 patients) as unique high-risk subgroups with similarly increased cardiac biomarkers ( $\chi^2$ -test, all  $P < 0.001$ ). This resulted in a shift towards higher risk stages and reduced median overall survival compared to those of patients with better kidney function or without atrial arrhythmia in univariate analyses (13 vs. 46 months and 17 vs. 53 months, respectively; both  $P < 0.001$ ). Performance analysis revealed that predictions in the entire cohort were least precise with the MAYO2004 staging system and most precise with the MAYO3b system. This performance pattern was almost preserved for patients with an estimated glomerular filtration rate <50 mL/min/1.73 m<sup>2</sup>, but less so for those with atrial arrhythmias. The MAYO3b staging system was most robust. Importantly, atrial arrhythmia retained its prognostic value in multivariable analysis including age, difference between involved and uninvolved free light chains, and any staging system, while estimated glomerular filtration rate <50 mL/min/1.73 m<sup>2</sup> was not statistically significant in multivariable analysis with the MAYO3b staging system. In conclusion, our results favor the MAYO3b staging system due to its consistently best performance and retained applicability in the subgroups with atrial arrhythmia and estimated glomerular filtration rate <50 mL/min/1.73 m<sup>2</sup>.

Selected contents of this manuscript were presented during the XVIth International Symposium on Amyloidosis in Kumamoto, Japan (oral presentation on 28th March 2018, OP48).

## Correspondence:

STEFAN O. SCHÖNLAND  
stefan.schoenland@med.uni-heidelberg.de

UTE HEGENBART  
ute.hegenbart@med.uni-heidelberg.de

Received: August 26, 2018.

Accepted: January 15, 2019.

Pre-published: January 17, 2019.

doi:10.3324/haematol.2018.205336

Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: [www.haematologica.org/content/104/7/1451](http://www.haematologica.org/content/104/7/1451)

©2019 Ferrata Storti Foundation

Material published in Haematologica is covered by copyright. All rights are reserved to the Ferrata Storti Foundation. Use of published material is allowed under the following terms and conditions:

<https://creativecommons.org/licenses/by-nc/4.0/legalcode>.

Copies of published material are allowed for personal or internal use. Sharing published material for non-commercial purposes is subject to the following conditions:

<https://creativecommons.org/licenses/by-nc/4.0/legalcode>, sect. 3. Reproducing and sharing published material for commercial purposes is not allowed without permission in writing from the publisher.



## Introduction

Systemic light chain amyloidosis (AL) is a life-threatening protein deposition disorder with an annual incidence in the western world of about 8-12 cases per one million inhabitants.<sup>1</sup> The disease is initiated by clonal plasma cells in the bone marrow,<sup>2</sup> which secrete the amyloid-precursor proteins: free light chains (FLC) with a propensity to misfold and aggregate. In the fatal course of this disease, excessive amyloid causes dysfunction of vital organs such as the heart, kidneys or liver, which results in substantial morbidity and mortality.<sup>1,3</sup>

In recent decades, biomarkers have gained essential importance for the diagnosis, risk stratification, response assessment and monitoring of AL patients. Besides the FLC test, which allows for highly sensitive quantification of the involved FLC and underlying clonal disorder,<sup>4-6</sup> a wide range of organ biomarkers are available which reflect the severity of organ dysfunction and amyloid load in AL patients.<sup>7-14</sup> Given the complexity of the disease, accurate estimates of prognosis are of great value to clinicians for optimal primary management and treatment.

Involvement of the heart and the consequent cardiac dysfunction are the major factors determining prognosis and, especially, early death of AL patients.<sup>15,16</sup> Elevated cardiac biomarkers are, therefore, the most powerful prognostic determinants in AL. Serum levels of N-terminal pro-B-type natriuretic peptide (NT-proBNP) and cardiac troponin T (cTnT) were first found to predict survival in several cohorts of AL patients,<sup>10,11,15,16</sup> and were consequentially incorporated into the first widely used staging system for AL, hereafter referred to as “MAYO2004”, which defined three stages.<sup>17</sup>

Subsequently, the composition and biomarker thresholds of the staging system for AL patients were refined.<sup>3,11,13,18-22</sup> Two modifications have been widely accepted: the subgroup of patients with the most unfavorable prognosis, who had elevated levels of both biomarkers, could be further separated by an additional higher cut-off of serum NT-proBNP level (>8,500 ng/L) as well as a low systolic blood pressure.<sup>13</sup> Patients classified as MAYO2004 stage III with NT-proBNP >8,500 ng/L had a dismal prognosis with a median overall survival of less than 5 months.<sup>13</sup> The MAYO2004 system is, therefore, frequently applied with the modification that patients in stage III with an NT-proBNP >8,500 ng/L are assigned a “stage IIIb”, resulting in a system of four stages (hereafter referred to as “MAYO3b”).<sup>23</sup> Concurrently, plasma cell factors were incorporated by another four-stage system (“MAYO2012”) utilizing serum FLC level (threshold  $\geq 180$  mg/L) as well as NT-proBNP and troponins with different thresholds (NT-proBNP  $\geq 1,800$  pg/mL and cTnT  $\geq 0.025$   $\mu$ g/L) compared to those of MAYO2004 (NT-proBNP  $\geq 332$  pg/mL and cTnT  $\geq 0.035$   $\mu$ g/L).<sup>17,21</sup>

Given the importance of these three widely used staging systems for the management of AL patients, a systematic comparison of their performance is warranted. Moreover, serum levels of troponins and NT-proBNP are influenced by renal dysfunction and atrial arrhythmia or pacemaker stimulation, which is questioning the reliability of these biomarkers in subgroups with these concomitant disorders.<sup>24-28</sup> Although impaired renal function and atrial arrhythmia are frequent in AL patients, neither of them was considered as a confounder during the development of the staging systems. We, therefore, systematically compared the performance of the MAYO2004, MAYO3b and MAYO2012 staging systems in a large cohort of newly diagnosed AL patients with special regard to the subgroups of patients with impaired renal function and atrial arrhythmia or pacemaker stimulation.

## Methods

### Patients and medical records

We identified 1,224 patients with confirmed AL, consecutively treated and followed at the Heidelberg Amyloidosis Center

between July 2002 and March 2017, for whom results of the required cardiac biomarkers and FLC assay were available before initial treatment. All demographic and clinical information, including age, gender, hematologic and organ-related laboratory test results, were obtained from medical records. Clinical data of half of the patients had not been used for statistical analysis before. This retrospective study was approved by the Ethics Committee of the University of Heidelberg and was conducted in accordance with the principles of the Helsinki declaration. All participants gave written informed consent to retrospective analysis of their clinical data.

### Definitions

Organ involvement was defined according to consensus standards.<sup>3</sup> Detailed information about how the staging systems are applied and how atrial arrhythmia and impaired renal function are defined are provided in the *Online Supplement*.

### Statistical analysis

Statistical analysis was performed using the R statistical environment,<sup>29</sup> version 3.3.2 on a x86\_64-w64-mingw32/x64 (64-bit) platform, together with the packages ‘survival’ (version 2.38), ‘VennDiagram’ (version 1.6.18) and ‘pec’ (version 2.5.4). Continuous data are described by their median and range. The Mann-Whitney U test was used to test differences in continuous variables and the  $\chi^2$  test with continuity correction was used to test differences in categorical variables between two groups.

Overall survival curves were constructed according to Kaplan-Meier estimates and comparisons were made using the log-rank test. The median estimated time of observation was calculated based on the median time to censoring (reverse Kaplan-Meier). Univariable and multivariable Cox models were fitted to evaluate the influence of possible prognostic factors on overall survival. The proportional hazards assumption was tested as proposed by Grambsch & Therneau.<sup>30</sup> To illustrate the results of the Cox models, hazard ratios and corresponding 95% confidence intervals (95% CI) were calculated.

For the evaluation of the predictive accuracy of prognostic models, a .632+ bootstrap estimate of the prediction error was calculated for all event times using a time-dependent adaption of the Brier score.<sup>31</sup> For the evaluation of discriminative accuracy, a .632 bootstrap estimate of the time-dependent adaption of the concordance index<sup>32</sup> for time intervals of 6 months was calculated.

All statistical tests were two-sided. Results with *P* values not larger than 5% were considered statistically significant.

Further statistical methods are described in detail in the *Online Supplement*.

### Laboratory methods

All biomarkers were measured by standard commercially available assays. Reference values were as follows: FLC $\kappa$  3.3 – 19.4 mg/L, FLC $\lambda$  5.7 – 26.3 mg/L, NT-proBNP <125 ng/L and <334 ng/L (different assays), cTnT <0.03  $\mu$ g/L, cardiac troponin I (cTNI) <0.04  $\mu$ g/L and <0.6  $\mu$ g/L (different assays), high sensitivity troponin (hsTnT) <50 pg/mL. The cTnT assay was applied mainly until 2009, when it was replaced by assays with higher sensitivity or specificity, cTNI (since 2007) and hsTnT (since 2009). At least one troponin value was available for each included patient. The Modification of Diet in Renal Disease Study Group formula was applied to calculate the estimated glomerular filtration rate (eGFR).

## Results

The baseline characteristics of the 1,224 patients included in the current study are reported in Table 1. The median age of the patients was 63 years and the majority (60%)

Table 1. Baseline characteristics of the patients and their treatment.

Characteristic	All patients	Missing (%)	EGFR ≥50 mL/min 1.73m <sup>2</sup>	EGFR <50 mL/min 1.73m <sup>2</sup>	P-value	No AF	Any AF	P-value
Number	1224		809 (66.1)	415 (33.9)		1037 (85.0)	183 (15.0)	
Age, years	63 [30, 86]	0	62 [30, 85]	67 [39, 86]	<b>&lt;0.001</b>	62 [30, 86]	68 [39, 83]	<b>&lt;0.001</b>
Male sex	738 (60.3)	0	503 (62.2)	235 (56.6)	0.069	605 (58.3)	129 (70.5)	0.003
Intact heavy chain	550 (45.0)	0.2	352 (43.6)	198 (47.8)	0.181	451 (43.6)	96 (52.5)	0.033
Heavy chain, g/L	10.6 [0.2, 1810.0]	4.9	10.0 [0.5, 1810.0]	12.3 [0.2, 51.2]	0.11	10.3 [0.2, 1810.0]	13.3 [0.8, 59.0]	0.008
Light chain = lambda	946 (77.3)	0	642 (79.4)	304 (73.3)	<b>0.019</b>	803 (77.4)	139 (76.0)	0.731
dFLC, mg/L	215 [0, 19808]	0	198 [0, 19252]	242 [0, 19808]	0.08	199 [0.00, 19808]	248 [0, 12078]	0.012
ECOG PS ≤1	817 (74.1)	9.9	599 (81.5)	217 (59.0)	<b>&lt;0.001</b>	712 (75.6)	102 (64.2)	0.003
N. of organs involved	3 [1, 7]	0	2 [1, 7]	3 [1, 7]	<b>0.004</b>	2 [1, 7]	3 [1, 7]	0.003
Heart involved	916 (74.8)	0	589 (72.8)	327 (78.8)	<b>0.027</b>	743 (71.6)	170 (92.9)	<b>&lt;0.001</b>
NT-proBNP, ng/L	3211 [20, 565442]	0	2072 [20, 53794]	8031 [40, 565442]	<b>&lt;0.001</b>	2652 [20, 565442]	7010 [202, 216187]	<b>&lt;0.001</b>
cTnT, µg/L	0.03 [0.01, 0.95]	75.7*	0.01 [0.01, 0.50]	0.07 [0.01, 0.95]	<b>&lt;0.001</b>	0.02 [0.01, 0.95]	0.07 [0.01, 0.67]	<b>&lt;0.001</b>
cTnI, µg/L	0.07 [0.00, 2.97]	57.2*	0.06 [0.00, 2.10]	0.10 [0.02, 2.97]	<b>&lt;0.001</b>	0.06 [0.00, 2.97]	0.12 [0.04, 0.87]	<b>&lt;0.001</b>
hsTnT, pg/mL	47 [3, 3245]	27*	39 [3, 3245]	75 [6, 1350]	<b>&lt;0.001</b>	43 [3, 1350]	71 [6, 3245]	<b>&lt;0.001</b>
Kidney involved	749 (61.2)	0	427 (52.8)	322 (77.6)	<b>&lt;0.001</b>	652 (62.9)	95 (51.9)	0.006
eGFR, mL/min/1.73 m <sup>2</sup>	62.5 [ESRD, 262.4]	0	77.7 [50.0, 262.4]	32.0 [ESRD, 49.9]	<b>&lt;0.001</b>	65.4 [ESRD, 262.4]	55.0 [ESRD, 218.83]	<b>&lt;0.001</b>
Proteinuria, g/day	1.85 [0.01, 28.00]	12.3	1.04 [0.01, 18.27]	3.79 [0.03, 28.00]	<b>&lt;0.001</b>	2.18 [0.01, 28.00]	0.69 [0.03, 16.99]	<b>&lt;0.001</b>
Albuminuria, mg/d	1106 [1, 17160]	19.7	436 [1, 14551]	2386 [1, 17160]	<b>&lt;0.001</b>	1490 [1, 17160]	243 [1, 13618]	0.001
MAYO2004 stage III	570 (46.6)	0	307 (37.9)	263 (63.4)	<b>&lt;0.001</b>	445 (42.9)	123 (67.2)	<b>&lt;0.001</b>
MAYO3b stage IIIb	288 (23.5)	0	105 (13.0)	183 (44.1)	<b>&lt;0.001</b>	216 (20.8)	71 (38.8)	<b>&lt;0.001</b>
MAYO2012 stage IV	367 (30.0)	0	183 (22.6)	184 (44.3)	<b>&lt;0.001</b>	283 (27.3)	82 (44.8)	<b>&lt;0.001</b>
Renal stage III	144 (13.4)	12.3	0 (0.0)	144 (40.8)	<b>&lt;0.001</b>	127 (13.8)	17 (11.0)	0.416
Chemotherapy		3.5			<b>&lt;0.001</b>			<b>&lt;0.001</b>
ASCT	161 (13.6)		139 (17.8)	22 (5.5)		156 (15.6)	5 (2.8)	
other	892 (75.5)		590 (75.4)	302 (75.9)		750 (74.9)	138 (78.4)	
none	128 (10.8)		54 (6.9)	74 (18.6)		95 (9.5)	33 (18.8)	

Categorical data are shown as counts (% of respective total), continuous data are shown as medians [range]. Statistically significant differences are represented in bold, italics. Bone marrow plasmacytosis was preferentially evaluated by cytology. \*At least one troponin value was available for each patient. eGFR: estimated glomerular filtration rate; AF: atrial arrhythmia or pacemaker rhythm; dFLC: difference between involved and uninvolved free light chains; ECOG PS: Eastern Cooperative Oncology Group Performance Status; NT-proBNP: N-terminal pro-B-type natriuretic peptide; cTnT: cardiac troponin T; cTnI: cardiac troponin I; hsTnT: high-sensitivity cardiac troponin T; ASCT: high-dose chemotherapy followed by autologous stem cell transplantation; ESRD: end-stage renal disease.

were male. The median estimated time of observation (reverse Kaplan-Meier) of the study population was 45.5 months (95% CI: 41.7 - 50.8). The median overall survival of all the patients was 38.3 months (95% CI: 31.9 - 44.5) (Figure 1).

We first characterized the subgroups of patients with impaired renal function, defined as an eGFR <50 mL/min/1.73 m<sup>2</sup> (eGFR<50) (n=415, 34%) and atrial arrhythmia or pacemaker stimulation (henceforth, referred to as AF) (n=183, 15%) and identified the potential influence of these conditions on the established staging systems. We then evaluated the additional prognostic value of eGFR<50 and AF in the context of these staging systems. Finally, we compared the performance of these staging systems in the entire cohort and in each of the subgroups.

#### Patients with impaired renal function and atrial arrhythmias represent unique high-risk subgroups

Both impaired renal function and AF were associated with older age, lower Performance Status and more heart involvement (Table 1). Although approximately 40% of patients with AF also had an eGFR<50, both subgroups

had unique characteristics and were not interchangeable: the AF subgroup was enriched for men with heart involvement, while female sex (with a trend to statistical significance) and kidney organ involvement were features of the subgroup with impaired renal function (Table 1).

As a consequence of the unfavorable characteristics of the subgroups with eGFR<50 and AF, fewer patients were eligible for conventional or high-dose chemotherapy (Table 1) and their overall survival was significantly worse than that of their respective controls (17.5 vs. 52.9 and 12.6 vs. 45.5, months, both  $P<0.0001$ ) (Figure 1). Univariable Cox models showed that the risk of death was almost double in the presence of either of these conditions (Table 2).

#### Patients with impaired renal function and atrial arrhythmias are correctly attributed to higher scores in each staging system

As expected, all cardiac biomarkers were significantly elevated in the subgroups with eGFR<50 and AF, while the difference between involved and uninvolved FLC (dFLC) was only slightly elevated among the patients with AF and not significantly elevated among those with eGFR<50

(Table 1). Accordingly, eGFR showed a highly significant negative correlation with all cardiac biomarkers (Online Supplementary Table S3). Serum NT-proBNP levels were significantly determined by both heart involvement status and renal function (Online Supplementary Figure S3). As renal function decreased, more patients both with and without heart involvement exceeded the NT-proBNP thresholds of any of the staging systems. As a consequence, patients with eGFR<50 and with AF were attributed higher scores than their respective controls in all the staging systems analyzed (Table 1, Figure 2B,C). As shown

above, eGFR<50 or AF are by themselves prognostically adverse factors. Thereafter we addressed the question of whether, within the high-risk cardiac biomarker categories, the patients with eGFR<50 or AF actually fared better, worse or equally (Online Supplementary Figure S6). Apparently both described effects counterbalance each other in the MAYO3b staging system. However, patients with eGFR<50 classified as highest risk by the MAYO2004 and MAYO2012 staging systems still had a worse prognosis than patients with eGFR ≥50 mL/min/1.73 m<sup>2</sup> (6.4 vs. 14.9 months and 4.7 vs. 9.6 months, both P<0.001).

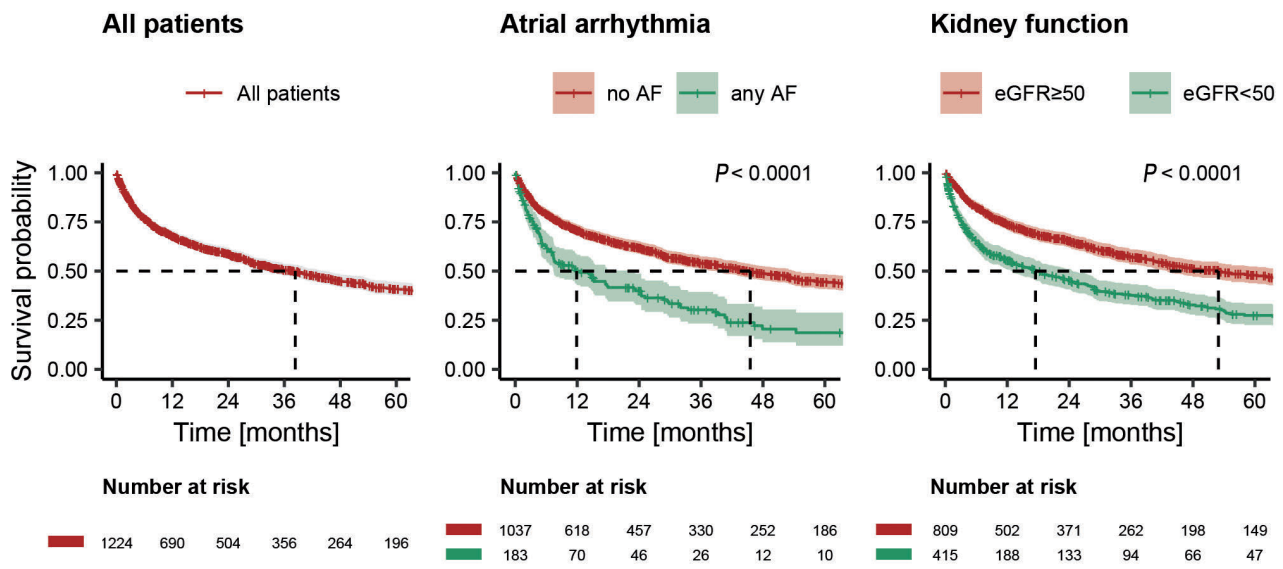


Figure 1. Survival of the whole cohort and the subgroups divided according to renal function and the presence or absence of atrial arrhythmias. Kaplan-Meier plots depicting overall survival (OS) from diagnosis. The shaded areas represent the 95% confidence interval estimates. The median OS (with corresponding 95% confidence intervals) were as follows: all patients, 38.3 (31.9 - 44.5) months; patients without an atrial arrhythmia or pacemaker rhythm (AF), 45.5 (38.9 - 54.2) months; patients with AF, 11.9 (7.4 - 23.0) months; patients with an estimated glomerular filtration rate (eGFR) ≥50 mL/min/1.73 m<sup>2</sup>, 52.9 (43.0 - 68.8) months, and patients with an eGFR <50 mL/min/1.73 m<sup>2</sup>, 18.3 (12.9 - 24.7) months.

Table 2. Results of univariable and multivariable analyses including different cardiac scoring systems.

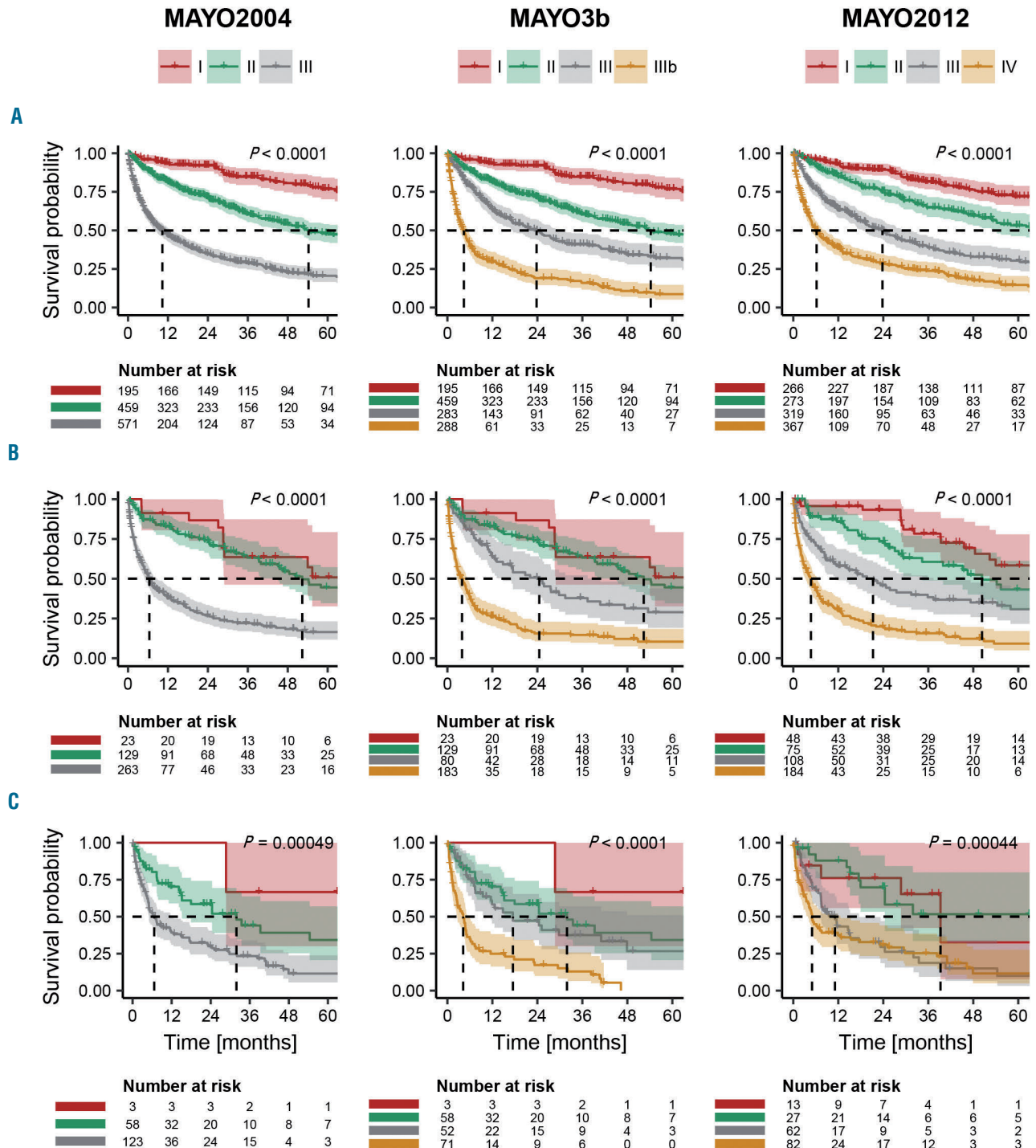
Prognostic factor	Comparison	HR	Univariable			MAYO2004			Multivariable MAYO3b			MAYO2012		
			95% CI	P	HR	95% CI	P	HR	95% CI	P	HR	95% CI	P	
Age, years	per 10 years	1.27	1.17 - 1.38	<b>&lt;0.001</b>	1.18	1.09 - 1.29	<b>&lt;0.001</b>	1.15	1.05 - 1.25	<b>0.002</b>	1.16	1.07 - 1.27	<b>&lt;0.001</b>	
Sex	male vs. female	1.13	0.96 - 1.33	0.133	-	-	-	-	-	-	-	-	-	
Light chain type	lambda vs. kappa	0.96	0.80 - 1.15	0.654	-	-	-	-	-	-	-	-	-	
dFLC, mg/L	<180 vs. ≥180	2.10	1.79 - 2.47	<b>&lt;0.001*</b>	1.60	1.35 - 1.89	<b>&lt;0.001</b>	1.53	1.29 - 1.81	<b>&lt;0.001</b>	-	-	-	
eGFR, mL/min/1.73m <sup>2</sup>	<50 vs. ≥50	1.85	1.58 - 2.17	<b>&lt;0.001</b>	1.33	1.13 - 1.57	<b>&lt;0.001</b>	1.15	0.97 - 1.37	0.105	1.32	1.12 - 1.56	<b>&lt;0.001</b>	
Atrial arrhythmia	any vs. no	1.96	1.61 - 2.40	<b>&lt;0.001</b>	1.28	1.04 - 1.57	<b>0.019</b>	1.24	1.01 - 1.53	<b>0.041</b>	1.24	1.00 - 1.52	<b>0.045</b>	
MAYO 2004 stage II	stage I	2.34	1.71 - 3.19	<b>&lt;0.001</b>	1.92	1.40 - 2.64	<b>&lt;0.001</b>	-	-	-	-	-	-	
MAYO 2004 stage III	stage I	6.36	4.71 - 8.58	<b>&lt;0.001</b>	4.55	3.31 - 6.24	<b>&lt;0.001</b>	-	-	-	-	-	-	
MAYO 3b stage II	stage I	2.36	1.73 - 3.22	<b>&lt;0.001</b>	-	-	-	2.05	1.49 - 2.81	<b>&lt;0.001</b>	-	-	-	
MAYO 3b stage IIIa	stage I	4.19	3.04 - 5.77	<b>&lt;0.001</b>	-	-	-	3.39	2.43 - 4.73	<b>&lt;0.001</b>	-	-	-	
MAYO 3b stage IIIb	stage I	11.26	8.22 - 15.42	<b>&lt;0.001</b>	-	-	-	8.10	5.75 - 11.40	<b>&lt;0.001</b>	-	-	-	
MAYO 2012 stage II	stage I	1.82	1.35 - 2.44	<b>&lt;0.001</b>	-	-	-	-	-	-	1.77	1.32 - 2.38	<b>&lt;0.001</b>	
MAYO 2012 stage III	stage I	3.74	2.84 - 4.92	<b>&lt;0.001</b>	-	-	-	-	-	-	3.44	2.61 - 4.55	<b>&lt;0.001</b>	
MAYO 2012 stage IV	stage I	7.08	5.44 - 9.21	<b>&lt;0.001</b>	-	-	-	-	-	-	6.26	4.77 - 8.21	<b>&lt;0.001</b>	

Summary of one univariable and three multivariable survival analyses using Cox proportional hazards regression. Statistically significant differences are represented in bold, italics. \*The threshold of 180 mg/L for the difference between involved and uninvolved free light chains (dFLC) was chosen to allow for comparison with the MAYO2012 system, which already includes dFLC at this threshold. HR: hazard ratio; 95% CI: 95% confidence interval; eGFR: estimated glomerular filtration rate.

**Impaired renal function and atrial arrhythmias add prognostic information to the staging systems**

To further evaluate whether, and if so to what extent, eGFR<50 and AF add prognostic information to the different staging systems, we performed multivariable survival analyses for each staging system. Age as well as dFLC (except in conjunction with the MAYO2012 stag-

ing system) were also included, as these factors had been found to be highly prognostic in the univariable analyses (Table 2). While eGFR<50 retained its prognostic value at the significance level of 5% in multivariable analyses with the MAYO2004 and MAYO2012 staging systems, it was not significant in multivariable analysis with the MAYO3b system, suggesting an improved prognostica-



**Figure 2. Overall survival according to the staging systems.** Kaplan-Meier-Plots, pairwise P-values and median overall survival are given in *Online Supplementary Table S5*. Shaded areas indicate the 95% confidence intervals. (A) All patients. (B) Subgroup of patients with an estimated glomerular filtration rate <50 mL/min/1.73 m<sup>2</sup>. (C) Subgroup of patients with atrial arrhythmia or pacemaker rhythm.

tion of the MAYO3b staging system compared to that of the other two models. However, AF retained its significant prognostic value in each condition.

### All evaluated staging systems provide a high prognostic value

Using Kaplan-Meier estimates, we could illustrate the ability of each analyzed staging system to identify subgroups of patients with distinct risks. There was virtually no overlap of the 95% confidence intervals of the different stages for each classification in the entire cohort (Figure 2A). With respect to the spread of the median overall survival, it seems that the MAYO3b system is superior in detecting the patients with the best and worse prognoses in the entire cohort and in each subgroup.

In the next step, we evaluated the overlap and respective prognosis of patients within each stage of the staging systems. The results of this analysis are shown in Table 3, in which proportions and median overall survival of patients are reported according to each stage of the MAYO2004/3b staging systems and subgrouped further by the MAYO2012 staging system. Both the MAYO3b and MAYO2012 staging systems were able to re-divide stages of the other system, identifying patients with better or worse outcomes. For example, 47% of the patients in MAYO2012 stage I were attributed to MAYO2004/3b stage II and 28% of the patients in MAYO2004/3b stage I were attributed to MAYO2012 stage II. In both cases, the higher stages were associated with an approximately 34 months shorter overall survival (Table 3).

To further evaluate and compare the prognostic value of the different staging systems, we computed time-dependent prediction errors, illustrating the overall model performance and time-dependent concordance indices, which mirror the systems' discriminative accuracy (Figure 3). All staging systems showed a high performance value (Figure 3A). However, the MAYO3b and MAYO2012 systems were superior to the MAYO2004 system with respect to discriminative accuracy in this analysis.

### The overall performance of each staging system is almost entirely preserved in patients with impaired renal function

In the subgroup of patients with eGFR<50, all of the analyzed staging systems retained highly significant global *P* values (Figure 2B). There was, however, only a poor discrimination of patients in stages I and II for all systems, as indicated by not significantly different median overall survivals (Figure 2B, *Online Supplementary Table S5*). This did not translate into a marked reduction of the predictive value according to prediction error curves, as well as con-

cordance indices, which was comparable to that of the entire cohort for each staging system (Figure 3B).

### The performance of all staging systems is reduced in patients with atrial arrhythmias, but best preserved in the MAYO3b system.

In the subgroup of patients with AF, the analyzed staging systems retained statistically significant global *P*-values, while there were overlaps between some stages (Figure 2C). Prediction error curves and concordance index plots did, however, confirm a reduced performance of all staging systems when compared to the performance in the entire cohort (Figure 3C). The MAYO2012 score appeared to be more affected than MAYO2004 and MAYO3b, especially with respect to discriminative accuracy for prediction beyond 18 months. Thus, the MAYO3b system clearly outperforms the MAYO2012 system in the subgroup of patients with AF.

## Discussion

This large retrospective study was aimed to: (i) evaluate the influence of two conditions with high prevalence among patients with AL amyloidosis patients, impaired renal function (eGFR<50 mL/min/1.73 m<sup>2</sup>, "eGFR<50") as well as atrial arrhythmia or pacemaker rhythm ("AF"), on the applicability of the most widely used AL amyloidosis cardiac staging systems (MAYO2004, MAYO3b and MAYO2012); (ii) evaluate the additional prognostic value of eGFR<50 and AF in the context of these staging systems; and (iii) compare the overall performance of these staging systems in the entire cohort and in the subgroups with eGFR<50 and AF.

(I) Patients with eGFR<50 or AF had a dismal prognosis; the risk of death was almost doubled by either comorbidity. The underlying reason seems to be that both conditions are surrogate parameters for heart involvement and accompanied by other unfavorable characteristics such as older age, lower Performance Status and, consequently, reduced eligibility for chemotherapy (Table 1). Interestingly, prerenal kidney injury due to heart involvement appears to be one of the most common reasons for reduced eGFR (*Online Supplementary Results, Online Supplementary Figure S5B*) and the survival of patients with decreased kidney function is mainly determined by heart involvement status and not negatively influenced by a higher degree of kidney organ involvement (*Online Supplementary Results, Online Supplementary Figure S6A*).

We verified in our cohort of AL patients that impaired renal function as well as AF are associated with increased

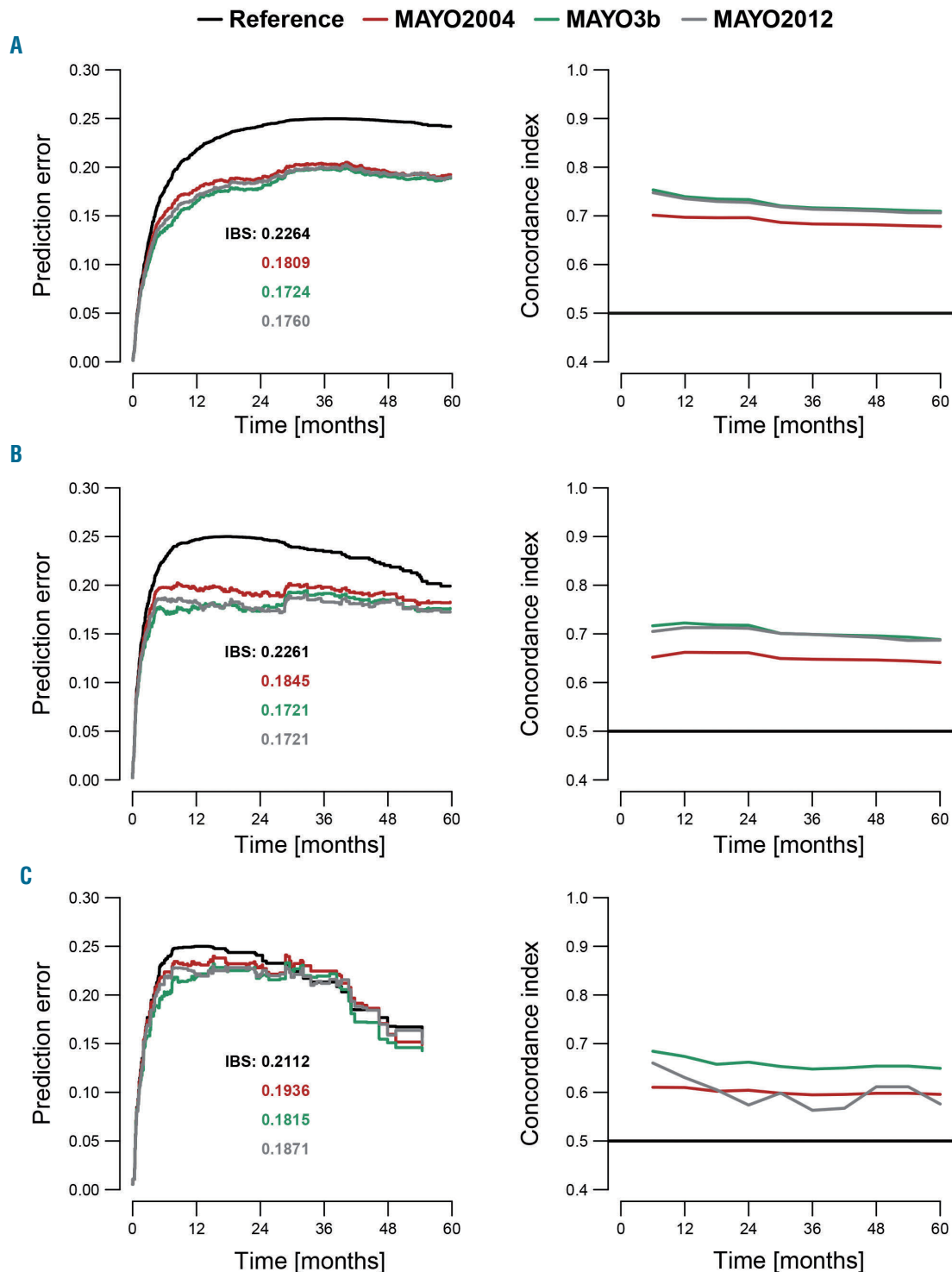
**Table 3.** Comparison of the MAYO2004/3b and MAYO2012 staging systems.

	MAYO2012 stage I			MAYO2012 stage II			MAYO2012 stage III			MAYO2012 stage IV		
	N.	Median OS	95% CI	N.	Median OS	95% CI	N.	Median OS	95% CI	N.	Median OS	95% CI
MAYO2004/3b stage I	141	131.7	130 - NR	54	97.8	62 - NR	-	-	-	-	-	-
MAYO2004/3b stage II	125	97.5	70 - NR	191	54.2	45 - 88	135	33	27 - 54	7	15.1	8 - NR
MAYO2004 stage III	-	-	-	27	72.1	25 - NR	183	17.1	14 - 24	360	5.9	5 - 8
MAYO3b stage IIIa	-	-	-	27	72.1	25 - NR	112	24.4	17 - 46	143	17.5	11 - 29
MAYO3b stage IIIb	-	-	-	-	-	-	71	7.7	6 - 18	217	3.4	3 - 5

N: number of patients; OS: overall survival; 95% CI: 95% confidence interval; NR: not reached.

serum levels of troponins and in particular NT-proBNP.<sup>24-26,28,33</sup> We were also able to show that, in contrast, dFLC, which is applied by the MAYO2012 system, is only slightly elevated in patients with eGFR<50 or AF. The elevation of cardiac biomarkers translated into patients in both sub-

groups being assigned to higher stages of all the analyzed scoring systems; as a consequence, only a few patients with either of these comorbidities were assigned stage I in any scoring system. However, the patients assigned to the respective highest stage in the eGFR<50 and AF subgroups



**Figure 3. Performance of the staging systems.** Curves of time-dependent prediction errors as well as time-dependent concordance indices of the scoring systems. The Integrated Brier Score (IBS) for each staging system is given with the respective color, as indicated by the legend. The reference is the marginal Kaplan-Meier estimator, ignoring the predictors. A concordance index of 0.5 indicates random chance. (A) All patients. The MAYO2012 and MAYO3b lines lie one above another. (B) Subgroup of patients with an estimated glomerular filtration rate <50 mL/min/1.73 m<sup>2</sup>. (C) Subgroup of patients with atrial arrhythmia or pacemaker rhythm.

did not have a favorable prognosis compared to that of patients assigned to the highest stage without these conditions. Indeed, patients with eGFR<50 classified as high-risk by the MAYO2004 and MAYO2012 staging systems had an even worse prognosis than that of patients with eGFR  $\geq$ 50 mL/min/1.73 m<sup>2</sup>, while there was no such difference in the MAYO3b system. This suggests that the higher stages due to the elevation of cardiac biomarkers are correctly assigned and provides evidence that the MAYO3b staging system is superior to the MAYO2004 and MAYO2012 systems in patients with impaired kidney function.

(II) Further evidence for a superiority of the MAYO3b system comes from the evaluation of the additional prognostic value of eGFR<50 and AF in the context of the staging systems. In multivariable analyses including age, dFLC and any of the staging systems, eGFR<50 and AF each retained their prognostic significance with the MAYO2004 and MAYO2012 systems, while eGFR<50 fell out with MAYO3b.

(III) We further evaluated the performance of the staging systems by synopsis of stratified Kaplan-Meier estimates (Figure 2) as well as predictive and discriminative accuracy using prediction errors and concordance indices (Figure 3). We confirmed the high prognostic value of each of the staging systems in our entire cohort. However, the MAYO3b and MAYO2012 systems showed a better discriminative accuracy than the MAYO2004 system, and the MAYO3b system was superior in detecting patients with the best and worst prognoses. In the subgroup with low eGFR, stages I and II within any of the staging systems could not discriminate patients with different survival. This did not, however, affect the overall predictive and discriminative accuracy of all staging systems, which again confirms that patients with eGFR<50 were correctly assigned to higher stages. In the AF subgroup, the systems were less precise, potentially also due to low numbers of patients assigned to the low-risk stages. The MAYO2012 system, in particular, lost discriminative accuracy, making the MAYO3b system superior in this subset. This might be because patients with very high risk of early death are well discriminated by stage IIIb, while in the MAYO2012 system the median survival of patients in the two highest stages is not significantly different. Hence, very high levels of NT-proBNP are prognostic in AL patients irrespective of eGFR or AF status. Herewith, we confirm recent findings made in other cohorts of patients with heart failure.<sup>26,34</sup>

Overall, our results suggest that the MAYO3b staging system should currently be preferred over the other wide-

ly used systems for the following reasons: (i) it showed the best performance and highest discriminative accuracy in the entire cohort; (ii) it was most applicable in the high-risk subsets of patients with eGFR<50 or AF; and (iii) it only requires two biomarkers compared to the three biomarkers required for the MAYO2012 score, which is an advantage from the perspective of model optimization.

A recently proposed staging system for cardiac transthyretin amyloidosis is based on NT-proBNP and eGFR.<sup>35</sup> Our results suggest, however, that the prognostic potential of reduced eGFR is already included in a staging system with multiple NT-proBNP thresholds such as MAYO3b. Interestingly, the superiority of the modified MAYO staging system with two NT-proBNP thresholds was also reported by the Mayo group itself at the annual International Symposium on Amyloidosis in March 2018 following calculation of the net reclassification improvement for pairs of prognostic models.<sup>36</sup>

On the other hand, our multivariable analysis suggests that the significant prognostic potential of atrial arrhythmia and dFLC is not entirely contained in the MAYO3b system. This is supported by the fact that the MAYO2012 system was able to identify patients with different outcomes from each stage of the MAYO3b system. In addition to AF and plasma cell-related factors, such as dFLC,<sup>21</sup> the inclusion of other promising and widely available biomarkers such as serum uric acid<sup>22</sup> might further improve the MAYO3b system. In contrast to novel cardiac biomarkers that have recently been suggested for inclusion into the AL amyloidosis staging system,<sup>19</sup> these parameters are readily available in every newly diagnosed AL patient worldwide. Finally, a potential modification of the MAYO3b staging system and validation of a novel staging system might be best performed with a multicenter approach, e.g. on behalf of the International Society of Amyloidosis, to achieve greatest validity and facilitate broad acceptance.

To our knowledge, this is the first report that extensively describes kidney impairment and atrial arrhythmia as independent prognostic factors in AL amyloidosis. Most importantly, the detailed performance analysis of the AL staging systems, including high-risk subgroups, will enable clinicians and investigators to make a confident choice within the broad range of currently available staging systems.

### Acknowledgments

The authors thank Rita Ziehl, medical data manager, for her assistance with the collection of clinical data.

### References

- Maurer MS, Elliott P, Comenzo R, Semigran M, Rapezzi C. Addressing common questions encountered in the diagnosis and management of cardiac amyloidosis. *Circulation*. 2017;135(14):1357–1377.
- Lisenko K, Schönland SO, Jauch A, et al. Flow cytometry-based characterization of underlying clonal B and plasma cells in patients with light chain amyloidosis. *Cancer Med*. 2016;5(7):1464–1472.
- Comenzo RL, Reece D, Palladini G, et al. Consensus guidelines for the conduct and reporting of clinical trials in systemic light-chain amyloidosis. *Leukemia*. 2012;26(11):2317–2325.
- Dittrich T, Bochtler T, Kimmich C, et al. AL amyloidosis patients with low amyloidogenic free light chain levels at first diagnosis have an excellent prognosis. *Blood*. 2017;130(5):632–642.
- Abraham RS, Katzmann JA, Clark RJ, Bradwell AR, Kyle RA, Gertz MA. Quantitative analysis of serum free light chains. A new marker for the diagnostic evaluation of primary systemic amyloidosis. *Am J Clin Pathol*. 2003;119(2):274–278.
- Milani P, Basset M, Russo F, Foli A, Merlini G, Palladini G. Patients with light-chain amyloidosis and low free light-chain burden have distinct clinical features and outcome. *Blood*. 2017;130(5):625–631.
- Kimmich C, Schönland S, Kräker S, et al. Amyloid in bone marrow smears in systemic light-chain amyloidosis. *Amyloid*. 2017;24(1):52–59.
- Sachchithanatham S, Hutt DF, Quigley AM, Hawkins P, Wechalekar AD. Role of 99mTc-DPD scintigraphy in imaging extra-cardiac light chain (AL) amyloidosis. *Br J Haematol*. 2018;183(3):506–509.
- Gertz MA, Kyle RA. Hepatic amyloidosis: clinical appraisal in 77 patients. *Hepatology*. 1997;25(1):118–121.



10. Merlini G, Lousada I, Ando Y, et al. Rationale, application and clinical qualification for NT-proBNP as a surrogate end point in pivotal clinical trials in patients with AL amyloidosis. *Leukemia*. 2016;30(10):1979–1986.
11. Dispenzieri A, Gertz MA, Kumar SK, et al. High sensitivity cardiac troponin T in patients with immunoglobulin light chain amyloidosis. *Heart Br Card Soc*. 2014;100(5): 383–388.
12. Palladini G, Hegenbart U, Milani P, et al. A staging system for renal outcome and early markers of renal response to chemotherapy in AL amyloidosis. *Blood*. 2014;124(15): 2325–2332.
13. Wechalekar AD, Schonland SO, Kastritis E, et al. A European collaborative study of treatment outcomes in 346 patients with cardiac stage III AL amyloidosis. *Blood*. 2013;121(17):3420–3427.
14. Mohty D, Petitalot V, Magne J, et al. Left atrial function in patients with light chain amyloidosis: a transthoracic 3D speckle tracking imaging study. *J Cardiol*. 2018;71(4):419–427.
15. Palladini G, Campana C, Klersy C, et al. Serum N-terminal pro-brain natriuretic peptide is a sensitive marker of myocardial dysfunction in AL amyloidosis. *Circulation*. 2003;107(19):2440–2445.
16. Dispenzieri A, Kyle RA, Gertz MA, et al. Survival in patients with primary systemic amyloidosis and raised serum cardiac troponins. *Lancet*. 2003;361(9371):1787–1789.
17. Dispenzieri A, Gertz MA, Kyle RA, et al. Serum cardiac troponins and N-terminal pro-brain natriuretic peptide: a staging system for primary systemic amyloidosis. *J Clin Oncol*. 2004;22(18):3751–3757.
18. Kristen AV, Giannitsis E, Lehrke S, et al. Assessment of disease severity and outcome in patients with systemic light-chain amyloidosis by the high-sensitivity troponin T assay. *Blood*. 2010;116(14):2455–2461.
19. Dispenzieri A, Gertz MA, Saenger A, et al. Soluble suppression of tumorigenicity 2 (sST2), but not galactin-3, adds to prognostication in patients with systemic AL amyloidosis independent of NT-proBNP and troponin T. *Am J Hematol*. 2015;90(6):524–528.
20. Palladini G, Barassi A, Klersy C, et al. The combination of high-sensitivity cardiac troponin T (hs-cTnT) at presentation and changes in N-terminal natriuretic peptide type B (NT-proBNP) after chemotherapy best predicts survival in AL amyloidosis. *Blood*. 2010;116(18):3426–3430.
21. Kumar S, Dispenzieri A, Lacy MQ, et al. Revised prognostic staging system for light chain amyloidosis incorporating cardiac biomarkers and serum free light chain measurements. *J Clin Oncol*. 2012;30(9):989–995.
22. Kumar S, Dispenzieri A, Lacy MQ, et al. Serum uric acid: novel prognostic factor in primary systemic amyloidosis. *Mayo Clin Proc*. 2008;83(3):297–303.
23. Kastritis E, Dimopoulos MA. Recent advances in the management of AL amyloidosis. *Br J Haematol*. 2016;172(2): 170–186.
24. Potter JM, Simpson AJ, Kerrigan J, et al. The relationship of plasma creatinine (as eGFR) and high-sensitivity cardiac troponin and NT-proBNP concentrations in a hospital and community outpatient population. *Clin Biochem*. 2017;50(15):813–815.
25. Chen M, Gerson H, Eintracht S, Nessim SJ, MacNamara E. Effect of hemodialysis on levels of high-sensitivity cardiac troponin T. *Am J Cardiol*. 2017;120(11):2061–2064.
26. Kristensen SL, Jhund PS, Mogensen UM, et al. Prognostic value of N-terminal pro-B-Type natriuretic peptide levels in heart failure patients with and without atrial fibrillation. *Circ Heart Fail*. 2017;10(10):e004409.
27. Thackray SDR, Witte K, Ghosh J, et al. N-terminal brain natriuretic peptide as a screening tool for heart failure in the pacemaker population. *Eur Heart J*. 2006;27(4): 447–453.
28. deFilippi CR, Herzog CA. Interpreting cardiac biomarkers in the setting of chronic kidney disease. *Clin Chem*. 2017; 63(1):59–65.
29. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.
30. Grambsch PM, Therneau TM. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*. 1994;81(3): 515–526.
31. Gerds TA, Schumacher M. Consistent estimation of the expected Brier score in general survival models with right-censored event times. *Biom J Biom. Z* 2006;48(6):1029–1040.
32. Gerds TA, Kattan MW, Schumacher M, Yu C. Estimating a time-dependent concordance index for survival prediction models with covariate dependent censoring. *Stat Med*. 2013;32(13):2173–2184.
33. Maisel A, Mueller C, Adams K, et al. State of the art: using natriuretic peptide levels in clinical practice. *Eur J Heart Fail*. 2008;10(9): 824–839.
34. Manzano-Fernández S, Januzzi JL, Boronat-García M, et al. Impact of kidney dysfunction on plasma and urinary N-terminal pro-B-type natriuretic peptide in patients with acute heart failure. *Congest Heart Fail*. 2010;16(5):214–220.
35. Gillmore JD, Damy T, Fontana M, et al. A new staging system for cardiac transthyretin amyloidosis. *Eur Heart J*. 2018;39(30):2799–2806.
36. Muchtar E, Therneau TM, Gertz MA, et al. Can we pick a winner in immunoglobulin light chain amyloidosis staging systems? ISA. Abstract 2018;PC073.