Network meta-analysis of randomized trials in multiple myeloma: efficacy and safety in relapsed/refractory patients

Cirino Botta,[1,]* Domenico Ciliberto,[1,]* Marco Rossi,[1] Nicoletta Staropoli,[1] Maria Cucè,[1] Teresa Galeano,[1] Pierosandro Tagliaferri,[1] and Pierfrancesco Tassone[1,2]

[1]Department of Experimental and Clinical Medicine, Magna Graecia University, Salvatore Venuta University Campus, Catanzaro, Italy; and [2]Sbarro Institute for Cancer Research and Molecular Medicine, Center for Biotechnology, College of Science and Technology, Temple University, Philadelphia, PA

Key Points

- Different therapeutic agents are currently available for the treatment of RRMM.

- By performing an NMA, we identified a lenalidomide-dexamethasone + mAb regimen as the most active therapeutic option in this setting.

Despite major therapeutic advancements, multiple myeloma (MM) is still incurable and relapsed/refractory multiple myeloma (RRMM) remains a challenge; the rational choice of the most appropriate regimen in this setting is currently undefined. We performed a systematic review and 2 standard pairwise meta-analyses to evaluate the efficacy of regimens that have been directly compared with bortezomib or immunomodulatory imide drugs (IMiDs) in head-to-head clinical trials and a network meta-analysis (NMA) to determine the relevance of each regimen on the basis of all the available direct and indirect evidence. Sixteen trials were included in the pairwise meta-analyses, and 18 trials were included in the NMA. Pairwise meta-analyses showed that a 3-drug regimen (bortezomib- or IMiD-based) was superior to a 2-drug regimen in progression-free-survival (PFS) and overall response rate (ORR). NMA showed that an IMiD backbone associated with anti-MM monoclonal antibodies (mAbs) (preferably) or proteasome inhibitors had the highest probability of being the most effective regimen with the lowest toxicity. The combination of daratumumab, lenalidomide, and dexamethasone ranked as the first regimen in terms of activity, efficacy, and tolerability according to the average value between surface under the cumulative ranking curve of PFS, overall survival, ORR, complete response rate, and safety. This is the first NMA comparing all currently available regimens evaluated in published randomized trials for the treatment of RRMM, but our results need to be interpreted taking into account differences in their patient populations. Our analysis suggests that IMiDs plus new anti-MM mAb–containing regimens are the most active therapeutic option in RRMM.

Introduction

Multiple myeloma (MM) is among the most common hematologic malignancies. Current milestones of MM therapy include either a triple- or double-drug combination, based on proteasome inhibitors (PIs) and/or immunomodulatory imide drugs (IMiDs) plus dexamethasone with or without chemotherapy. Eligible patients further undergo autologous stem cell transplantation and consolidation therapy,[1,2] and patients not eligible for autologous stem cell transplantation enter follow-up.[3,4] However, virtually all patients still relapse and require further treatment. Thus, relapsed/refractory MM (RRMM) currently represents the main focus of intensive clinical research. Indeed, a plethora of new agents, including second-generation PIs, histone deacetylase (HDAC) inhibitors, and monoclonal antibodies (mAbs), have shown consistent activity in prospective phase 2/3 clinical trials.[5,6] On the basis of these premises, we undertook a systematic review and 2 pairwise meta-analyses of phase 2/3 randomized trials in the

The full-text version of this article contains a data supplement.

RRMM setting to evaluate the impact and safety of new therapeutic strategies compared with bortezomib with or without dexamethasone or IMiDs + dexamethasone regimens. Because each trial usually compares only 2 regimens, it was challenging to evaluate the relative efficacy of all the regimens we investigated. To overcome these limitations, we further performed a network meta-analysis (NMA), which is a recently introduced Bayesian statistical approach that allows combining direct and indirect evidence to rank the different treatments according to their efficacy and safety.

Methods

Search strategy

An electronic search for relevant publications was performed by using PubMed, Embase, Ovid, the Cochrane collaboration database, and proceedings from the major international meetings in hematology and oncology. Only prospective studies were allowed in this analysis.[7] The following search headings were used: "multiple myeloma", "relapse", "refractory", "randomized clinical trial", "management", "bortezomib", "lenalidomide", "thalidomide", and "therapy" in different combinations (eg, "relapsed/refractory multiple myeloma"). All titles were screened and abstracts were reviewed. The related articles function, article references, and Google Scholar were also screened for other applicable publications and were used for searching related studies, abstracts, and citations. Only articles in English were considered. The last date of the search was June 24, 2016. A systematic review was performed according to the guidelines and recommendations from the preferred reporting items for systematic reviews and meta-analyses (PRISMA) checklist.[8,9]

Inclusion criteria

To be included in the analysis, the studies had to meet several criteria. They had to involve only patients with a diagnosis of RRMM (including patients in second-line treatment), be randomized controlled trials (RCTs) with or without blinding, be abstracts or unpublished data (to be included only if they had sufficient information on study design, characteristics of participants, interventions, and outcomes), have patients who received an unconventional or new regimen in the experimental arm, and have patients who received a standard regimen for RRMM in the control arm.

Exclusion criteria

Studies were excluded from the analysis if they were not comparative or not prospective, if outcomes of interest were not reported, if the methodology was not clearly reported, and if only patients after the second line of treatment were included.

Data extraction and quality assessment

Two reviewers (C.B. and N.S.) independently reviewed the literature according to the above predefined strategy and criteria. Each reviewer extracted the following data: title and reference details (first author, year), study population characteristics (number of patients in study, number treated by each approach, number of treatment lines, previous exposure to bortezomib or IMiDs), type of interventions, and outcome data. For each trial, we evaluated hazard ratios (HRs) of progression-free survival (PFS) and overall survival (OS); odds ratio (OR) of overall response rate (ORR), very good partial response rate (VGPR), and complete response (CR); and risk ratio (RR) for safety (evaluation of common toxicities). If the HR of survival curves was not reported,[10,11] it was derived from the graph by using the method of Tierney et al[12] (see supplemental Data).

All data were recorded independently in separate databases and were compared at the end of the reviewing process to limit selection bias. The database was also reviewed by 2 additional investigators (D.C. and M.C.). Duplicates were removed and any disparities were clarified.

Selected studies were assessed for quality according to the *Cochrane Handbook for Systematic Reviews of Interventions*, as described elsewhere,[13] by assigning 1 point for each of the following 5 requirements:

method of randomization, allocation concealment, blindness, withdrawal or dropout, and adequacy of follow-up. Studies were ranked as A if they had 4 to 5 points, B if they had 2 to 3 points, or C if they had 0 to 1 point of a total of 5 points.[14] The presence of publication bias was investigated by using Egger's test[15] or Begg's test[16] and by visual inspection of funnel plots.

Pairwise meta-analyses

Traditional pairwise meta-analyses were carried out as described elsewhere.[17-19] The following comparisons were used: experimental therapy vs conventional bortezomib-based therapy or experimental therapy vs conventional IMiD-based therapy (additional details are provided in the supplemental Data). Survival data were extracted as HRs of OS and PFS with relative confidence intervals (95% CIs); response rates and toxicity rate were calculated by using the method for dichotomous data (OR and RR assessment with 95% CIs). Cochrane's Q test and I^2 statistics were used to assess heterogeneity between studies, and both fixed-effects (FE) and random-effects (RE) models were used for the analysis. Pooled data analysis was performed according to the DerSimonian and Laird test. $P < .05$ was considered statistically significant. Statistical analysis was conducted by using STATA software (Version 14.1, STATA, College Station, TX) and R (meta package).

NMA

We performed an NMA by using both Bayesian and frequentist approaches to compare the different therapeutic regimens simultaneously. The Bayesian NMA was performed with STATA software by using the mvmeta package, and the frequentist NMA was performed with R software by using the netmeta package.[20] NMA synthesizes data from a network of trials that involve multiple interventions and therefore has the potential to rank the treatments according to the outcome. This method integrates direct and indirect comparisons. Within the framework of NMA, we ranked the evaluated regimens based on survival outcomes (OS and PFS), treatment efficacy (ORR and CR), and safety (the most frequent grade 3-4 adverse event in each trial). For each outcome, we performed a Bayesian NMA with an (RE) model by using a Markov chain Monte Carlo simulation technique with up to 30 000 iterations. Loop inconsistency and heterogeneity were assessed by evaluating the logarithm of the ratio of 2 odds ratios (RoR) from direct and indirect evidence in the loop with the ifplot command in STATA.[21,22] RoR values close to 0 indicate that both direct and indirect evidence are in agreement. Heterogeneity of the loop was then assessed through the restricted maximum likelihood method.[21,22]

Relative effects of treatments are reported as HRs for survival outcomes (PFS and OS) and OR or RR for binary outcomes (ORR, CR, and toxicity) along with corresponding 95% credible intervals, the Bayesian equivalent of 95% CIs. Ranking probabilities and surface under the cumulative ranking curve (SUCRA), which provides a numerical summary of the rank distribution of each treatment schedule on the different end points, were used to provide hierarchy probabilities.[21-23] The larger the SUCRA value (ie, closer to 1), the better the rank of the intervention. The frequentist equivalent of the SUCRA, the P-score, was further used to confirm the results of the NMA.[20]

Results

Study selection

Figure 1 shows a PRISMA chart of RCT selection and search strategy. In the time frame covered by this systematic review, a total of 1679 studies, including full articles or meeting abstracts, were retrieved. Nineteen trials with a total of 8997 patients were finally selected and included in the analysis (Table 1).[10,11,24-42] Two trials[41,42] did not report data or provide a graph showing OS and were therefore excluded from all OS analyses. At least 1 data comparison in terms of survival, ORR, or toxicity was reported in all selected RCTs, which were therefore deemed eligible for the

Figure 1. PRISMA flowchart for study selection and review. MA, meta-analysis; PRISMA, preferred reporting items for systematic reviews and meta-analyses (checklist); SEER, Surveillance, Epidemiology, and End Results (program).
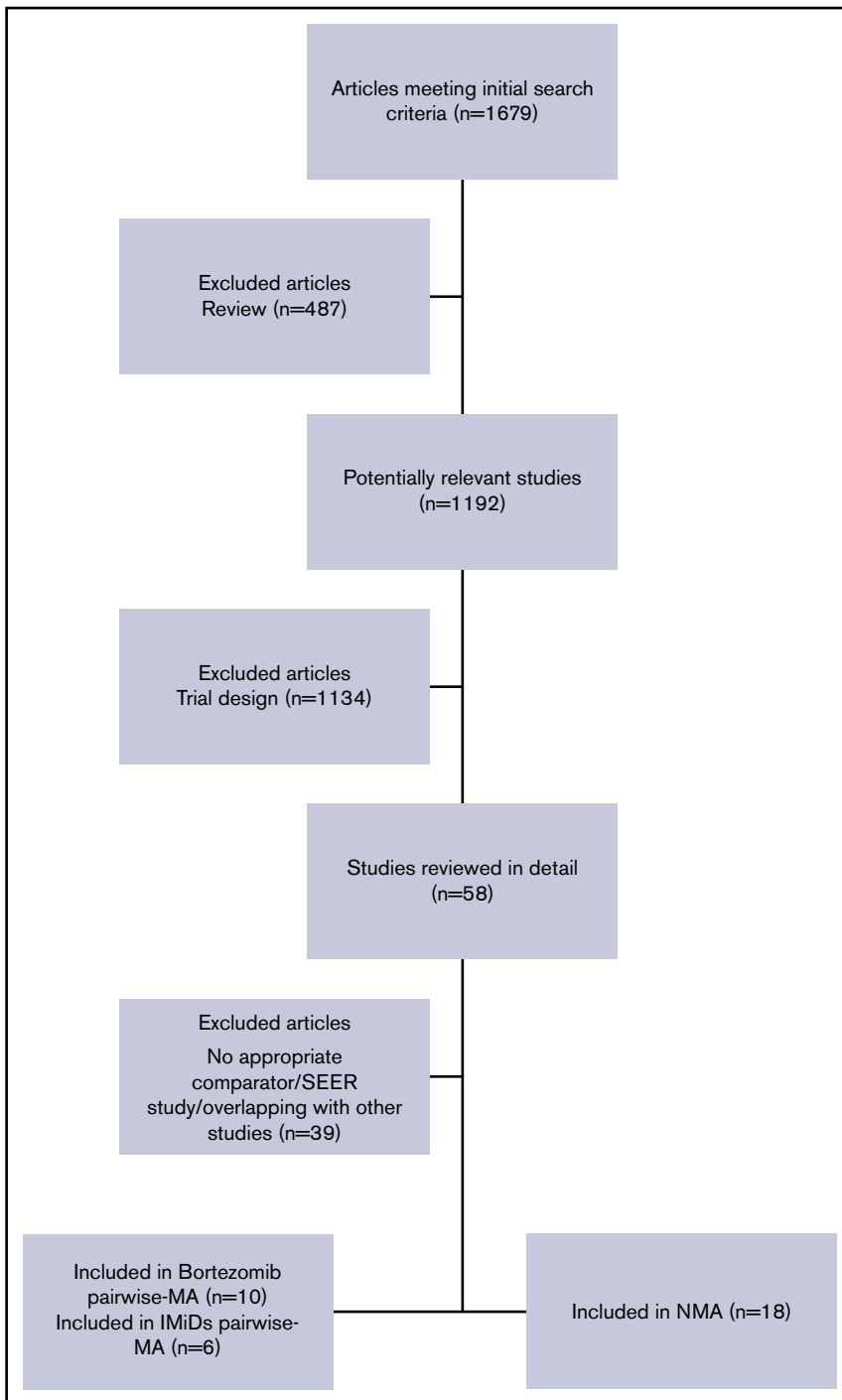
analysis. A summary of the 19 trials included in the final analyses showed that 17 trials were eligible for OS analysis; 19 were eligible for PFS, ORR, and safety analyses; 16 were included in the pairwise meta-analyses; and 18 were included in the NMA.

Quality assessment

As depicted in Table 2, 15 trials were scored as A (low risk of bias), and 4 trials were scored as B (intermediate risk of bias). None of the trials evaluated was scored as C (high risk of bias).

Pairwise meta-analyses

Our analyses included 16 trials for pairwise meta-analyses (involving 7500 patients). In particular, there were 10 trials (4371 patients)[10,11,33-37,39-41] in which an experimental treatment was compared with a bortezomib-based conventional treatment and 6 trials (3129 patients)[11,24-27,42] in which an experimental treatment was compared with an IMiD-based conventional treatment. Funnel plot visual inspection (supplemental Figure 1), Egger's test ($t = 0.57$; $P = .58$), and Begg's test ($z = 0.80$; $P = .42$) showed no evidence

Table 1. Characteristics of the studies included in the meta-analyses

Reference	Trial name	Year	Phase	No. of patients	Drug combinations in arm 1	Drug combinations in arm 2	Patients with 1 previous treatment (%)	Patients with 2 or more previous treatments (%)
40	VANTAGE 088	2013	3	637	VOR/BOR	BOR	42	58
39	DOXIL-MMY-3001	2007	3	646	PLD/BOR	BOR	34	66
10	NCT00401843	2015	2	281	SIL/BOR	BOR	52	48
41	MMY 3021	2011	3	222	subBOR	BOR	63	37
37	AMBER	2012	2	102	BOR + BEV	BOR	48	52
36	ENDEAVOR	2015	3	929	CAR/DEX	BOR/DEX	50	50
35	PANORAMA 1	2014	3	768	PAN/BOR/DEX	BOR/DEX	51	49
34	CASTOR	2016	3	498	DAR/BOR/DEX	BOR/DEX	47	53
33	NCT01478048	2016	2	152	ELO/BOR/DEX	BOR/DEX	66	34
11	NCT00602511	2012	3	131	THA/DEX	BOR/DEX	NA	NA
32	MM-010	2007	3	351	LEN/DEX	DEX	32	68
31	MM-009	2007	3	353	LEN/DEX	DEX	38	62
29	APEX	2009	3	669	BOR	DEX	38	62
28	OPTIMUM	2012	3	499	THA	DEX	57	43
27	ASPIRE	2014	3	792	CAR/LEN/DEX	LEN/DEX	43	57
26	ELOQUENT-2	2015	3	646	ELO/LEN/DEX	LEN/DEX	48	52
42	TOURMALINE-MM1	2015	3	722	IXA/LEN/DEX	LEN/DEX	61	39
25	POLLUX	2016	3	569	DAR/LEN/DEX	LEN/DEX	52	48
24	MMVAR/IFM 2005-04	2012	3	269	BOR/THA/DEX	THA/DEX	53	47

BEV, bevacizumab; BOR, bortezomib; CAR, carfilzomib; DAR, daratumumab; DEX, dexamethasone; ELO, elotuzumab; IXA, ixazomib; LEN, lenalidomide; NA, not applicable; PAN, panobinostat; PLD, pegylated liposomal doxorubicin; SIL, siltuximab; subBOR, subcutaneous bortezomib; THA, thalidomide; VOR, vorinostat.

of publication bias for bortezomib comparison. For IMiD comparison (supplemental Figure 1), only funnel plot visual inspection was performed because of the low number of studies (6) included in the analysis, again suggesting no evidence for publication bias. All these tests should be considered to have low power because of the small number of studies involved in our analysis.

Table 2. Quality assessment of studies included in the meta-analysis

Reference	Trial name	Year	Method of randomization	Allocation concealment	Blinded	Withdrawal and dropout	Baseline	Quality level*
40	VANTAGE 008	2013	Centralized	Central office	Yes	Detailed criteria	Identical baseline	A
39	DOXIL-MMY-3001	2007	Centralized	Central office	No	Detailed criteria	Identical baseline	A
10	NCT00401843	2015	Centralized	Central office	Yes	Detailed criteria	Identical baseline	A
41	MMY 3021	2011	Centralized	Central office	No	Detailed criteria	Identical baseline	A
37	AMBER	2012	Centralized	Central office	Yes	Detailed criteria	Identical baseline	A
36	ENDEAVOR	2015	Centralized	Central office	No	Detailed criteria	Identical baseline	A
35	PANORAMA 1	2014	Centralized	Central office	Yes	Detailed criteria	Identical baseline	A
34	CASTOR	2016	Centralized	Central office	No	Detailed criteria	Identical baseline	A
33	NCT01478048	2016	Centralized	No details	No	Detailed criteria	Identical baseline	B
11	NCT00602511	2012	Centralized	Central office	No	Detailed criteria	Identical baseline	A
32	MM-010	2007	Centralized	Central office	Yes	Detailed criteria	Identical baseline	A
31	MM-009	2007	Centralized	Central office	Yes	Detailed criteria	Identical baseline	A
29	APEX	2009	No details	No details	No	Detailed criteria	Identical baseline	B
28	OPTIMUM	2012	No details	No details	No	Detailed criteria	Identical baseline	B
27	ASPIRE	2014	Centralized	Central office	No	Detailed criteria	Identical baseline	A
26	ELOQUENT-2	2015	Centralized	Central office	No	Detailed criteria	Identical baseline	A
42	TOURMALINE-MM1	2015	Centralized	Central office	Yes	Detailed criteria	Identical baseline	A
25	POLLUX	2016	Centralized	Central office	No	Detailed criteria	Identical baseline	A
24	MMVAR/IFM 2005-04	2012	No details	No details	No	Detailed criteria	Identical baseline	B

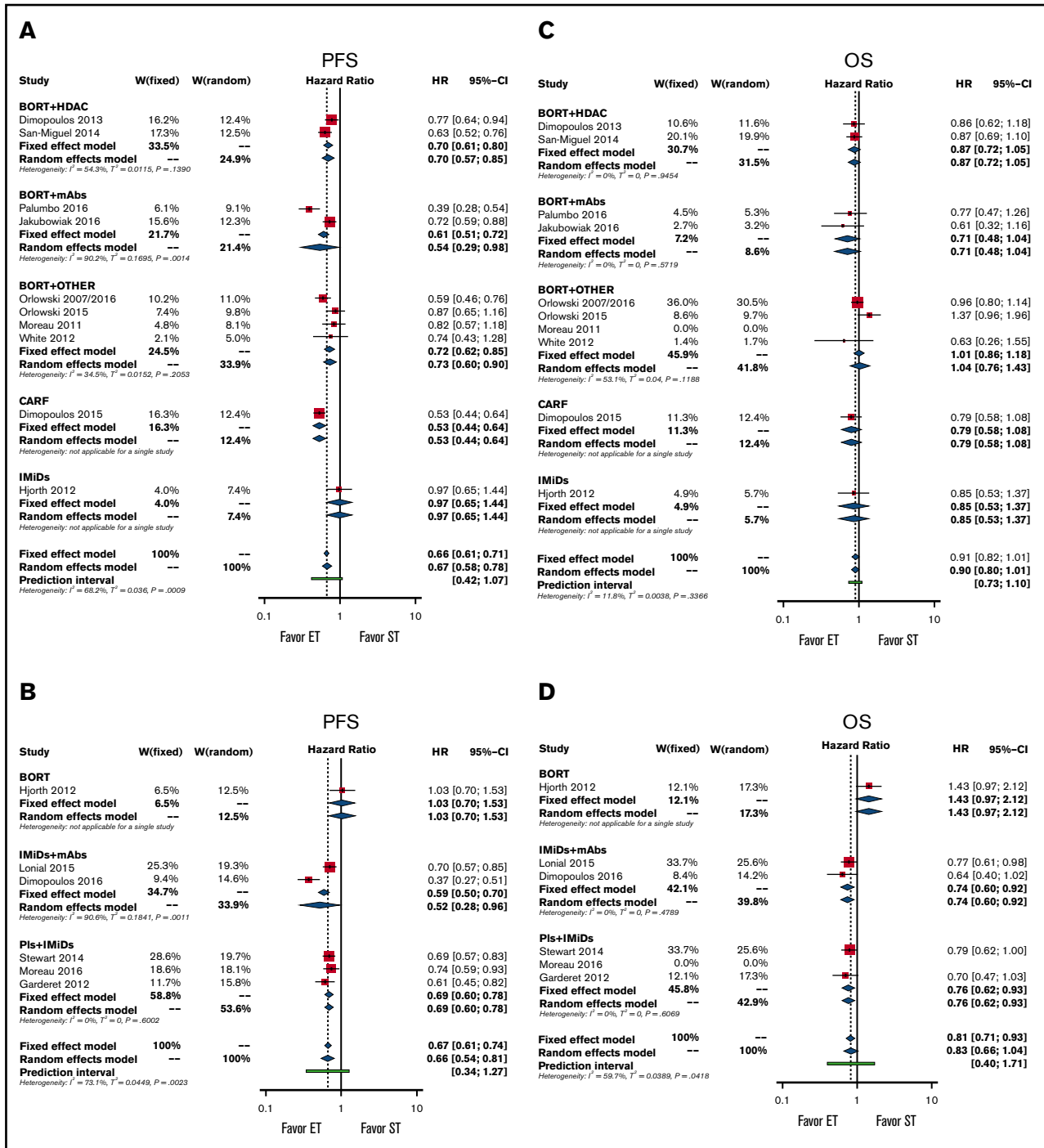*A, quality score of 4 to 5 of a total score of 5; B, quality score of 2 to 3 of a total score of 5.

Figure 2. Forest plots of comparisons between experimental treatments and standard treatments in terms of PFS and OS. Bortezomib (BORT) with or without dexamethasone represents the standard treatment (ST) in (A) PFS and (C) OS; IMiDs represent the standard treatment in (B) PFS and (D) OS. Subgroups have been created according to drug classes. CARF, carfilzomib; ET, experimental treatment; W, weight.

PFS analyses

Comparison of bortezomib-based regimens. The experimental treatment, when compared with bortezomib-based regimens, significantly improved PFS by using either an FE or RE model (pooled HRs for RE, 0.67; 95% CI, 0.58-0.78; *P* < .001; Figure 2A). By using subgroup analysis, we demonstrated significant PFS benefit in

doublet therapy subgroups and in the subgroup for therapy with all classes of drugs except IMiDs (pooled HRs, 0.97; 95% CI, 0.65-1.44; *P* = .880; Figure 2A; supplemental Figure 2A).

Comparison of IMiD-based regimens. The experimental treatment, when compared with IMiD-based regimens, significantly improved PFS (pooled HRs for RE, 0.66; 95% CI, 0.54-0.81; *P* < .001;

Table 3. Most common grade 3 to 4 adverse events analyzed in the pairwise meta-analyses

Adverse event subgroup	No. of patients with grade 3-4 adverse events		Overall RR	95% CI	P
	ET	CT			
Anemia					
BOR	280	240	1.124	0.946-1.335	.184
IMiD	208	252	0.843	0.659-1.079	.175
Thrombocytopenia					
BOR	734	482	1.147	0.867-1.516	.337
IMiD	277	197	1.594	1.059-2.397	.025
Neutropenia					
BOR	461	251	1.672	1.198-2.334	.003
IMiD	479	462	1.023	0.807-1.296	.851
Peripheral neuropathy					
BOR	117	127	0.686	0.425-1.107	.123
IMiD	60	36	1.118	0.527-2.374	.771
Fatigue					
BOR	196	122	1.717	1.031-2.861	.038
IMiD	80	65	1.224	0.890-1.683	.213
Diarrhea					
BOR	199	119	1.190	0.636-2.224	.586
IMiD	55	39	1.379	0.919-2.070	.121
Serious adverse event					
BOR	870	679	1.219	1.083-1.372	.001
IMiD	668	621	1.086	1.000-1.180	.050

BOR, bortezomib; CT, conventional therapy; ET, experimental therapy.

Figure 2B). Subgroup analysis demonstrated significant PFS benefit in the doublet therapy subgroup and in the subgroup for therapy with all classes of drugs except bortezomib (pooled HRs, 0.78; 95% CI, 0.47-1.30; P = .341; Figure 2B; supplemental Figure 2B).

In our analyses, both models showed high heterogeneity, and the prediction intervals suggest an overall nonsignificant result. Specifically, the highest level of heterogeneity was mainly a result of the study by Hjorth et al.[11] However, this study had a low weight in both FE and RE models, and we can conclude that the addition of another drug to either bortezomib or the IMiD backbone or the use of carfilzomib (only in the case of bortezomib) is significantly better than the backbone alone.

OS analyses

Comparison of bortezomib-based regimens. In our analysis, the experimental treatment showed a trend toward benefit in terms of OS compared with conventional treatment (pooled HRs for RE, 0.90; 95% CI, 0.80-1.01; P = .06). This result was confirmed in subgroup analyses for all classes of drugs and for doublet therapy and single-drug therapy (Figure 2C; supplemental Figure 2C).

Comparison of IMiD-based regimens. In our analysis, the experimental treatment showed a trend toward benefit in terms of OS

(pooled HRs for RE, 0.83; 95% CI, 0.66-1.44; P = .101). The models showed high heterogeneity, and the very wide prediction intervals suggest overall uncertain results. Subgroup analyses suggest again that heterogeneity is mainly dependent on the study by Hjorth et al.[11] Indeed, each subgroup analysis showed no heterogeneity, and we observed a significant advantage for doublet, IMiD + mAb, and PI + IMiD groups (Figure 2D; supplemental Figure 2D). Note that the doublet group included all studies except the study by Hjorth et al,[11] thus further underscoring the relevance of this study in increasing the overall heterogeneity of the model.

From the previous findings, we can conclude that the addition of a second drug to a bortezomib backbone or the use of carfilzomib does not improve patients' survival compared with the bortezomib backbone alone, whereas the addition of another drug to the IMiD backbone significantly improves patients' OS.

RR analyses

Comparison of bortezomib-based regimens. In terms of ORR, we found a significant advantage for the experimental treatment (OR, 1.45; 95% CI, 1.16-1.82; P = .001). This advantage was clearly evident with HDAC inhibitor, carfilzomib, and doublet subgroups (supplemental Figure 3A-B). Differences in terms of VGPR are shown in supplemental Figure 3C-D. In terms of CR, we demonstrated a significant advantage for the experimental treatment (OR, 1.84; 95% CI, 1.50-2.26; P < .001). This advantage was evident in all subgroups except BORT + OTHER (supplemental Figure 3E). As for doublet therapy and single-drug treatment subgroups, a significant CR benefit was observed for doublet therapy (supplemental Figure 3F).

Comparison of IMiD-based regimens. In terms of ORR, we demonstrated a significant advantage for the experimental treatment (OR, 2.23; 95% CI, 1.56-3.20; P < .001) across all subgroup analyses except for bortezomib (supplemental Figure 4A-B). In terms of VGPR, a significant difference was found in favor of the experimental arm (pooled VGPRs, 2.49; 95% CI, 1.70-3.64; P < .001) as confirmed for all subgroups except for the IMiD + mAb group (supplemental Figure 4C-D). In terms of CR, we found a significant advantage for the experimental treatment (OR for CR, 2.30; 95% CI, 1.29-4.10; P = .005; supplemental Figure 4E-F).

Altogether, our results suggest that the addition of a second drug to the IMiD backbone or bortezomib backbone (or the use of carfilzomib) significantly increased the probability of obtaining a highest ORR, VGPR, or CR and that this result is more evident in patients treated with HDAC inhibitors, anti-MM mAbs, or PIs + IMiDs.

Toxicity analyses

Common adverse events were similar in both meta-analyses, as detailed in Table 3. We observed a significant difference for thrombocytopenia in the experimental arm of IMiD meta-analysis and a major risk of neutropenia, fatigue, and serious adverse events in the experimental arm of bortezomib meta-analysis.

NMA

All treatments analyzed were divided into 9 groups: dexamethasone, bortezomib, IMiDs, bortezomib + HDACs, bortezomib + mAbs, bortezomib + other drugs, carfilzomib, PI + IMiDs, and IMiDs + mAbs according to drug class (details are provided in supplemental Table 1) or as independent treatments (except for bortezomib or
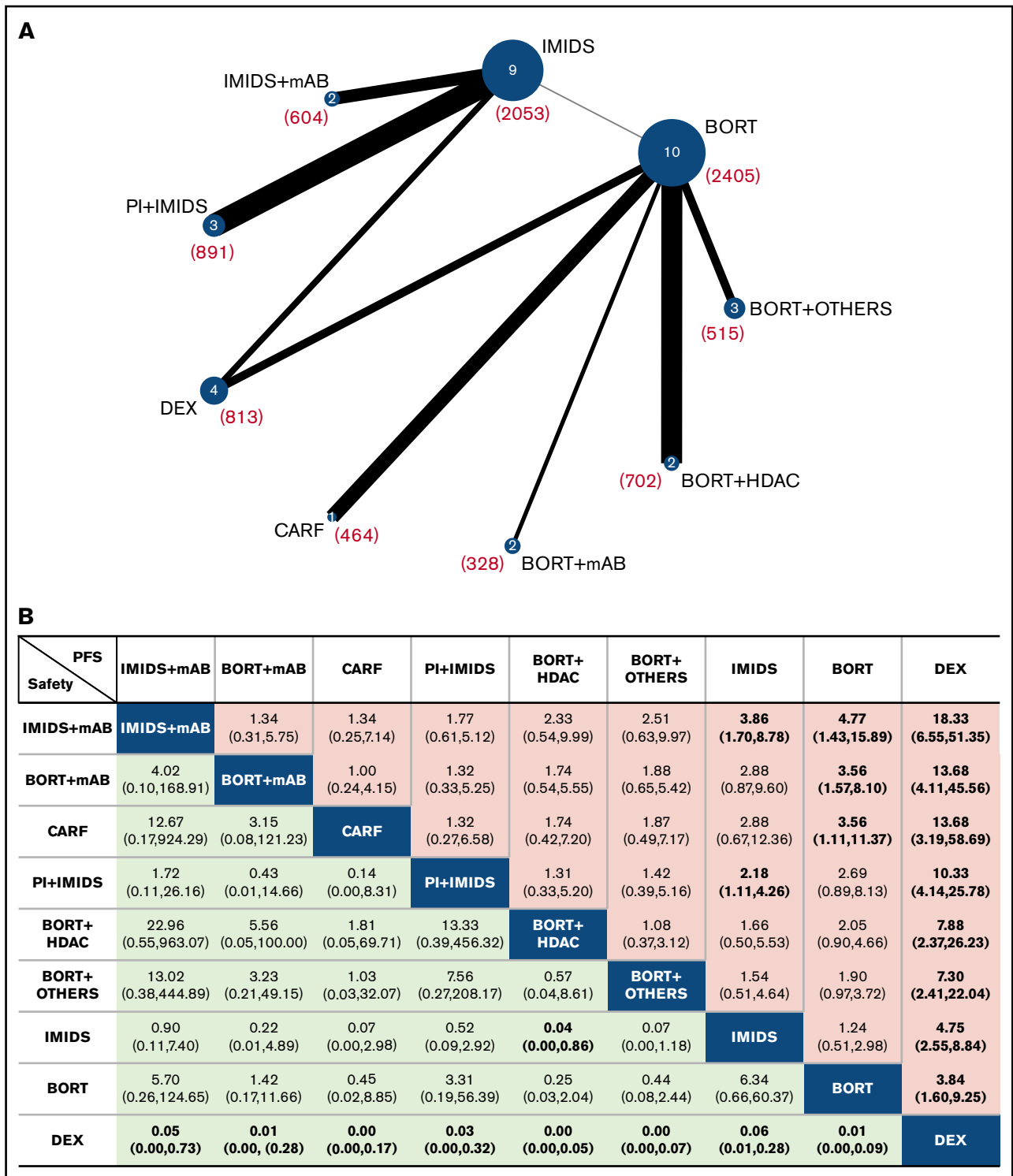
Figure 3. Network of the comparisons and comparative efficacy and tolerance results. (A) Network plot of all treatment groups evaluated in the NMA for PFS. The number of studies that analyzed each treatment group is shown inside the circles; the overall numbers of patients included in the analysis for each group are provided in parentheses. (B) Effect estimates of the treatment in terms of PFS (column headings being compared with row headings) and safety (row headings being compared with column headings). PFS is reported as HR and safety is reported as RR (95% credible intervals are in parentheses). Statistically significant comparisons are shown in bold. DEX, dexamethasone.

bortezomib + dexamethasone regimens grouped into the "bortezomib with or without dexamethasone" group; for thalidomide, thalidomide + dexamethasone and lenalidomide + dexamethasone were grouped into the "IMiDs with or without dexamethasone" group). All regimens or groups were evaluated for differences in PFS, OS, ORR, CR, and safety (network plots are shown in Figure 3A; supplemental Figure 5A-B).
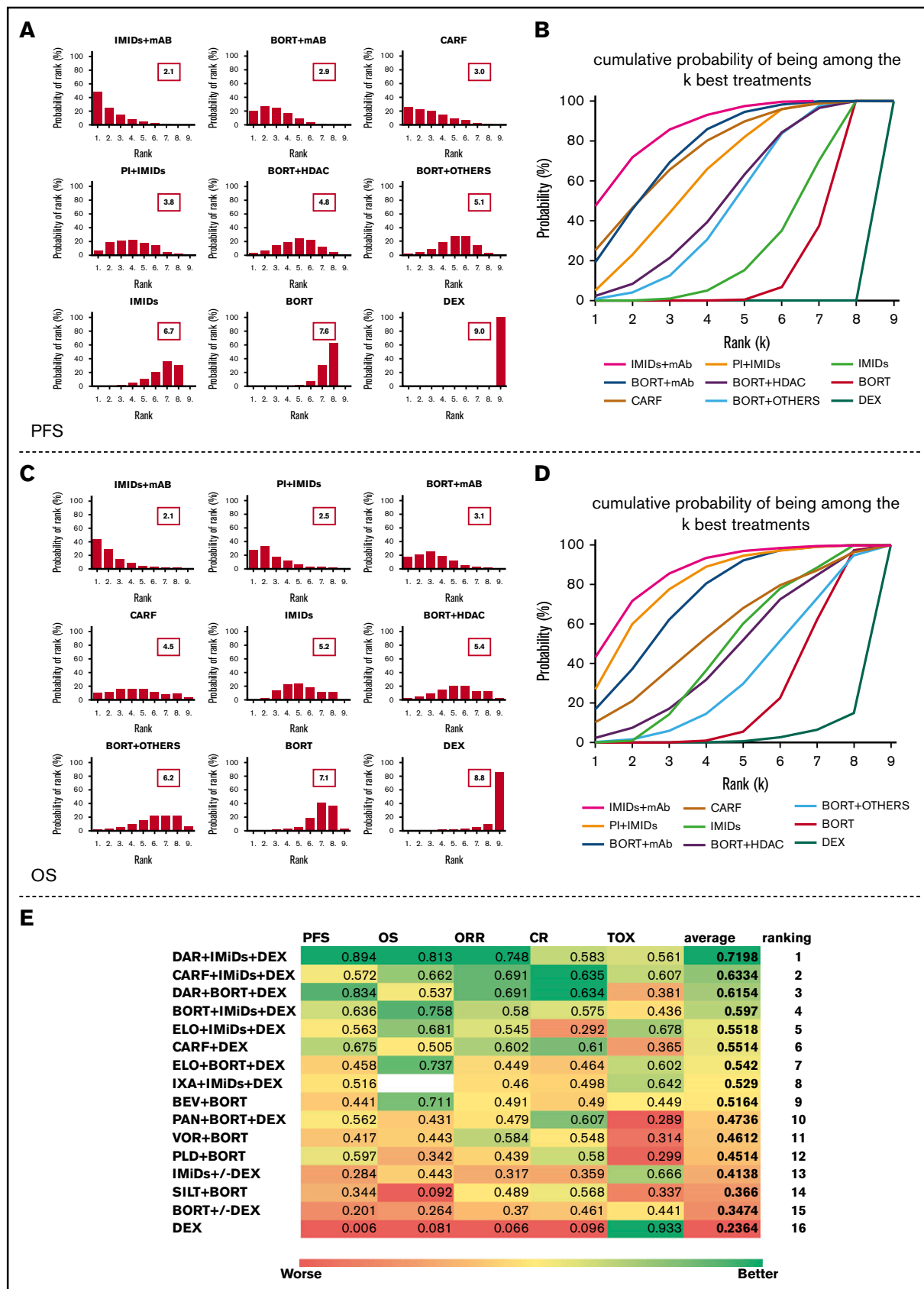
Figure 4.

Each group was simultaneously compared against all other groups through a Bayesian NMA, and efficacy results for PFS and safety are shown in Figure 3B in terms of HRs with 95% credible intervals (efficacy results in terms of OS, ORR, and CR are shown in supplemental Figure 6). All groups evaluated had a significantly lower risk for progression compared with dexamethasone, whereas dexamethasone was found to be significantly safer than any other group. Bortezomib + mAbs and carfilzomib groups were significantly better than the bortezomib alone group, whereas the PI + IMiD group was significantly better than IMiDs alone. Interestingly, IMiDs + mAbs was the only group significantly better than both IMiDs and bortezomib. Regarding safety, the only significant value (except for dexamethasone) was reached by the IMiDs group, which shows a significantly lower RR for toxicity compared with the bortezomib + HDAC group.

We determined which group had the highest probability of being the most effective in term of PFS, OS, ORR, CR, and safety (supplemental Figure 7A). The IMiD + mAB group achieved the highest probability of being the most effective in term of PFS, OS, and ORR, whereas the carfilzomib group reported the highest probability of achieving CR. The bortezomib + HDAC group had the highest probability of inducing a grade 3 to 4 toxicity in MM patients. However, ranking the treatment solely on the basis of the probability of being the best does not account for the uncertainty in the relative treatment effects and may lead to overinterpretation of the results. Thus, we evaluated the probability of each group being at each rank (rank probabilities) and summarized them in a rankogram together with the median rank. We also evaluated the cumulative probability of each group being among the k ranks (k ranges from 1 to 9 in our analysis) together with the SUCRA.[22,23,43]

Figure 4A-B clearly shows that, regarding PFS, the IMiD + mAb group has the highest probability of being the most effective and that almost certainly it is among the 4 best treatments. It should be noted, however, that the bortezomib + mAb and carfilzomib groups are only slightly inferior to the best-ranked treatment and that the magnitude of this difference is very difficult to estimate.

In term of OS, the scenario is slightly different (Figure 4C-D). The IMiD + mAb group is again the highest ranked treatment and has a more than 90% probability of being among the 4 best treatments. In this case, however, the PI + IMiD group performed similarly, followed by the bortezomib + mAb group. Thus, there is considerable uncertainty regarding the ranking of the carfilzomib group, which has similar probabilities of being at each of the 9 ranks. We then estimated SUCRA values for PFS, OS, ORR, CR, and safety and calculated an average value to rank all the groups included in the analysis by using a multiparametric approach (supplemental Figure 7B). This analysis revealed IMiD + mAb, PI + IMiD, bortezomib + mAb, and carfilzomib as the best 4 options for RRMM patients, but we should take into account that the carfilzomib group, while achieving the highest score in term of CR, seems to be inferior to other groups in term of OS and toxicity. Moreover, when examining the rank probabilities (supplemental Figure 7C), the addition of a second drug to an IMiD

backbone (IMiD + mAb and PI + IMiD together) accounts for an impressive 70% probability of being the best regimen in term of OS.

Finally, we investigated which regimen, among all regimens included in the NMA, scores as the overall best regimen. To find the answer, we determined the probability of being the best, cumulative rank probabilities, and SUCRA values for PFS, OS, ORR, CR, and safety and estimated an average value to rank all the treatment options included in the analysis (Figure 4E; supplemental Figure 8). According to average SUCRA values, the daratumumab + lenalidomide-dexamethasone regimen (within the IMiD + mAB group) achieved the highest score (average SUCRA, 0.7198) followed by carfilzomib + lenalidomide-dexamethasone (average SUCRA, 0.6334) and daratumumab + bortezomib-dexamethasone (average SUCRA, 0.6154). These results were further confirmed by the fact that the daratumumab + lenalidomide-dexamethasone regimen reported the highest probability of being the best regimen in terms of PFS (44.56%), OS (27.57%), and ORR (22.26%), with a very low probability of being the worst regimen in terms of grade 3 to 4 toxicities (3.70%) (supplemental Figure 8A). Again, it is relevant that 4 of the 5 best regimens in our ranking include the addition of a second drug to an IMiD backbone, and that 3 of the 5 best regimens include the new anti-MM mAbs daratumumab and elotuzumab.

No significant inconsistency or loop-specific heterogeneity were found in our NMA as shown by the inconsistency factor plots in supplemental Figure 9A (based on the 95% CI of the logRoR crossing the null value of 0).

For exploratory purposes, we performed the same NMA by using a frequentist approach. All regimens were ranked by using the P-score according to the method recently described by Rücker et al.[20] We observed similar results in terms of all variables (supplemental Figure 9B) and the daratumumab + lenalidomide-dexamethasone regimen was again found to be the highest scoring regimen (average P-score, 0.733).

Discussion

The aim of our study was to systematically review and compare all therapeutic options available for RRMM, taking into account the high activity of regimens that included novel agents such as HDAC inhibitors or mAbs. Indeed, although most of these agents have undergone randomized clinical trials, a clear understanding of their relative efficacy is still lacking, and in the absence of predictive biomarkers, treatment choice is exclusively based on patient characteristics (prior treatment, absolute refractoriness of the disease, comorbidities, and the patient's ability to tolerate the regimen) and tumor-specific characteristics (disease burden, risk status, cytogenetics, and score according to the International Scoring System [ISS] and Revised ISS) characteristics. Thus, ranking the value of all these opportunities might be of help in designing therapeutic algorithms to be validated in prospective clinical trials.

In this study, we performed 2 traditional pairwise meta-analyses and an NMA. Indeed, although pairwise meta-analysis can offer a robust

Figure 4. Ranking of treatments based on NMA results. (A,C) Distribution of the probabilities of being at each rank, together with mean rank. Cumulative ranking probabilities for each treatment evaluated in term of (B) PFS and (D) OS. (E) All of the SUCRA values for each regimen in regard to PFS, OS, ORR, CR, and toxicity (TOX; in this case, the higher the SUCRA, the safer the regimen is for patients). An average SUCRA and the average ranking are provided. BEV, bevacizumab; DAR, daratumumab; DEX, dexamethasone; ELO, elotuzumab; IXA, ixazomib; PAN, panobinostat; PLD, pegylated liposomal doxorubicin; SILT, siltuximab; VOR, vorinostat.

tool for evaluating the strength of evidence derived from direct comparative studies, Bayesian NMA seems to be the best tool for investigating the strength of evidence for regimens that have not undergone direct comparison.

The pairwise meta-analyses showed that a 3-drug combination that includes bortezomib or lenalidomide reduces the risk of disease progression and increases the probability of achieving at least a partial response, even at the cost of greater toxicity compared with doublet therapy. However, only IMiD-based regimens plus mAbs or PIs demonstrated a clear advantage in terms of OS. To overcome the limitation of the lack of direct comparisons, we used the NMA approach and grouped all available regimens into 9 subgroups. In our analysis, the IMiD + mAb group achieved the highest probability of being the most effective regimen for RRMM in term of PFS followed by the bortezomib + mAb and the carfilzomib group with very close SUCRA scores; OS followed closely with the PI + IMiD and bortezomib + mAb groups. When considering specific regimens, we observed that the daratumumab + lenalidomide-dexamethasone regimen reached the highest SUCRA values in term of PFS, OS, and ORR with a good safety profile, ranking as the average best treatment. We also observed that a regimen that adds a second drug to an IMiD backbone has a 70% probability of being the best treatment option and that, among the 5 regimens (of 16 evaluated) scoring the highest average SUCRA, 4 included an IMiD backbone.

Thus, the results of our pairwise comparison indicate that whenever possible, a 3-drug regimen must be preferred over a 2-drug regimen, and our NMA further suggests that this regimen should have an IMiD backbone (preferentially combined with a new anti-MM mAb or alternatively with a PI). Furthermore, we observed that the combination of an HDAC inhibitor or pegylated liposomal doxorubicin with bortezomib significantly increases the toxicity of the treatment without adding additional benefit compared with regimens based on PIs + IMiDs, carfilzomib, or mAbs. It is clear that these regimens should be reserved for selected very fit MM patients. Furthermore, a correlation between activity and efficacy outcomes was not clearly demonstrated. Although the highest CR rates were attained with carfilzomib or PI-based 3-drug regimens, the 4 regimens that included elotuzumab and daratumumab ranked highest in terms of OS. These data are consistent with clinical outcomes observed in other tumors undergoing immunotherapy-based regimens[44-48] and are supported by recent experimental findings on MM pathogenesis. Indeed, the encouraging results of immunotherapy are, to a degree, not surprising when taking into account that MM is a malignancy strongly dependent on microenvironment features, including inflammation and immunosuppression events.[49-51] In such a scenario, it is likely that new immunotherapeutic agents will quickly become available for RRMM patients. Immune checkpoint inhibitors such as programmed cell death protein 1 and programmed death-ligand 1 blocking antibodies have shown promising preliminary results in small phase 1/2 trials that demonstrate the benefit of IMiDs in combination with pembrolizumab; several phase 2/3 trials with agents of this class are currently ongoing.[52,53]

Unfortunately, these novel drugs are costly, thus leading to sustainability concerns for health care systems. It is therefore imperative to prioritize treatment schedules that have demonstrated improvement in survival and/or quality of life and to develop and implement new methodologies of comparative analysis such as NMA for constructing decision-making therapeutic algorithms. This information might benefit health technology assessment by providing a novel benchmark (NMA) ranking for cost-effectiveness. One of the major strengths of our analysis is the simultaneous comparison of the results from all published trials, thus achieving information regarding hierarchy even in the absence of direct comparisons.

However, our work has limitations. First, data were retrieved from published studies rather than from individual patients' records. Second, and most important, potential biases can be produced by the heterogeneity of the agents and patient populations included in the analysis. An optimal meta-analysis in this setting should compare studies that are similar in terms of patient characteristics (such as number of prior regimens, percentages of IMiD and PI refractoriness, age, MM risks) and treated with the same regimens in the same treatment line, which cannot be done with current evidence. In addition, no information is provided on the best sequencing of the combinations that were investigated. Finally, this work is intrinsically heuristic in nature and should therefore be considered a snapshot of current evidence, taking into account that some of these regimens are candidates for first-line therapy.

To the best of our knowledge, this is the first NMA on RRMM that includes all regimens currently evaluated in randomized trials in MM. Our findings suggest that a 3-drug regimen containing the lenalidomide-dexamethasone backbone, preferentially combined with anti-MM mAbs daratumumab or elotuzumab, has the highest probability of being ranked as the best treatment in this setting, underlying the role of immunotherapy in MM. Prospective randomized trials are eagerly awaited to clarify the optimal sequencing of treatments for MM.

Authorship

Contribution: C.B. and D.C. designed the research, conducted the literature search, performed data extraction, conducted the statistical analysis, interpreted the results, and wrote the manuscript; N.S. reviewed the analysis, assessed methodologic quality, and wrote the manuscript; M.C. and T.G. reviewed the analysis; M.R., P. Tagliaferri and P. Tassone reviewed the protocol, assessed methodologic quality, reviewed and edited the manuscript; and all authors approved the final version of the manuscript.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: C.B., 0000-0002-1522-4504.

Correspondence: Pierfrancesco Tassone, Department of Experimental and Clinical Medicine, Magna Graecia University, Salvatore Venuta University Campus, Viale Europa, 88100 Catanzaro, Italy; e-mail: tassone@unicz.it.

References

1. Laubach J, Garderet L, Mahindra A, et al. Management of relapsed multiple myeloma: recommendations of the International Myeloma Working Group. *Leukemia*. 2016;30(5):1005-1017.

2. Ludwig H, Miguel JS, Dimopoulos MA, et al. International Myeloma Working Group recommendations for global myeloma care. *Leukemia*. 2014;28(5):981-992.

3. Moreau P, Attal M, Facon T. Frontline therapy of multiple myeloma. *Blood*. 2015;125(20):3076-3084.

4. Engelhardt M, Terpos E, Kleber M, et al; European Myeloma Network. European Myeloma Network recommendations on the evaluation and treatment of newly diagnosed patients with multiple myeloma. *Haematologica*. 2014;99(2):232-242.

5. Liu L, Zhao N, Xu W, Sheng Z, Wang L. Pooled analysis of the reports of carfilzomib, panobinostat, and elotuzumab combinations in patients with refractory/relapsed multiple myeloma. *J Hematol Oncol*. 2016;9(1):54.

6. Liu JD, Sun CY, Tang L, et al. Efficacy and safety of panobinostat in relapsed or/and refractory multiple myeloma: meta analyses of clinical trials and systematic review. *Sci Rep*. 2016;6:27361.

7. Parmar MK, Torri V, Stewart L. Extracting summary statistics to perform meta-analyses of the published literature for survival endpoints. *Stat Med*. 1998;17(24):2815-2834.

8. Shamseer L, Moher D, Clarke M, et al; PRISMA-P Group. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ*. 2015;349:g7647.

9. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med*. 2015;162(11):777-784.

10. Orlowski RZ, Gercheva L, Williams C, et al. A phase 2, randomized, double-blind, placebo-controlled study of siltuximab (anti-IL-6 mAb) and bortezomib versus bortezomib alone in patients with relapsed or refractory multiple myeloma. *Am J Hematol*. 2015;90(1):42-49.

11. Hjorth M, Hjertner Ø, Knudsen LM, et al; Nordic Myeloma Study Group (NMSG). Thalidomide and dexamethasone vs. bortezomib and dexamethasone for melphalan refractory myeloma: a randomized study. *Eur J Haematol*. 2012;88(6):485-496.

12. Tierney JF, Stewart LA, Ghersi D, Burdett S, Sydes MR. Practical methods for incorporating summary time-to-event data into meta-analysis. *Trials*. 2007;8:16.

13. Staropoli N, Ciliberto D, Botta C, et al. Pegylated liposomal doxorubicin in the management of ovarian cancer: a systematic review and metaanalysis of randomized trials. *Cancer Biol Ther*. 2014;15(6):707-720.

14. Detsky AS, Naylor CD, O'Rourke K, McGeer AJ, L'Abbé KA. Incorporating variations in the quality of individual randomized trials into meta-analysis. *J Clin Epidemiol*. 1992;45(3):255-265.

15. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ*. 1997;315(7109):629-634.

16. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics*. 1994;50(4):1088-1101.

17. Staropoli N, Ciliberto D, Chiellino S, et al. Is ovarian cancer a targetable disease? A systematic review and meta-analysis and genomic data investigation. *Oncotarget*. 2016;7(50):82741-82756.

18. Ciliberto D, Staropoli N, Chiellino S, Botta C, Tassone P, Tagliaferri P. Systematic review and meta-analysis on targeted therapy in advanced pancreatic cancer. *Pancreatology*. 2016;16(2):249-258.

19. Ciliberto D, Botta C, Correale P, et al. Role of gemcitabine-based combination therapy in the management of advanced pancreatic cancer: a meta-analysis of randomised trials. *Eur J Cancer*. 2013;49(3):593-603.

20. Rücker G, Schwarzer G. Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Med Res Methodol*. 2015;15:58.

21. Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JP. Evaluating the quality of evidence from a network meta-analysis. *PLoS One*. 2014;9(7):e99682.

22. Chaimani A, Higgins JP, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. *PLoS One*. 2013;8(10):e76654.

23. Jansen JP, Trikalinos T, Cappelleri JC, et al. Indirect treatment comparison/network meta-analysis study questionnaire to assess relevance and credibility to inform health care decision making: an ISPOR-AMCP-NPC Good Practice Task Force report. *Value Health*. 2014;17(2):157-173.

24. Garderet L, Iacobelli S, Moreau P, et al. Superiority of the triple combination of bortezomib-thalidomide-dexamethasone over the dual combination of thalidomide-dexamethasone in patients with multiple myeloma progressing or relapsing after autologous transplantation: the MMVAR/IFM 2005-04 Randomized Phase III Trial from the Chronic Leukemia Working Party of the European Group for Blood and Marrow Transplantation. *J Clin Oncol*. 2012;30(20):2475-2482.

25. Dimopoulos MA, Oriol A, Nahi H, et al; POLLUX Investigators. Daratumumab, lenalidomide, and dexamethasone for multiple myeloma. *N Engl J Med*. 2016;375(14):1319-1331.

26. Lonial S, Dimopoulos M, Palumbo A, et al; ELOQUENT-2 Investigators. Elotuzumab therapy for relapsed or refractory multiple myeloma. *N Engl J Med*. 2015;373(7):621-631.

27. Stewart AK, Rajkumar SV, Dimopoulos MA, et al; ASPIRE Investigators. Carfilzomib, lenalidomide, and dexamethasone for relapsed multiple myeloma. *N Engl J Med*. 2015;372(2):142-152.

28. Kropff M, Baylon HG, Hillengass J, et al. Thalidomide versus dexamethasone for the treatment of relapsed and/or refractory multiple myeloma: results from OPTIMUM, a randomized trial. *Haematologica*. 2012;97(5):784-791.

29. Richardson PG, Sonneveld P, Schuster MW, et al; Assessment of Proteasome Inhibition for Extending Remissions (APEX) Investigators. Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. *N Engl J Med*. 2005;352(24):2487-2498.

30. Dimopoulos MA, Chen C, Spencer A, et al. Long-term follow-up on overall survival from the MM-009 and MM-010 phase III trials of lenalidomide plus dexamethasone in patients with relapsed or refractory multiple myeloma. *Leukemia*. 2009;23(11):2147-2152.

31. Weber DM, Chen C, Niesvizky R, et al; Multiple Myeloma (009) Study Investigators. Lenalidomide plus dexamethasone for relapsed multiple myeloma in North America. *N Engl J Med*. 2007;357(21):2133-2142.

32. Dimopoulos M, Spencer A, Attal M, et al; Multiple Myeloma (010) Study Investigators. Lenalidomide plus dexamethasone for relapsed or refractory multiple myeloma. *N Engl J Med*. 2007;357(21):2123-2132.

33. Jakubowiak A, Offidani M, Pégourie B, et al. Randomized phase 2 study: elotuzumab plus bortezomib/dexamethasone vs bortezomib/dexamethasone for relapsed/refractory MM. *Blood*. 2016;127(23):2833-2840.

34. Palumbo A, Chanan-Khan A, Weisel K, et al; CASTOR Investigators. Daratumumab, bortezomib, and dexamethasone for multiple myeloma. *N Engl J Med*. 2016;375(8):754-766.

35. San-Miguel JF, Hungria VT, Yoon SS, et al. Panobinostat plus bortezomib and dexamethasone versus placebo plus bortezomib and dexamethasone in patients with relapsed or relapsed and refractory multiple myeloma: a multicentre, randomised, double-blind phase 3 trial. *Lancet Oncol*. 2014;15(11):1195-1206.

36. Dimopoulos MA, Moreau P, Palumbo A, et al; ENDEAVOR Investigators. Carfilzomib and dexamethasone versus bortezomib and dexamethasone for patients with relapsed or refractory multiple myeloma (ENDEAVOR): a randomised, phase 3, open-label, multicentre study. *Lancet Oncol*. 2016;17(1):27-38.

37. White D, Kassim A, Bhaskar B, Yi J, Wamstad K, Paton VE. Results from AMBER, a randomized phase 2 study of bevacizumab and bortezomib versus bortezomib in relapsed or refractory multiple myeloma. *Cancer*. 2013;119(2):339-347.

38. Orlowski RZ, Nagler A, Sonneveld P, et al. Final overall survival results of a randomized trial comparing bortezomib plus pegylated liposomal doxorubicin with bortezomib alone in patients with relapsed or refractory multiple myeloma. *Cancer*. 2016;122(13):2050-2056.

39. Orlowski RZ, Nagler A, Sonneveld P, et al. Randomized phase III study of pegylated liposomal doxorubicin plus bortezomib compared with bortezomib alone in relapsed or refractory multiple myeloma: combination therapy improves time to progression. *J Clin Oncol*. 2007;25(25):3892-3901.

40. Dimopoulos M, Siegel DS, Lonial S, et al. Vorinostat or placebo in combination with bortezomib in patients with multiple myeloma (VANTAGE 088): a multicentre, randomised, double-blind study. *Lancet Oncol*. 2013;14(11):1129-1140.

41. Moreau P, Pylypenko H, Grosicki S, et al. Subcutaneous versus intravenous administration of bortezomib in patients with relapsed multiple myeloma: a randomised, phase 3, non-inferiority study. *Lancet Oncol*. 2011;12(5):431-440.

42. Moreau P, Masszi T, Grzasko N, et al; TOURMALINE-MM1 Study Group. Oral ixazomib, lenalidomide, and dexamethasone for multiple myeloma. *N Engl J Med*. 2016;374(17):1621-1634.

43. Cope S, Jansen JP. Quantitative summaries of treatment effect estimates obtained with network meta-analysis of survival curves to inform decision-making. *BMC Med Res Methodol*. 2013;13:147.

44. Cusi MG, Botta C, Pastina P, et al. Phase I trial of thymidylate synthase poly-epitope peptide (TSPP) vaccine in advanced cancer patients. *Cancer Immunol Immunother*. 2015;64(9):1159-1173.

45. Correale P, Botta C, Rotundo MS, et al. Gemcitabine, oxaliplatin, levofolinate, 5-fluorouracil, granulocyte-macrophage colony-stimulating factor, and interleukin-2 (GOLFIG) versus FOLFOX chemotherapy in metastatic colorectal cancer patients: the GOLFIG-2 multicentric open-label randomized phase III trial. *J Immunother*. 2014;37(1):26-35.

46. Correale P, Botta C, Martino EC, et al. Phase Ib study of poly-epitope peptide vaccination to thymidylate synthase (TSPP) and GOLFIG chemo-immunotherapy for treatment of metastatic colorectal cancer patients. *OncoImmunology*. 2015;5(4):e1101205.

47. Herbst RS, Baas P, Kim DW, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet*. 2016;387(10027):1540-1550.

48. Schadendorf D, Hodi FS, Robert C, et al. Pooled analysis of long-term survival data from phase II and Phase III trials of ipilimumab in unresectable or metastatic melanoma. *J Clin Oncol*. 2015;33(17):1889-1894.

49. Rossi M, Botta C, Correale P, Tassone P, Tagliaferri P. Immunologic microenvironment and personalized treatment in multiple myeloma. *Expert Opin Biol Ther*. 2013;13(Suppl 1):S83-S93.

50. Botta C, Gullà A, Correale P, Tagliaferri P, Tassone P. Myeloid-derived suppressor cells in multiple myeloma: pre-clinical research and translational opportunities. *Front Oncol*. 2014;4:348.

51. Botta C, Di Martino MT, Ciliberto D, et al. A gene expression inflammatory signature specifically predicts multiple myeloma evolution and patients survival. *Blood Cancer J*. 2016;6(12):e511.

52. Zagouri F, Terpos E, Kastritis E, Dimopoulos MA. Emerging antibodies for the treatment of multiple myeloma. *Expert Opin Emerg Drugs*. 2016;21(2):225-237.

53. Rosenblatt J, Avigan D. Targeting the PD-1/PD-L1 axis in multiple myeloma: a dream or a reality? *Blood*. 2017;129(3):275-279.