Ferrata Storti Foundation

Efficacy of first-line treatments for multiple myeloma patients not eligible for stem cell transplantation: a network meta-analysis

Hedwig M. Blommestein,[1,2*] Chrissy H.Y. van Beurden-Tan,[3*]
Margreet G. Franken,[1] Carin A. Uyl-de Groot,[1,2] Pieter Sonneveld[3]
and Sonja Zweegman[4]

*HB and CvBT contributed equally to this work

[1]Erasmus School of Health Policy & Management, Institute for Medical Technology Assessment, Erasmus University Rotterdam; [2]Comprehensive Cancer Organisation, Utrecht; [3]Erasmus MC Cancer Institute, Rotterdam and [4]Department of Hematology, Amsterdam UMC, the Netherlands

Check the online version for the most updated information on this article, online supplements, and information on authorship & disclosures: www.haematologica.org/content/104/5/1026

ABSTRACT

Decision making for patients with multiple myeloma (MM) not transplant eligible (NTE) is complicated by a lack of head-to-head comparisons of standards of care, the increase in the choice of treatment modalities, and the promising results that are rapidly evolving from studies with novel regimens. To support evidence-based decision making, we performed a network meta-analysis for NTE MM patients that synthesizes direct and indirect evidence and enables a comparison of all treatments. Relevant randomized clinical trials were identified by a systematic literature review in EMBASE®, MEDLINE®, MEDLINE®-in-Process and the Cochrane Central Register of Controlled Trials for January 1999 to March 2016. Efficacy outcomes [i.e. the hazard ratio (HR) and 95% confidence interval (95%CI) for progression-free survival] were extracted and synthesized in a random effects network-meta analysis. In total, 24 studies were identified including 21 treatments. According to the network-meta analysis, the HR for progression-free survival was favorable for all NTE MM treatments compared to dexamethasone (HR: 0.19-0.90). Daratumumab-bortezomib-melphalan-prednisone and bortezomib-melphalan-prednisone-thalidomide with bortezomib-thalidomide maintenance were identified as the most effective treatments (HR: 0.19, 95%CI: 0.08-0.45 and HR: 0.22, 95%CI: 0.10-0.51, respectively). HR and 95%CI for currently recommended treatments, bortezomib-lenalidomide-dexamethasone, bortezomib-melphalan-prednisone, and lenalidomide-dexamethasone compared to dexamethasone, were 0.31 (0.16-0.59), 0.39 (0.20-0.75), and 0.44 (0.29-0.65), respectively. In addition to identifying the most effective treatment options, we illustrate the additional value and evidence of network meta-analysis in clinical practice. In the current treatment landscape, the results of network meta-analysis may support evidence-based decisions and ultimately help to optimize treatment and outcomes of NTE MM patients.

Introduction

Multiple myeloma (MM) is a hematologic disease characterized by the proliferation of malignant plasma cells, causing disease-related symptoms such as anemia, hypercalcemia, renal and bone disease. The age-standardized incidence rate is 4.5 per 100,000.[1] Incidence increases with age and two-thirds of the patients diagnosed with MM are over 65 years of age.[2] The treatment armamentarium has been greatly increased in the last decade, with novel proteasome inhibitors (PIs), immunomodulatory drugs (IMiDs), and monoclonal antibodies now being incorporated in first-line treatment regimens, which have considerably improved progression-free survival (PFS) and overall survival (OS) of MM. Given the median age

of 70 years at diagnosis, the majority of newly diagnosed (ND) MM patients are not transplant eligible (NTE) for stem cell transplant (SCT). Current standards of care for NTE NDMM patients are bortezomib-melphalan-prednisone (VMP), lenalidomide-dexamethasone (Rd), and in the USA bortezomib-Rd (Vrd),[3] supported by randomized phase III trials.[4-6] Recently, better PFS was demonstrated for daratumumab-VMP (DaraVMP) compared to VMP.[7]

Although randomized clinical trials (RCTs) remain the gold standard to define standards of care, we predict that in the current treatment landscape the role of network meta-analysis (NMA) will become increasingly important. Firstly, today there is more than one standard of care, but a randomized study between two registered standards of care is highly unlikely to be performed because of the reluctance of pharmaceutical industries to support such studies.[8] Therefore, head-to-head comparisons of VMP *versus* Rd or VRd *versus* VMP are not likely to be initiated.[9] NMA can help to discriminate between efficacy of non-head-to-head compared regimens. Secondly, with the increase in the number of treatment modalities, the number of smaller randomized Phase II studies is expected to increase at the cost of Phase III RCTs. NMA provides more solid estimates of treatment effects by combining RCTs that provide direct and indirect evidence for effectiveness and allows competing treatments to be ranked.[10] Thirdly, with the high number of studies currently enrolling patients, standard of care arms are expected to change within short time frames.[8] This hampers the development of classical phase III trials, as at the end of the study it might appear that the standard arm of the study no longer reflects the reality in the clinic. Finally, the heterogeneous biological characteristics of MM and the clonal evolution of the disease will lead to studies with a smaller sample size that will not allow for randomization, increasing the need for indirect comparisons.

There are currently two systematic literature reviews (SLRs) and NMAs available for first-line NTE NDMM treatments.[11,12] Due to the timing of their searches and selection criteria, these reviews did not, however, include all currently available treatments (e.g. VRd, VMPT-VT, DaraVMP) and RCT evidence (e.g. HOVON87 comparing MPT-T and MPR-R[13]). To support evidence-based decision making in clinical practice, we performed an SLR and an NMA synthesizing all direct and indirect evidence from phase III RCTs that is currently available and compared the outcome of all treatment options for NTE NDMM patients.

Methods

Systematic literature review

An SLR was conducted in the databases EMBASE®, MEDLINE®, MEDLINE®-in-Process and the Cochrane Central Register of Controlled Trials for the period January 1, 1999 to March 1, 2016 to identify relevant studies (*Online Supplementary Appendix 1*). Studies were included if they described a Phase III RCT among newly diagnosed adult patients with MM. Furthermore, one of the pre-specified treatments (*Online Supplementary Appendix 2*) had to be part of the regimens of the RCT. After removing duplicates, citations were first screened on the basis of title and abstract and then screened on the contents of their full text. Citations were excluded due to the following reasons: not in English, review, study phase, intervention, dis-

ease, study design, meta-analysis, patient population, economic outcomes, meta-analysis, and other. (For a detailed description of the exclusion categories see *Online Supplementary Appendix 2*). To incorporate the latest clinical developments, the publication of the pre-specified interim analysis of the phase III ALCYONE RCT comparing DaraVMP to VMP[7] was added as additional record.

Data extraction

Data were extracted on trial details (i.e. publication source, trial ID, trial number, research, and comparator treatment(s), number of patients, median age, and primary outcome, and follow up) and efficacy outcomes. Efficacy outcomes included PFS and OS. For OS we obtained median survival. For PFS we obtained the median survival, 95% confidence interval (CI) and hazard ratio (HR) and 95%CI of the HR. In cases in which HRs and/or 95%CI for PFS were not reported, we estimated the missing data with the available Kaplan-Meier curves using the methods described by Tierney *et al*.[14] In cases in which multiple sources reported on the same trial, the most recently published PFS data were extracted. Risk of bias in randomized trials was assessed using the Cochrane Collaboration tool[15] (*Online Supplementary Appendix 3*).

Network meta-analysis

A network was made from the identified treatment options in the SLR. It includes the HRs for PFS from the trials for treatments that were compared head-to-head. A comparison between all treatments can be made based on a common comparator (i.e. reference treatment). The choice of the reference treatment does not influence the outcomes of the study and final results can be presented relative to all included treatments. The oldest treatment (i.e. dexamethasone) was selected as a reference treatment from which the relative effectiveness of all treatments was estimated. We performed a similar analysis with MPT as reference treatment, given that this regimen was used as (comparator) treatment in several RCTs. Treatments were sorted based on their *P*-score. This *P*-score measures the average proportion of treatments worse than the respective treatment where 1 means theoretically the best and 0 means the worst.[16]

To conduct an NMA for 2- and multi-arm studies, we used the netmeta package v.0.9-7 in R version 3.3.1 (*Online Supplementary Appendix 4*). We ran a random effects model assuming that the included studies represent a random sample of effect sizes that could have been observed and that the effect can best be estimated by the mean of all available studies. A random effects model was deemed appropriate since there were multiple trials available for some comparisons (e.g. MPT with MP) and sampling error was not considered to be the most plausible explanation for the observed variation. With a random effects model we allow for differences in the patient population and implementations of interventions.[17] The netmeta package uses a frequentist approach based on the graph-theoretical methods routinely applied in electrical networks.[18,19] In contrast to the Bayesian approach that produces credible intervals, analysis based on the frequentist approach produces 95%CIs and, as all CIs, these should be interpreted as follows: 95% of the produced CIs would contain the true value if the analysis were repeated many times.[20]

Face-validity of the NMA results was checked by comparing the computed HRs by the NMA with the HRs reported in the publications of the trials. To validate our outcomes to a previously reported NMA,[12] we performed a scenario analysis with different treatment groups (separating MPT and MPT-T) and a scenario with a limited number of studies. In the third scenario analysis, we used a fixed effect model instead of a random

Table 1. Data extraction of the included trials.

Trial reference Trial ID NCT number	Primary outcome	Randomized / enrolled patients	Treatment	Median age research treatment years (range)	N itt	Median PFS 95% CI	HRs (95% CI) {research vs. comparator treatment}		Median OS	Median follow up
Facon 2006	OS	500	D	70 (67-73)	127	12.2 (10.2-14.2)			33.4	82.8
IFM 95/01 n/r			MP	70 (68-72)	122	21.1 (17.8-24.4)	0.75 (0.62-0.91) {MP vs. D}	1.15 (0.93-1.42) {MP vs. MD}	34	82.8
			MD	69 (68-72)	118	22.9 (19.0-26.8)	0.66 (0.53-0.81) {MD vs. D}	1.45 (1.17-1.79) {DI vs MD}	39.6	82.8
			DI	69 (67-72)	121	15.2 (9.9-20.5)	0.92 (0.76-1.11) {DI vs. D}	1.26 (1.04-1.53) {DI vs. MP}	32	82.8
Facon 2007	OS	447	MPT	n/r (65-75[1])	125	27.5 (23.4-31.6)	0.59 (0.44-0.78) {MPT vs. M100}		51.6	51.5
IFM 99–06 NCT00367185			MP	n/r (65-75[2])	196	17.8 (15.1-20.5)	0.51 (0.39-0.66) {MPT vs. MP}		33.2	51.5
			M100	n/r (65-75[3])	126	19.4 (17.4-21.4)	0.87 (0.68-1.1) {M100 vs. MP}		38.3	51.5
Morgan 2013	PFS, OS	856	MP	73 (57-89)	423	12 n/r	0.81 (0.69-0.94) {CTD(a) vs. MP}		32	70.8
MRC M IX ISRCTN68454111			CTDa	73 (58-87)	426	13 n/r			34	70.8
Rajkumar 2008	TTP	470	TD	64 (39-86)	235	14.9 n/r	0.5 (0.38-0.64) {TD vs. D}		NR	17
MM-003 NCT00057564			D	64 (31-84)	235	6.5 n/r			30	18
Ludwig 2009	PFS, tolerance	289	TD	72 (54-86)	145	16.7 n/r	1.3 (0.95-1.78) {TD vs. MP}		41.5	28.1
NCT00205751			MP	72 (55-86)	144	20.7 n/r			49.4	28.1
Palumbo 2008	RR, PFS	331	MPT-T	72	167	21.8 (19.6-26.1)	0.63 (0.48-0.81) {MPT vs. MP}		45	38.4
GIMEMA NCT00232934			MP	72	164	14.5 (12.2-17)			47.6	37.7
Hulin 2009	OS	232	MPT	79 (75-89)	115	24.1 (19.4-29)	0.61 (0.46-0.82) {MPT vs. MP}		44	47.5
IFM 01/01 Trial n/r			MP		117	18.5 (14.6-21.3)			29.1	47.5
Waage 2010	OS	363	MPT-T	75	184	15 (12-19)	0.89 (0.7-1.13) {MPT vs. MP}		29	42
NMSG NCT00218855			MP	74	179	14 (11-18)			32	42
Beksac 2010	Treatment response, toxicities	122	MPT	69	60	n/r	0.7 (0.42-1.17) {MPT vs. MP}		26	35
TMSG NCT00934154			MP	72	62	n/r			28	23
Wijermans 2010	EFS	344	MPT-T	72 (65-87)	171	15 n/r	0.79 (0.62-1) {MPT vs. MP}		40	39
HOVON-49 ISRCTN90692740			MP	73 (65-84)	173	11 n/r			31	39
Sacchi 2011	n/r	135	MPT	76 (66–89)	70	33 n/r	0.67 (0.38-1.18) {MPT vs. MP}		52	30
	n/r		MP	79 (68–88)	79	22 65 n/r			32	30
Hungria 2016	ORR	82	CTD	70	32	25.9 n/r	0.89 (0.48-1.64) {MPT vs. CTD}		32.4	37.5
NCT01532856			TD	72	18	21.5 n/r	1.1 (0.53-2.31) {TD vs. CTD}		54.6	37.5
			MPT	72	32	38.5	0.73 (0.34-1.59)		42	37.5

continued from the previous page

Study	Endpoint	N	Arm	Age (range)	Events	Median (CI)	HR (CI) {comparison}	OS	OS
						n/r	{MPT *vs.* TD}		
San Miguel 2008	TTP	682	VMP	71 (57–90)	344	21.7 n/r	0.56 (0.4-0.79) {VMP *vs.* MP}	56.4	60.1
VISTA NCT00111319			MP	71 (48–91)	338	15.2 n/r		43.1	60.1
Mateos 2014	n/r	260	VTP	73 (69–76)	130	23 n/r	0.8 (0.61-1.04) {VMP *vs.* VTP}	43	72
GEM2005 NCT00443235			VMP	73 (68–77)	130	32 n/r		63	72
Niesvizky 2015	PFS	502	VD	75 (67-79)	168	14.7 (12-18.6)	1.12 (0.83-1.51) {VD *vs.* VTD}	49.8	44.3
UPFRONT NCT00507416			VTD	73 (66-77)	167	15.4 (12.6-24.2)	0.89 (0.66-1.21) {VTD *vs.* VMP}	51.5	41.3
			VMP	72 (68-77)	167	17.3 (14.8-20.3)	1.11 (0.84-1.48) {VD *vs.* VMP}	53.1	43.4
Palumbo 2014	PFS	511	VMPT-VT	71 (68-75)	254	35.3 n/r	0.58 (0.47-0.71) {VMPT-VT *vs.* VMP}	NR	54
GIMEMA0305 NCT01063179			VMP	71 (68-75)	257	24.8 n/r		60.6	54
Zonder 2011	PFS	198	RD	n/r[4]	99	39 (26-53)	0.56 (0.39-0.79) {RD *vs.* D}	NR	45.4
S0232 NCT00064038			D	n/r[5]	99	15 (8-23)		NR	45.4
Benboubkher 2014	PFS	1623	Rd	73 (44–91)	535	25.5 n/r	0.97 (0.83-1.12) {MPT *vs.* RD18}	58.9	45.5
FIRST/MM-020 NCT00689936			Rd18	73 (40–89)	541	20.7 n/r	1.43 (1.22-1.67) {RD18 *vs.* RD}	56.7	45.5
			MPT	73 (51–92)	547	21.2 n/r	1.39 (1.18-1.64) {MPT *vs.* RD}	48.5	45.5
Zweegman 2016	PFS	568	MPT-T	72 (60-91)	280	20 (18-23)	0.87 (0.72-1.04) {MPR-R *vs.* MPT-T}	49	32.6
HOVON-87 EUDRACT 2007-004007-34			MPR-R	73 (60-87)	280	22 (19-27)		50	32.6
Stewart 2015	PFS	306	MPT-T	76 (54-92)	154	21 (18-27)	0.84 (0.64-1.09) {MPT-T *vs.* MPR-R}	52.6	40.7[7]
ECOG E1A06 NCT00602641			MPR-R	77 (63-92)	152	18.7 (16-22)		47.7	
Magarotto 2016	PFS	654	MPR-R	74 (63-91)	218	24 n/r	0.81 (0.63-1.03) {MPR-R *vs.* RD}	NR	39
EMN01 NCT01093196			CPR	73 (63-87)	222	20 n/r	1.01 (0.9-1.13) {CPR *vs.* RD}	NR	39
			Rd	73 (50-89)	222	21 n/r	0.8 (0.63-1.02) {MPR-R *vs.* CPR}	NR	39
Palumbo 2012	PFS	459	MPR-R	71 (65–87)	152	31 n/r	0.49 (0.35-0.69) {MPR-R *vs.* MPR}	56	53
MM-015 NCT00405756			MP	72 (65–91)	154	13 n/r	1.19 (0.94-1.5) {MP *vs.* MPR}	52	53
			MPR	71 (65–86)	153	14 n/r	0.4 (0.29-0.54) {MPR-R *vs.* MP}	54	53
Durie 2017 ||	PFS	525	VRd	n/r (≥18[6])	264	43 (39-52)	0.71 (0.56-0.91) {VRd *vs.* Rd}	52	54
SWOG S0777 NCT00644228			Rd	n/r (≥18[7])	261	30 (25-39)		38	56
Mateos 2018	PFS	706	DaraVMP	71 (40-93)	350	NR	0.50 (0.38-0.65) {DaraVMP *vs.* VMP}	NR	16.5
ALCYONE NCT02195479			VMP	71 (50-91)	356	18.1 (16.5-19.9)		NR	

CI: confidence interval; N: number; itt: intention to treat; n/r: not reported; NR: not reached; PFS: progression-free survival; OS: overall survival; TTP: time to progression; EFS: event-free survival; ORR: overall response rate; RR: response rate; D: Dexamethasone; DI: Dexamethasone-Interferon alpha; M100: Melphalan 100; MD: Melphalan-Dexamethasone; MP: Melphalan-Prednisone; TD: Thalidomide-Dexamethasone; CTD: Cyclophosphamide-Thalidomide-Dexamethasone; CTD(a): Cyclophosphamide-Thalidomide-Dexamethasone (attenuated); Melphalan-Prednisone-Thalidomide/Melphalan-Prednisone-Thalidomide and Thalidomide maintenance (MPT/MPT-T); VD: Bortezomib-Dexamethasone; VTD: Bortezomib-Thalidomide-Dexamethasone; VMP: Bortezomib-Melphalan-Prednisone; VTP: Bortezomib-Thalidomide-Prednisone; VMPT-VT: Bortezomib-Melphalan-Prednisone-Thalidomide and Bortezomib-Thalidomide; CPR: Cyclophosphamide-Prednisone-Lenalidomide; Rd: Lenalidomide-Dexamethasone; Rd18: 18 cycles Lenalidomide-Dexamethasone; MPR: Melphalan-Prednisone-Lenalidomide; MPR-R: Melphalan-Prednisone-Lenalidomide and Lenalidomide maintenance; VRd: Bortezomib-Lenalidomide-Dexamethasone; DaraVMP: Daratumumab-Bortezomib-Melphalan-Prednisone; [1] 40% ≥70 years [2] 43% ≥70 years [3] 39% ≥70 years. [4] 49% ≥65 years. [5] 47% ≥65 years. [6] 38% ≥65 years. [7] 48% ≥65 years Source HR: from published trial (MM-003, Ludwig 2009, GIMEMA, MRC-MIX, GIMEMA0305, HOVON87, S0777, E1A06, ALCYONE, IFM-99/06, EMN01, FIRST), obtained from a previous patient-level meta-analysis[5] (IFM-01/01, NMSG, TMSG, HOVON49), from a previous NMA15 (Sacchi 2011) and data on file from investigators (Hungria 2016). Calculations were made using the published HR and *P* value (VISTA), Kaplan-Meier curves (IFM95/01 and the MM-15) and *P*-value and number of events (GEM2005, Upfront, s0232). Table 1 presents the extracted and calculated data.

effects model. Heterogeneity and inconsistency were assessed by decomposing the Q statistic[21,22] and quantified by the I^2-statistic,[23] which presents the percentage of the variability in effects due to heterogeneity rather than chance.[24]

Results

Systematic literature review

Figure 1 presents the PRISMA flow diagram; the PRISMA checklist is presented in *Online Supplementary Appendix 3*. The SLR identified a total of 19,773 citations from the databases. One additional recent record was included (i.e. the ALCYONE trial[7]). After removing duplicates, 18,752 citations remained. Based on title and abstract, 17,741 citations were excluded for further analysis. The full texts of 1011 citations were reviewed and, based on this assessment, 944 citations were excluded. In the second full text review of the remaining 67 citations, 43 citations were excluded because these did not report the most recent results (e.g. extended follow-up results were available). After the entire assessment, 24 RCTs remained and were included for data extraction and the NMA. See Figure 1 for further details of the reasons for exclusion.

These 24 RCTs included 21 treatment options: 1) Dexamethasone (D), 2) Dexamethasone-Interferon alpha (DI), 3) Melphalan 100 (M100), 4) Melphalan-Dexamethasone (MD), 5) Melphalan-Prednisone (MP), 6) Thalidomide-Dexamethasone (TD), 7) Cyclophosphamide-Thalidomide-Dexamethasone (CTD), 8) Cyclophosphamide-Thalidomide-Dexamethasone (attenuated) [CTD(a)], 9) Melphalan-Prednisone-Thalidomide/Melphalan-Prednisone-Thalidomide and Thalidomide maintenance (MPT/MPT-T), 10) Bortezomib-Dexamethasone (VD), 11) Bortezomib-Thalidomide-Dexamethasone (VTD), 12) Bortezomib-Melphalan-Prednisone (VMP), 13) Bortezomib-Thalidomide-Prednisone (VTP), 14) Bortezomib-Melphalan-

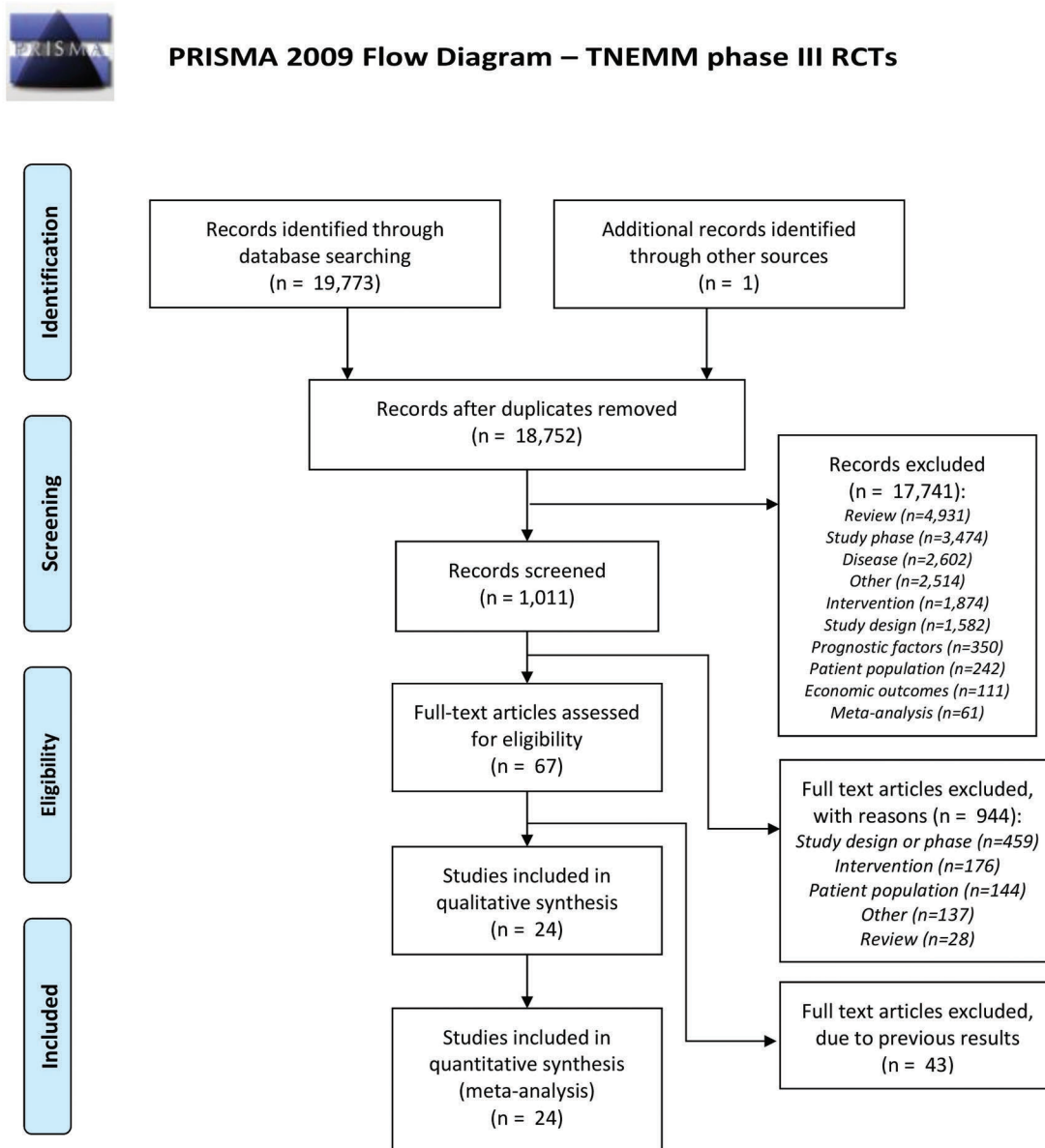## PRISMA 2009 Flow Diagram – TNEMM phase III RCTs

Figure 1. PRISMA 2009 flow diagram: transplant not eligible multiple myeloma (TNEMM) Phase III randomized controlled trials (RCTs). n: number. From Moher *et al*. 2009.[52]

Prednisone-Thalidomide and Bortezomib-Thalidomide (VMPT-VT), 15) Cyclophosphamide-Prednisone-Lenalidomide (CPR), 16) Lenalidomide-Dexamethasone (Rd), 17) 18 cycles Lenalidomide-Dexamethasone (Rd18), 18) Melphalan-Prednisone-Lenalidomide (MPR), 19) Melphalan-Prednisone-Lenalidomide and Lenalidomide maintenance (MPR-R), 20) Bortezomib-Lenalidomide-Dexamethasone (VRd), 21) Daratumumab-Bortezomib-Melphalan-Prednisone (DaraVMP).

Data extraction

Table 1 provides the details, and extracted and calculated data of the included trials. Most trials (21 out of 24) investigated iMID-based regimens (thalidomide or lenalidomide). Since MP has been the standard treatment for decades,[25] MP was the comparator in 12 trials. PFS was the primary end point for 13 trials. The median age of the patient population was reported by most trials and ranged from 64 to 79 years. While some trials included patients aged <65 years, either because of choosing broader age limits or because of including patients who were not eligible for SCT independent of age, most trials only included patients aged ≥65 years. The IFM99-06[26] and IFM01/01[27] only focused on patients aged ≥70 and ≥75, respectively.

Network meta-analysis network

All identified RCTs (n=24) and treatments (n=21) were incorporated within one network (Figure 2). We combined MPT and MPT-T. The duration of induction therapy with thalidomide var-ied leading to a clear overlap in planned thalidomide use between protocols with and without maintenance, preventing a clear discrimination between MPT with and without thalidomide maintenance.

Figure 2 presents the obtained HR(s) from the trial(s) and the HR obtained from the NMA for each of the connections (i.e. treatment comparisons) in our network. In order to validate our data, we compared the HR from treatments for which only direct evidence from a single RCT was available. The HR obtained from the NMA should be equal to the HR obtained from the RCT. The HR from the NMA was indeed similar to the HR from the trials for six comparisons[5-7,28-30] [i.e. CTD(a) vs. MP, VMP vs. MP, DaraVMP vs. VMP, VRd vs. Rd, VMPT-VT vs. VMP and VMP vs. VTP] (*Online Supplementary Appendix 5*). In addition, our network includes several treatments for which both direct and indirect evidence were available. *Online Supplementary Appendix 5* presents the HRs based on direct and indirect evidence and shows that none of the *P*-values for disagreement was lower than 0.05.

The percentage of the variability in effect estimates due to heterogeneity rather than sampling error (=I^2) was 72% indicating substantial between-study heterogeneity (i.e. within the 50-90% range can be quantified as substantial heterogeneity[24]). We allowed for between-study heterogeneity by using the random effects model. Heterogeneity could be reduced by excluding some of the trials; however, because of a lack of valid reasons (e.g. patients' characteristics, treatment dosing or follow up) for excluding trials, we decided not to perform analyses of this kind.

Figure 2. Network of the studies included in the network meta-analysis (NMA). White boxes represent treatments and reference numbers using the following abbreviations. 1) Dexamethasone (D); 2) Dexamethasone-Interferon alpha (DI); 3) Melphalan 100 (M100); 4) Melphalan-Dexamethasone (MD); 5) Melphalan-Prednisone (MP); 6) Thalidomide-Dexamethasone (TD); 7) Cyclophosphamide-Thalidomide-Dexamethasone (CTD); 8) Cyclophosphamide-Thalidomide-Dexamethasone (attenuated) [CTD(a)]; 9) Melphalan-Prednisone-Thalidomide / Melphalan-Prednisone-Thalidomide and Thalidomide maintenance (MPT/MPT-T); 10) Bortezomib-Dexamethasone (VD); 11) Bortezomib-Thalidomide-Dexamethasone (VTD); 12) Bortezomib-Melphalan-Prednisone (VMP); 13) VTP: Bortezomib-Thalidomide-Prednisone (VTP); 14) Bortezomib-Melphalan-Prednisone-Thalidomide and Bortezomib-Thalidomide (VMPT-VT); 15) Cyclophosphamide-Prednisone-Lenalidomide (CPR); 16) Lenalidomide-Dexamethasone (Rd); 17) 18 cycles Lenalidomide-Dexamethasone (Rd18); 18) Melphalan-Prednisone-Lenalidomide (MPR); 19) Melphalan-Prednisone-Lenalidomide and Lenalidomide maintenance (MPR-R); 20) Bortezomib-Lenalidomide-Dexamethasone (VRd); 21) Daratumumab-Bortezomib-Melphalan-Prednisone (DaraVMP). Black box represents the reference treatment in the network meta-analysis. Gray boxes include the trial reference and hazard ratio (HR) for progression-free survival on the top row(s). Bottom row shows HR according to the NMA. *HR not statistically significant at 5%.

Results *versus* dexamethasone

Figure 3 presents the HRs with the corresponding 95%CI for PFS and the *P*-score of the NMA in which dexamethasone was used as comparator for the remaining 20 "other treatment" options. HRs above 1 indicate that the "other treatment" is less effective than the comparator treatment, dexamethasone; HRs below 1 indicate that the "other treatment" is more effective than dexamethasone. All first-line NTE NDMM treatment options were better compared to the reference treatment dexamethasone (i.e. reducing the risk of progression or death compared to dexamethasone). HRs ranged between 0.19-0.90; however, not all treatments were statistically significantly different from dexamethasone, because of wide 95%CIs. DaraVMP and VMPT-VT were identified as the most effective treatment options as they had the highest and almost similar *P*-scores (i.e. a 96% and 93% certainty that this treatment is better than another treatment, averaged over all competing treatments) and most favorable relative treatment effects compared to dexamethasone (i.e. HR: 0.19, 95%CI: 0.08-0.45 and HR 0.22, 95%CI: 0.10-0.51 for DaraVMP and VMPT-VT, respectively). The HRs and 95%CIs for currently recommended treatments, VRd, VMP and Rd compared to dexamethasone, were 0.31 (95%CI: 0.16-0.59), 0.39 (95%CI: 0.20-0.75), and 0.44 (95%CI: 0.29-0.65), respectively. Selecting MPT as a reference treatment does not change the hierarchy of the treatments as the *P*-score values do not change if one considers a different reference treatment. Compared to MPT, only DaraVMP had a statistically lower HR for PFS (HR 0.41, 95%CI: 0.19-0.91; *P*<0.05) (*Online Supplementary Appendix 6*).

Scenario analysis network meta-analysis

In order to rule out that grouping of MPT and MPT-T would affect the outcome of the analysis, we performed a scenario in which we grouped IFM 01/01, IFM 99/06 and Sacchi *et al.* 2011, as MPT and GIMEMA, HOVON49, TMSG and NMSG as MPT-T. The MPT-T group was connected in the network to the MPT-T arm from the HOVON87 trial and the ECOG E1A06 trial. Overall, the results were comparable to the base case (*Online Supplementary Appendix 7*). We found similar results for MPT (HR 0.46, 95%CI: 0.30-0.71) and MPT-T (HR 0.47, 95% CI 0.30-0.73) compared to D.

The second scenario, based on the trials included by Weisel *et al.*,[12] showed lower HRs for PFS for Rd compared to VMP, MPT and MP, but the 95%CI for VMP overlapped with Rd [Rd *vs.* VMP: HR 0.73, 95%CI: 0.48-1.11 (*Online Supplementary Appendix 8*)].

Results from the third scenario analysis (fixed effect model instead of random effects model) are presented in *Online Supplementary Appendix 9*. While the HRs from the fixed effect model are quite similar, the 95%CIs are smaller, as is typical for fixed effect models.
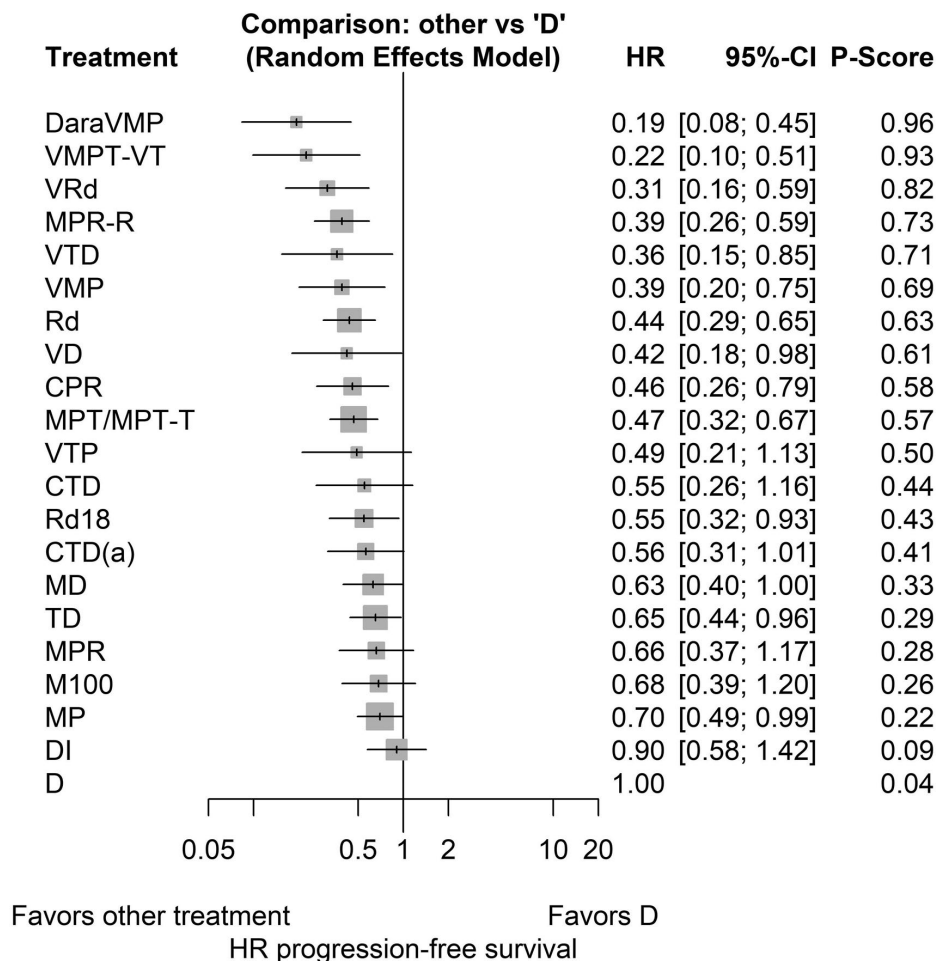

Treatment	Comparison: other vs 'D' (Random Effects Model)	HR	95%-CI	P-Score
DaraVMP		0.19	[0.08; 0.45]	0.96
VMPT-VT		0.22	[0.10; 0.51]	0.93
VRd		0.31	[0.16; 0.59]	0.82
MPR-R		0.39	[0.26; 0.59]	0.73
VTD		0.36	[0.15; 0.85]	0.71
VMP		0.39	[0.20; 0.75]	0.69
Rd		0.44	[0.29; 0.65]	0.63
VD		0.42	[0.18; 0.98]	0.61
CPR		0.46	[0.26; 0.79]	0.58
MPT/MPT-T		0.47	[0.32; 0.67]	0.57
VTP		0.49	[0.21; 1.13]	0.50
CTD		0.55	[0.26; 1.16]	0.44
Rd18		0.55	[0.32; 0.93]	0.43
CTD(a)		0.56	[0.31; 1.01]	0.41
MD		0.63	[0.40; 1.00]	0.33
TD		0.65	[0.44; 0.96]	0.29
MPR		0.66	[0.37; 1.17]	0.28
M100		0.68	[0.39; 1.20]	0.26
MP		0.70	[0.49; 0.99]	0.22
DI		0.90	[0.58; 1.42]	0.09
D		1.00		0.04

Favors other treatment   Favors D
HR progression-free survival

Figure 3. Results of the network meta-analysis in which dexamethasone was used as comparator. HR: hazard ratio; CI: confidence interval. For abbreviations for treatments, see Figure 2.

Discussion

Current clinical decision making in MM is complicated by a lack of head-to-head comparisons of standards of care, an increasing number of treatment modalities, and rapidly evolving promising results of studies with novel regimens (among smaller subpopulations). In this treatment landscape, we believe the role of NMA will become increasingly important, although it cannot replace RCTs, and these will still represent the gold standard.

Firstly, NMAs are able to provide data where head-to-head comparisons are lacking.[20,24] For NTE NDMM, there have been no head-to-head comparisons between the current three standard of care regimens (i.e. VRd, VMP and Rd). Only VRd has been compared head-to-head with Rd, and there are no studies comparing VMP with VRd or Rd. With our NMA, we show that the HR of VRd was lower than VMP and Rd, and VRd also had the highest *P*-score. We present similar HRs and *P*-scores for VMP and Rd. However, we also show considerable overlap of the 95%CIs of VRd, Rd and VMP. Our NMA does not support the use of one over the other regimens, thus leaving three valuable options for clinical practice. The choice of therapy will be guided by the patient's characteristics. For example, a proteasome inhibitor (PI)-based regimen in high-risk cytogenetic disease, and a preference for lenalidomide without bortezomib in patients with neuropathy.[31-34]

According to the ranking based on their *P*-scores and comparative effectiveness estimates, DaraVMP and VMPT-VT were identified as the most effective treatments. Although there is one RCT already showing better PFS and OS[28] for VMPT-VT when compared with VMP, we now add data showing comparable efficacy to DaraVMP, which is expected to become an important standard of care. This finding is important given the pronounced differences in global access to expensive treatment regimens. As all drugs in the VMPT-VT regimen will soon be available as generic compounds, this regimen is a valuable option in clinical practice as well. In addition, the pronounced efficacy of VMPT-VT highlights the use of maintenance therapy following PI-based induction regimens. In addition, in a non-head-to-head comparison with VMP, the study of the PETHEMA group showed that maintenance therapy did result in a substantially longer PFS.[35] We now add further evidence for maintenance therapy with PIs by showing high efficacy of VMPT-VT as compared to VMP. This is important because the European Medicines Agency has still not approved maintenance therapy with bortezomib, given that no head-to-head comparisons of maintenance *versus* no maintenance therapy have been made.

Secondly, NMAs provide more solid and precise effectiveness estimates when head-to-head data from multiple RCTs are available.[20,24] Our network included several trials investigating MPT/MPT-T *versus* MP. Some of these trials showed superiority of MPT/MPT-T over MP,[26,27,36] while other trials found no difference.[37-40] NMA enables this evidence to be synthesized and, according to our analysis, MPT/MPT-T was superior over MP (HR 0.67, 95%CI: 0.55-0.81).

Thirdly, NMA calculates effectiveness estimates including direct and indirect evidence from RCTs providing additional evidence when head-to-head data are only available from one single RCT. Due to the rapid evolution of the treatment armamentarium, efficacy evidence is increasingly based on a single RCT, and this is often from only one institute or region in the world. There is increasing evidence of RCTs investigating a similar treatment comparison providing contradictory results[41] and this may increase the interest in indirect evidence. Indirect evidence may confirm or alter the results from a single RCT, as we have shown for MPR-R compared to MPT. Although there was no statistically significant difference between MPR-R and MPT-T based on direct evidence from two RCTs, synthesizing direct and indirect evidence resulted in a statistically significant HR for MPR-R compared to MPT/MPT-T. Favorable indirect evidence for MPR-R compared to MPT-T was obtained through the comparison with MP. MPR-R compared more favorably to MP (according to the MM-15 HR MPR-R *vs.* MP 0.4) than MPT (HR MPT *vs.* MP 0.67 according to multiple trials). However, it should be noted that the direct evidence for MPR-R compared to MP was based on a single RCT while MPT/MPT-T *versus* MP was studied in seven RCTs, and therefore the evidence for the latter comparison is believed to be more solid.[24,41] Indirect evidence is not always available, for example, for the comparison between VRd and Rd there is only direct evidence from a single study.[6] While a fixed effect NMA will produce similar results to the trial (HR 0.71, 95%CI: 0.57-0.9), a random effects NMA obtains larger 95%CIs (HR 0.71, 95%CI: 0.43-1.17), as it includes two levels of uncertainty: within and between study variances.[17] Therefore, there is less likelihood of significant differences between treatments.

Two other NMAs are available for newly diagnosed NTE NDMM patients. Our results align with the results from Kuhr *et al.*[11] in that VMP and MPT are more effective than MP. Our results also confirm the conclusion from Weisel *et al.*[12] that Rd is more favorable than MP [HR 0.63, 95%CI: 0.44-0.89 (*Online Supplementary Appendix 5*)]. However, in contrast to their findings, we found that Rd and VMP have comparable effectiveness outcomes (i.e. small difference in HR for PFS compared to D but largely overlapping CIs). The primary analysis of Weisel *et al.* included a limited number of treatments (i.e. VMP, MP, MPT and Rd) and RCTs (i.e. VISTA, IFM01/01, IFM 99/06, Sacchi, FIRST) as Phase III trials not using dosing schemes in line with the summary of product characteristics (SmPC) were excluded. There are several arguments against this restriction. Firstly, although dosing schemes in line with the SmPC might be recommended in the selected trials by Weisel *et al.*, it is debatable whether this ensures treatments are identical within a network, especially because of variation in clinical practice due to either physician's preference or patient-related factors such as age, co-morbidities and toxicities. For example, the trial of Sacchi *et al.* 2011 was grouped with MPT studies while maintenance was only provided in a limited number of centers. Furthermore, the administered and planned dose may differ, as, for example, illustrated by the HOVON87 in which relative dose intensity varied between 0.54-0.96.[13] Since there is no evidence on the impact of dosing schemes, we believe that a more comprehensive network (e.g. our network, including 19 additional trials) provides more solid evidence. The reason Weisel *et al.*[12] did not find an overlap between VMP and Rd in their sensitivity analyses including six and twelve additional studies, is most likely because they used a fixed effect model for their analysis. A random effects model, like that used in our

analysis and by Kuhr *et al.*,[11] is, however, more appropriate, as this model allows for the between study-heterogeneity in the additional studies.

One might argue that while our NMA provides additional evidence in different circumstances, we had to make assumptions in order to conduct the analysis, and that this introduces a level of uncertainty. Firstly, we grouped MPT and MPT-T studies together since we could not make an unambiguous distinction between them. For example, thalidomide was prescribed until disease progression in the HOVON49 and GIMEMA trial but prescribed "continuously" for up to a maximum of 12 months in the TMSG trial. In the NMSG trial, it was even recommended to continue thalidomide maintenance until second relapse. However, most investigators discontinued thalidomide at first relapse. Prescription of thalidomide was also not consistent within a trial.[38] Sacchi *et al.*[38] described that, although planned, maintenance was only provided to 18% of the patients and in a limited number of centers. Their results, however, showed that there was no difference in PFS between maintenance and no-maintenance approaches.[42] Therefore, we believe that it is appropriate to combine these trials, as performed previously,[11,43] and the results of our sensitivity analysis confirm this assumption (see *Online Supplementary Appendix 7*).

Secondly, the validity of the outcomes of NMA depends on the comparability between studies. Our analysis focused on treatments for NTE NDMM patients studied in phase III RCTs. Although it is possible to include non-randomized evidence in NMA,[45] and this could have provided additional information regarding effectiveness in clinical practice[46-48] or treatments not analyzed in a phase III RCT (e.g. bortezomib-cyclophosphamide-dexamethasone, VCD49), we believe that limiting our analysis to the relative effectiveness of RCT evidence reduces the risk of bias and systematic errors.[44] Further research to improve methodologies for conducting, evaluating and interpreting non-randomized evidence is recommended.[44] We focused on NTE NDMM treatment to increase homogeneity between the patient populations in the study. We observed between-study heterogeneity comparable to the proportions previously reported by Kuhr *et al.*[11] We allow for this heterogeneity by using a random effects instead of a fixed effect model. However, the consequence of this is larger 95%CIs.

A potential limitation of our search strategy is that we only included English language publications. To the best of our knowledge, this does not, however, lead to the exclusion of relevant studies or treatments. Furthermore, our NMA was limited to the intermediate outcome PFS and did not include other outcomes of interest such as OS, adverse events, quality of life, cost, and cost-effectiveness. While OS may even be the most important subject of investigation for patients and health care decision makers, we believe a comparison of OS for first-line therapies with the currently available data is prone to bias due to crossover, heterogenous and limited follow up (e.g. especially for DaraVMP: median OS was not reached at 16.5 months follow up), and different subsequent treatment lines.[50,51] In the context of increasing health care expenditures, cost-effectiveness is also another relevant and important outcome, and this remains a subject for further research. Several treatment options showed comparable effectiveness outcomes, but costs could very well differ due to drug prices, treatment duration, and route of administration. Our study facilitates cost-effectiveness research on first-line NTE treatments.

The treatment armamentarium is rapidly increasing and evolving for NTE NDMM patients, and NMAs will, therefore, become increasingly important. We illustrate the additional value and evidence that can be provided. NMAs support evidence-based decision making and may help optimize treatment and outcomes of NTE NDMM patients in clinical practice.

References

1. Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, et al. Cancer incidence and mortality patterns in europe: Estimates for 40 countries in 2012. Eur J Cancer. 2013;49(6):1374-1403.
2. Mateos MV, San Miguel JF. How should we treat newly diagnosed multiple myeloma patients? Hematology Am Soc Hematol Educ Program. 2013;2013:488-495.
3. Moreau P, San Miguel J, Sonneveld P, et al. Multiple myeloma: ESMO clinical practice guidelines for diagnosis, treatment and follow-updagger. Ann Oncol. 2017;128(suppl_4):iv52-iv61.
4. Benboubker L, Dimopoulos MA, Dispenzieri A, et al. Lenalidomide and dexamethasone in transplant-ineligible patients with myeloma. N Engl J Med. 2014;371(10):906-917.
5. San Miguel JF, Schlag R, Khuageva NK, et al. Bortezomib plus melphalan and prednisone for initial treatment of multiple myeloma. N Engl J Med. 2008;359(9):906-917.
6. Durie BG, Hoering A, Abidi MH, et al. Bortezomib with lenalidomide and dexamethasone versus lenalidomide and dexamethasone alone in patients with newly diagnosed myeloma without intent for immediate autologous stem-cell transplant (SWOG S0777): A randomised, open-label, phase 3 trial. Lancet. 2017;389(10068):519-527.
7. Mateos MV, Dimopoulos MA, Cavo M, et al. Daratumumab plus bortezomib, melphalan, and prednisone for untreated myeloma. N Engl J Med. 2018;378(6):518-528.
8. Bothwell LE, Greene JA, Podolsky SH, Jones DS. Assessing the gold standard--lessons from the history of RCTs. N Engl J Med. 2016;374(22):2175-2181.
9. Gentile M, Magarotto V, Offidani M, et al. Lenalidomide and low-dose dexamethasone (rd) versus bortezomib, melphalan, prednisone (VMP) in elderly newly diagnosed multiple myeloma patients: A comparison of two prospective trials. Am J Hematol. 2017;92(3):244-250.
10. Neupane B, Richer D, Bonner AJ, Kibret T, Beyene J. Network meta-analysis using R: A review of currently available automated packages. PLoS One. 2014;9(12):e115065.
11. Kuhr K, Wirth D, Srivastava K, Lehmacher W, Hellmich M. First-line therapy for non-transplant eligible patients with multiple myeloma: Direct and adjusted indirect comparison of treatment regimens on the existing market in germany. Eur J Clin Pharmacol. 2016;72(3):257-265.
12. Weisel K, Doyen C, Dimopoulos M, et al. A systematic literature review and network meta-analysis of treatments for patients with untreated multiple myeloma not eligible for stem cell transplantation. Leuk Lymphoma. 2017;58(1):153-161.
13. Zweegman S, van der Holt B, Mellqvist UH, et al. Melphalan, prednisone, and lenalidomide versus melphalan, prednisone, and thalidomide in untreated mul-

tiple myeloma. Blood. 2016;127(9):1109-1116.

14. 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.

15. Higgins JP, Altman DG, Gotzsche PC, et al. The cochrane collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.

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

17. Borenstein M, Hedges LV, Higgins JP, Rothstein HR. A basic introduction to fixed-effect and random-effects models for meta-analysis. Res Synth Methods. 2010;1(2):97-111.

18. Rucker G. Network meta-analysis, electrical networks and graph theory. Res Synth Methods. 2012;3(4):312-324.

19. Rucker G, Schwarzer G. Reduce dimension or reduce weights? comparing two approaches to multi-arm studies in network meta-analysis. Stat Med. 2014;33(25):4353-4369.

20. Bhatnagar N, Lakshmi PV, Jeyashree K. Multiple treatment and indirect treatment comparisons: An overview of network meta-analysis. Perspect Clin Res. 2014;5(4):154-158.

21. Higgins JP, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: Concepts and models for multi-arm studies. Res Synth Methods. 2012;3(2):98-110.

22. Krahn U, Binder H, Konig J. A graphical tool for locating inconsistency in network meta-analyses. BMC Med Res Methodol. 2013;13:35.

23. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327(7414):557-560.

24. Higgins JP, Green S, eds. Cochrane handbook for systematic reviews of interventions version 5.1.0. [updated March 2011] ed. The Cochrane Collaboration; 2011. Available from www.handbook.cochrane.org. Available from www.handbook.cochrane.org.

25. Alexanian R, Haut A, Khan AU, et al. Treatment for multiple myeloma. combination chemotherapy with different melphalan dose regimens. JAMA. 1969;208(9):1680-1685.

26. Facon T, Mary JY, Hulin C, et al. Melphalan and prednisone plus thalidomide versus melphalan and prednisone alone or reduced-intensity autologous stem cell transplantation in elderly patients with multiple myeloma (IFM 99-06): A randomised trial. Lancet. 2007; 370(9594):1209-1218.

27. Hulin C, Facon T, Rodon P, et al. Efficacy of melphalan and prednisone plus thalidomide in patients older than 75 years with newly diagnosed multiple myeloma: IFM 01/01 trial. J Clin Oncol. 2009;27(22):3664-3670.

28. Palumbo A, Bringhen S, Larocca A, et al. Bortezomib-melphalan-prednisone-thalidomide followed by maintenance with bortezomib-thalidomide compared with bortezomib-melphalan-prednisone for initial treatment of multiple myeloma: Updated follow-up and improved survival. J Clin Oncol. 2014;32(7):634-640.

29. Morgan GJ, Davies FE, Gregory WM, et al. Long-term follow-up of MRC myeloma IX trial: Survival outcomes with bisphosphonate and thalidomide treatment. Clin Cancer Res. 2013;19(21):6030-6038.

30. Mateos MV, Richardson PG, Schlag R, et al. Bortezomib plus melphalan and prednisone compared with melphalan and prednisone in previously untreated multiple myeloma: Updated follow-up and impact of subsequent therapy in the phase III VISTA trial. J Clin Oncol. 2010;28(13):2259-2266.

31. Dimopoulos MA, Terpos E, Chanan-Khan A, et al. Renal impairment in patients with multiple myeloma: A consensus statement on behalf of the international myeloma working group. J Clin Oncol. 2010;28(33):4976-4984.

32. Larocca A, Offidani M, Musto P, et al. 744 impact of bortezomib- or lenalidomide-based induction treatment on high risk cytogenetic transplant-ineligible patients with newly diagnosed multiple myeloma enrolled in the gimema-MM-03-05 and EMN01 trials. Blood. 2017;130(Suppl 1):744.

33. Sonneveld P, Avet-Loiseau H, Lonial S, et al. Treatment of multiple myeloma with high-risk cytogenetics: A consensus of the international myeloma working group. Blood. 2016;127(24):2955-2962.

34. Terpos E, Kleber M, Engelhardt M, et al. European myeloma network guidelines for the management of multiple myeloma-related complications. Haematologica. 2015;100(10):1254-1266.

35. Mateos MV, Oriol A, Martinez-Lopez J, et al. Outcomes with two different schedules of bortezomib, melphalan, and prednisone (VMP) for previously untreated multiple myeloma: Matched pair analysis using long-term follow-up data from the phase 3 VISTA and PETHEMA/GEM05 trials. Ann Hematol. 2016;95(12):2033-2041.

36. Palumbo A, Bringhen S, Liberati AM, et al. Oral melphalan, prednisone, and thalidomide in elderly patients with multiple myeloma: Updated results of a randomized controlled trial. Blood. 2008;112(8):3107-3114.

37. Wijermans P, Schaafsma M, Termorshuizen F, et al. Phase III study of the value of thalidomide added to melphalan plus prednisone in elderly patients with newly diagnosed multiple myeloma: The HOVON 49 study. J Clin Oncol. 2010;28(19):3160-3166.

38. Sacchi S, Marcheselli R, Lazzaro A, et al. A randomized trial with melphalan and prednisone versus melphalan and prednisone plus thalidomide in newly diagnosed multiple myeloma patients not eligible for autologous stem cell transplant. Leuk Lymphoma. 2011;52(10):1942-1948.

39. Beksac M, Haznedar R, Firatli-Tuglular T, et al. Addition of thalidomide to oral melphalan/prednisone in patients with multiple myeloma not eligible for transplantation: Results of a randomized trial from the turkish myeloma study group. Eur J Haematol. 2011;86(1):16-22.

40. Waage A, Gimsing P, Fayers P, et al. Melphalan and prednisone plus thalidomide or placebo in elderly patients with multiple myeloma. Blood. 2010; 116(9):1405-1412.

41. Ioannidis JP. Why most published research findings are false. PLoS Med. 2005; 2(8):e124.

42. Mateos MV, Oriol A, Martinez-Lopez J, et al. Bortezomib, melphalan, and prednisone versus bortezomib, thalidomide, and prednisone as induction therapy followed by maintenance treatment with bortezomib and thalidomide versus bortezomib and prednisone in elderly patients with untreated multiple myeloma: A randomised trial. Lancet Oncol. 2010;11(10):934-941.

43. Fayers PM, Palumbo A, Hulin C, et al. Thalidomide for previously untreated elderly patients with multiple myeloma: Meta-analysis of 1685 individual patient data from 6 randomized clinical trials. Blood. 2011;118(5):1239-1247.

44. Efthimiou O, Mavridis D, Debray TP, et al. Combining randomized and non-randomized evidence in network meta-analysis. Stat Med. 2017;36(8):1210-1226.

45. Mohty M, Terpos E, Mateos MV, et al. Multiple myeloma treatment in real-world clinical practice: Results of a prospective, multinational, noninterventional study. Clin Lymphoma Myeloma Leuk. 2018; 18(10):e401-e419.

46. Schmitz S, Maguire A, Morris J, et al. The use of single armed observational data to closing the gap in otherwise disconnected evidence networks: A network meta-analysis in multiple myeloma. BMC Med Res Methodol. 2018;18(1):66.

47. Verelst SGR, Blommestein HM, de Groot S, et al. Long-term outcomes in patients with multiple myeloma: A retrospective analysis of the dutch population-based HAematological registry for observational studies (PHAROS). HemaSphere. 2018; 2(4):1.

48. Jimenez-Zepeda VH, Duggan P, Neri P, Tay J, Bahlis NJ. Bortezomib-containing regimens (BCR) for the treatment of non-transplant eligible multiple myeloma. Ann Hematol. 2017;96(3):431-439.

49. Arditi C, Burnand B, Peytremann-Bridevaux I. Adding non-randomised studies to a cochrane review brings complementary information for healthcare stakeholders: An augmented systematic review and meta-analysis. BMC Health Serv Res. 2016; 16(1):598.

50. Blommestein HM, Verelst SG, de Groot S, Huijgens PC, Sonneveld P, Uyl-de Groot CA. A cost-effectiveness analysis of real-world treatment for elderly patients with multiple myeloma using a full disease model. Eur J Haematol. 2015;96(2):198-208.

51. Zheng Y, Pan F, Sorensen S. Modeling treatment sequences in pharmacoeconomic models. Pharmacoeconomics. 2017; 35(1):15-24.

52. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med. 2009;6(7):e1000097.