

# Thalidomide for previously untreated elderly patients with multiple myeloma: meta-analysis of 1685 individual patient data from 6 randomized clinical trials

Peter M. Fayers,<sup>1,2</sup> Antonio Palumbo,<sup>3</sup> Cyrille Hulin,<sup>4</sup> Anders Waage,<sup>2,5</sup> Pierre Wijermans,<sup>6</sup> Meral Bektaş,<sup>7</sup> Sara Brinthen,<sup>8</sup> Jean-Yves Mary,<sup>9</sup> Peter Gimsing,<sup>10</sup> Fabian Termorshuizen,<sup>11</sup> Rauf Haznedar,<sup>12</sup> Tommaso Caravita,<sup>13</sup> Philippe Moreau,<sup>14</sup> Ingemar Turesson,<sup>15</sup> Pellegrino Musto,<sup>16</sup> Lotfi Benboubker,<sup>17</sup> Martijn Schaafsma,<sup>18</sup> Pieter Sonneveld,<sup>19</sup> and Thierry Facon,<sup>20</sup> on behalf of the Nordic Myeloma Study Group, Italian Multiple Myeloma Network, Turkish Myeloma Study Group, Hemato-Oncologie voor Volwassenen Nederland, Intergroupe Francophone du Myélome, and European Myeloma Network

<sup>1</sup>Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, United Kingdom; <sup>2</sup>Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway; <sup>3</sup>Divisione di Ematologia dell'Università di Torino, Azienda Ospedaliera Universitaria San Giovanni Battista, Torino, Italy; <sup>4</sup>Hematology Department, University Hospital, Nancy, France; <sup>5</sup>Department of Hematology, St Olavs University Hospital, Trondheim, Norway; <sup>6</sup>Department of Hematology, Haga Hospital, The Hague, The Netherlands; <sup>7</sup>Department of Hematology, Ankara University, Ankara, Turkey; <sup>8</sup>Divisione di Ematologia dell'Università di Torino, Azienda Ospedaliera Universitaria Senese Giovanni Battista, Torino, Italy; <sup>9</sup>Inserm U717, Université Paris 7, Hôpital Saint-Louis, Paris, France; <sup>10</sup>Department of Hematology, Rigshospitalet and University of Copenhagen, Copenhagen, Denmark; <sup>11</sup>Erasmus Medical Center and University, Rotterdam, The Netherlands; <sup>12</sup>Department of Hematology, Gazi University School of Medicine, Ankara, Turkey; <sup>13</sup>Cattedra e Divisione di Ematologia, Università Tor Vergata, Ospedale S. Eugenio, Rome, Italy; <sup>14</sup>Centre Hospitalier Universitaire, Nantes, France; <sup>15</sup>Department of Hematology, Skane University Hospital, Malmö, Sweden; <sup>16</sup>Department of Onco-Hematology, Istituti di Ricovero e Cura a Carattere Scientifico Centro di Riferimento Oncologico della Basilicata Rionero in Vulture, Potenza, Italy; <sup>17</sup>Onco-hématologie Centre Hospitalier Régional Universitaire, Hôpital Bretonneau, Tours, France; <sup>18</sup>Medical Spectrum Twente, Enschede, The Netherlands; <sup>19</sup>Erasmus Medical Center and University, Rotterdam, The Netherlands; and <sup>20</sup>Department of Hematology, Hôpital Claude Huriez, Lille, France

**The role of thalidomide for previously untreated elderly patients with multiple myeloma remains unclear. Six randomized controlled trials, launched in or after 2000, compared melphalan and prednisone alone (MP) and with thalidomide (MPT). The effect on overall survival (OS) varied across trials. We carried out a meta-analysis of the 1685 individual patients in these trials. The primary endpoint was OS, and progression-free survival (PFS) and 1-year response rates**

**were secondary endpoints. There was a highly significant benefit to OS from adding thalidomide to MP (hazard ratio = 0.83; 95% confidence interval 0.73-0.94,  $P = .004$ ), representing increased median OS time of 6.6 months, from 32.7 months (MP) to 39.3 months (MPT). The thalidomide regimen was also associated with superior PFS (hazard ratio = 0.68, 95% confidence interval 0.61-0.76,  $P < .0001$ ) and better 1-year response rates (partial response or better**

**was 59% on MPT and 37% on MP). Although the trials differed in terms of patient baseline characteristics and thalidomide regimens, there was no evidence that treatment affected OS differently according to levels of the prognostic factors. We conclude that thalidomide added to MP improves OS and PFS in previously untreated elderly patients with multiple myeloma, extending the median survival time by on average 20%. (*Blood*. 2011;118(5):1239-1247)**

## Introduction

In recent years, there has been considerable progress in the treatment of multiple myeloma. Thalidomide was demonstrated to be effective in patients with relapsing multiple myeloma in 1999.<sup>1</sup> This discovery stimulated a number of studies, which confirmed the initial results that approximately 30% of patients with relapsing multiple myeloma respond to thalidomide.<sup>2</sup> A further step was to test the effect of thalidomide in newly diagnosed multiple myeloma; and from the year 2000 onwards, 5 groups in Europe launched 6 randomized studies with similar design, aimed at patients who had not received previous treatment and were not eligible for high-dose therapy. An Italian randomized study with melphalan, prednisone, and thalidomide (MPT) versus melphalan and prednisone (MP) demonstrated a significant difference in progression-free survival (PFS),<sup>3</sup> but this did not translate into an overall survival (OS) benefit.<sup>4</sup> One French study demonstrated a significant difference in both PFS and OS in patients 65 to 75 years

of age,<sup>5</sup> and the same result was found in a second, parallel study of patients older than 75 years.<sup>6</sup> The Dutch/Belgian study has demonstrated differences in event-free survival, OS, and PFS.<sup>7</sup> A Nordic trial found small and nonsignificant differences in OS and PFS, although there was a significant difference in response during the first year.<sup>8</sup> Finally, a Turkish trial has been completed, demonstrating a significant response and an early but nonsignificant survival advantage after 6 months of MPT treatment.<sup>9</sup> Thus, studies have shown a favorable effect on response rates, but it is less clear whether this results in improved OS as only 4 trials demonstrated a significant advantage of MPT over MP. Similarly, a meta-analysis using only published summary data suggested a “trend toward improvement in OS.”<sup>10</sup>

To provide a definitive estimate of the survival benefits of MPT over MP, if any, we performed an individual patient meta-analysis based on data from all 6 completed trials. This analysis, based on

Submitted March 14, 2011; accepted May 30, 2011. Prepublished online as *Blood* First Edition paper, June 13, 2011; DOI 10.1182/blood-2011-03-341669.

The publication costs of this article were defrayed in part by page charge

payment. Therefore, and solely to indicate this fact, this article is hereby marked “advertisement” in accordance with 18 USC section 1734.

© 2011 by The American Society of Hematology

individual data from 1685 patients, also allows us to explore possible reasons for the heterogeneity in the reported results of the separate studies, including the impact of interactions of patient characteristics with treatment.

## Methods

### Objectives

The primary aim of this meta-analysis is to examine the effect of thalidomide on patients with newly diagnosed multiple myeloma, using the individual patient data from the 6 recently completed trials that compared MPT versus MP. OS, defined as the time from randomization until death from any cause, was the primary endpoint, with the hypothesis that thalidomide improves OS. PFS and best response at 12 months were defined as secondary outcomes.

Secondary aims were prespecified: To explore with interaction tests the hypotheses that thalidomide benefits only patients with good performance status and that thalidomide has the same effect irrespective of renal failure defined by elevated serum creatinine levels.

### Data extraction

Individual patient data were obtained for the 6 trials of the 5 collaborating groups. Best response was defined as being the best recorded level of response within the first 12 months, and patients were classified into: dead within 12 months or, for those surviving one year or longer, very good partial response or better (VGPR), partial response (PR), and no response reported. The definition of complete response varied, and data were not available for all studies, so complete response was lumped with VGPR. All trials received institutional review board approval from their respective institutions.

Based on the literature,<sup>11,12</sup> the following factors, which were available across studies, were also collected for each patient to investigate study heterogeneity and for consideration as important risk factors that might account for study-to-study variation in outcomes: International Staging System (ISS),<sup>11</sup> Durie-Salmon Staging (DS stage),<sup>12</sup> Eastern Cooperative Oncology Group/World Health Organization performance status (WHO-PS),<sup>13</sup> serum  $\beta_2$ -microglobulin, serum creatinine, serum calcium, serum lactate dehydrogenase (LDH), and hemoglobin. Age and sex were also recorded.

### Statistical analysis

All randomly assigned patients were included on an intention-to-treat basis. Living patients were censored at the date they were last confirmed to be alive. All analyses were stratified by study. "Forest plots" of estimated hazard ratio (HR) and 95% confidence intervals (CIs) were used for the basic analysis. Treatment effect was estimated from survival curves<sup>14</sup> with HR estimated from proportional odds Cox models.<sup>15</sup> Survival curves are simple (nonstratified) Kaplan-Meier curves.<sup>14</sup>

It was anticipated that survival might be affected by baseline heterogeneity as the studies used differing patient eligibility and selection criteria. Heterogeneity was assessed and tested by  $I^2$ ,<sup>16,17</sup> which measures inconsistency (percentage of total variation across studies resulting from heterogeneity) of effects. It was specified that, in the event of heterogeneity, a random effect model<sup>17,18</sup> would be used, as described in the *Cochrane Handbook for Systematic Reviews of Interventions*,<sup>19</sup> although this approach remains controversial.<sup>20</sup> The risk factors were used to stratify HR forest plots of the MPT versus MP, with the continuous factors being categorized at usual thresholds ( $\beta_2$ -microglobulin < 3.5 mg/L and > 5.5 mg/L, LDH < 300 U/L, creatinine < 176  $\mu$ M, albumin < 35 g/L, and hemoglobin < 8.5 g/dL or  $\geq$  10.5 g/dL).<sup>11,12</sup> The importance of risk factors on survival was also explored using Cox models, and their impact on treatment effect was examined using interaction tests. For

the Cox models, continuous factors were explored as continuous; for  $\beta_2$ -microglobulin and creatinine, a logarithmic transformation reflecting a log-linear or percentage-change relationship was found to be appropriate. It was anticipated that the residual variation and study-to-study heterogeneity might be reduced by allowing for important prognostic factors.

All statistical tests were 2-sided, and  $P$  values < .05 were considered to be statistically significant for the main analyses. When exploring covariates and interactions, a more conservative  $P$  < .01 was used as an informal protection against including an excess of false-positive covariates and treatment interactions on outcome through multiplicity of significance tests. Analyses were carried out in STATA, Version 11,<sup>21</sup> with the stratified random effects models implemented using a shared frailty extension to the Cox model.<sup>22</sup> The proportional hazards assumption was confirmed by visual inspection of survival plots and complementary log-log plots for individual trials. The analyses were effected by analyzing the individual patient data in a one-stage analysis (Cox regressions) and also using 2-stage analyses in which summary HRs were first estimated for each subgroup and then in the second stage combined with random effects analyses of forest plots using the STATA "metan" macro.<sup>23</sup>

## Results

### Description of trials

The 6 trials are shown in Table 1. A total of 1685 patients were recruited to the combined trials and randomized between MP and MPT. The trials vary with respect to treatment policy (dose and the number of cycles/duration of thalidomide, and similarly for MP), and study design (only Intergroupe Francophone du Myélome-II [IFM-II] and Nordic Myeloma Study Group [NMSG] were blinded, with placebo added to MP in the control arm). Trials also varied in the selection of patients (WHO-PS, stage of disease, and age range). The characteristics of the recruited patients (Table 2) reflect the impact of these selection criteria on stage and performance status: for example, 31% in IFM-1 were ISS stage I, but only 16% in NMSG; 52% in Italian Multiple Myeloma Network (GIMEMA) had WHO-PS of 0, compared with 7% in Turkish Myeloma Study Group (TMSG). These are well recognized prognostic factors for survival, and the impact of this baseline heterogeneity is explored in the survival analyses.

### OS

A total of 920 deaths (55%) have been observed in the 1685 patients. Information about date of death or censoring was available for all patients, and the OS curves are shown in Figure 1. Median survival time on MP was 32.7 months (95% CI, 30.5-36.6 months), and on MPT 39.3 months (95% CI, 35.6-44.6 months). The absolute increase in 2-year survival was 5.1%, with 63.7% on MP alive at 2 years (95% CI, 60.3%-66.9%) and 68.8% on MPT (95% CI, 65.4%-71.9%); larger absolute differences were observed at 3 and 4 years. Using an unadjusted Cox model, stratified by study, the estimated HR is 0.83 (95% CI, 0.73-0.94,  $P$  = .004), representing a statistically significant benefit to MPT. However, Figure 1 also suggests that there may be early deaths in the MPT arm, with no benefit OS benefit in the first year.

Figure 2 summarizes the HRs by study and indicates significant trial-to-trial heterogeneity ( $I^2$  = 61.3%,  $P$  = .024). This heterogeneity may jeopardize the  $P$  value and CI of the estimated overall effect (0.82). When the prognostic factors were explored individually as potential covariates using Cox models, ISS, DS stage, WHO-PS,  $\beta_2$ -microglobulin, hemoglobin, albumin, creatinine, LDH, and age were all important risk factors for survival; however,

**Table 1. Characteristics of the trials**

Characteristic	GIMEMA	HOVON	IFM-I	IFM-II	NMSG	TMSG
Trial label	GIMEMA	HOVON49	IFM99-06	IFM01/01	NMSG12	TMSG
Country	Italy	The Netherlands/Belgium	France	France	Nordic	Turkey
No. of patients	331	333	321	229	357	114
Inclusion year	2002-05	2002-07	2000-05	2002-06	2002-07	2006-09
Age, y	> 65	> 65	65-75	> 75	> 65	> 55
SD stage	II, III	Ib, II, III	II, III, and I high-risk	II, III, and I high-risk	I-III symptomatic	I-III symptomatic
WHO status	0-4	0-3	0-4	0-4	0-4	0-2
Placebo	No	No	No	Yes	Yes	No
Melphalan	4 mg/m <sup>2</sup>	0.25 mg/kg	0.25 mg/kg	0.20 mg/kg	0.25 mg/kg	9  mg/m <sup>2</sup>
Prednisone	days 1-4 40 mg/m <sup>2</sup> days 1-7	days 1-5 1 mg/kg days 1-5	days 1-4 2 mg/kg days 1-4	days 1-4 2 mg/kg days 1-4	days 1-4 100 mg days 1-4	days 1-4 60 mg/m <sup>2</sup> days 1-4
No. of cycles/interval (weeks)	6/4	8/4	12/6	12/6	Until plateau/6	8/6
Thalidomide, mg/day	100	200	100-400*	100	200-400	100
Duration	Until relapse	8 cycles	12 cycles	12 cycles	Until relapse†	8 cycles
Crossover to MPT from MP	No	No	No	No	No	18%
Second-line thalidomide after relapse, % of MP patients	34‡	57	38	45	45	7

HOVON indicates Hemato-Oncologie voor Volwassenen Nederland.

\*Recommended dose was 200 mg/day with rapid increase if well tolerated, up to maximum 400 mg/day.

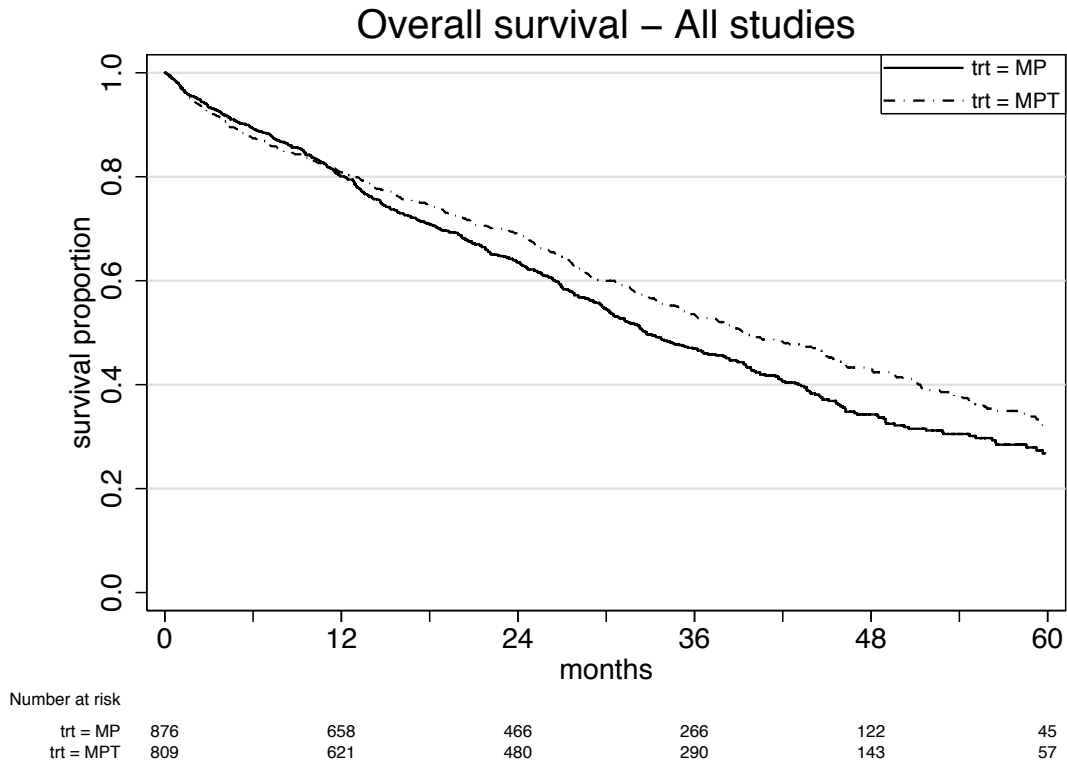
†A total of 200 mg/day from plateau until relapse.

‡Second-line thalidomide or bortezomib.

**Table 2. Patient characteristics**

Characteristic	GIMEMA (n = 331)	HOVON (n = 333)	IFM-I (n = 321)	IFM-II (n = 229)	NMSG (n = 357)	TMSG (n = 114)
Median age, y	72.5	72.6	69.4	78.5	74.4	70.6
<b>Sex, %</b>						
Male	54	56	54	45	55	54
Female	46	44	46	55	45	46
<b>DS stage, %</b>						
DS I	0	0	10	9	15	5
DS II	38	26	26	23	36	53
DS III	62	74	64	68	45	40
Not known	0	0	0	0	4	2
<b>ISS, %</b>						
ISS 1	19	25	31	22	16	18
ISS 2	38	24	34	38	40	35
ISS 3	23	18	27	28	33	38
Not known	20	32	8	12	11	10
<b>WHO-PS, %</b>						
0	52	36	24	19	15	7
1	24	46	50	45	23	43
2	19	14	19	28	28	46
3	5	4	6	7	23	4
4	1	0	2	0	7	0
Not known	0	0	0	0	4	1
<b>Ig heavy chains, %</b>						
IgG	64	60	67	58	62	68
IgA	25	30	21	28	21	15
IgD	0	0	0	0	1	3
Light only	11	10	10	12	8	2
Other	0	0	2	0	0	1
Not known	0	0	0	2	8	11
<b>Ig light chains, %</b>						
K	0	62	65	65	50	52
L	0	38	32	34	31	41
Not known	100	0	2	2	19	7
Median β <sub>2</sub> -microglobulin, mg/L	3.7	4.0	3.9	4.3	4.2	5.0
Median hemoglobin, g/dL	10.6	10.5	10.5	10.6	10.6	9.8
Median albumin, g/L	36.0	34.0	38.0	38.0	33.0	36.0
Median creatinine, μM	70.7	96.0	92.5	90.0	94.5	97.3
Median LDH, U/L	NR	293.0	337.0	331.5	223.5	238.5

HOVON indicates Hemato-Oncologie voor Volwassenen Nederland; and NR, not recorded.



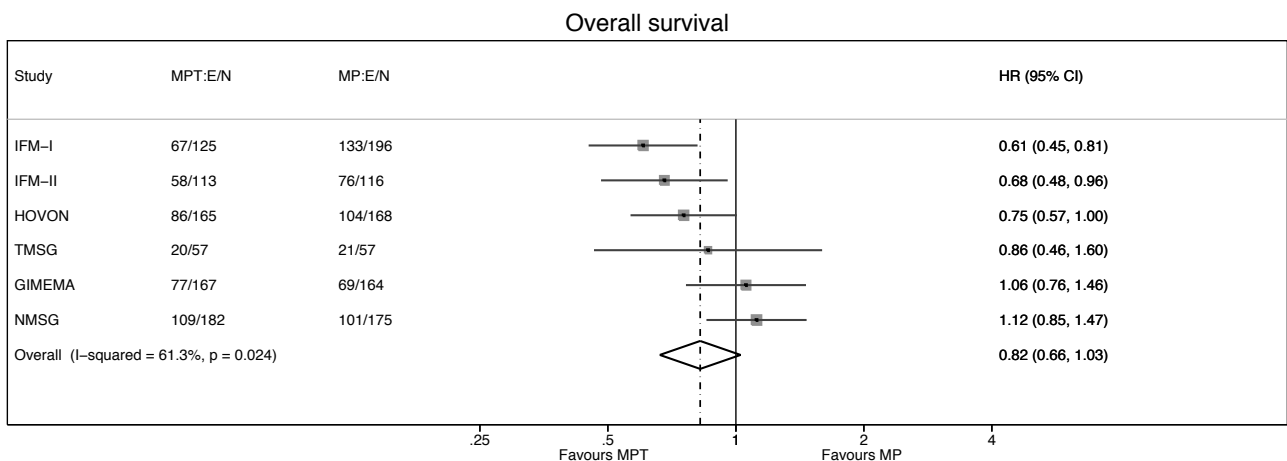
**Figure 1. OS: all patients, by treatment.** Cox model for treatment, with analysis stratified by study using a random effects (frailty) model: HR = 0.83 (95% CI, 0.73-0.94) in favor of MPT;  $P = .004$ . Median survival: MP 32.7 months (95% CI, 30.5-36.6 months); MPT 39.3 months (95% CI, 35.6-44.6 months).

not all factors were independent, and a multivariate Cox model resulted in retaining ISS, DS stage I or II versus III, WHO-PS, age, creatinine, and albumin (all  $P < .007$ ). These baseline covariates appeared to explain the heterogeneity that was observed in the unadjusted model. The covariate-adjusted model confirmed the estimated treatment effect (HR = 0.79,  $P = .002$ , 95% CI, 0.68-0.92,  $P = .002$ ). Additional Cox models were used to explore treatment-covariate interactions, but no clear evidence of interactions was found (all  $P > 0.036$ ). Thus, for example, ISS stage is one of the most important prognostic factors. Figure 3 shows HRs for treatment effect in each study, divided by ISS;

the values of HR are similar for each stage, and no interaction effect is apparent. Figure 4 summarizes the effects of each prognostic factor on OS.

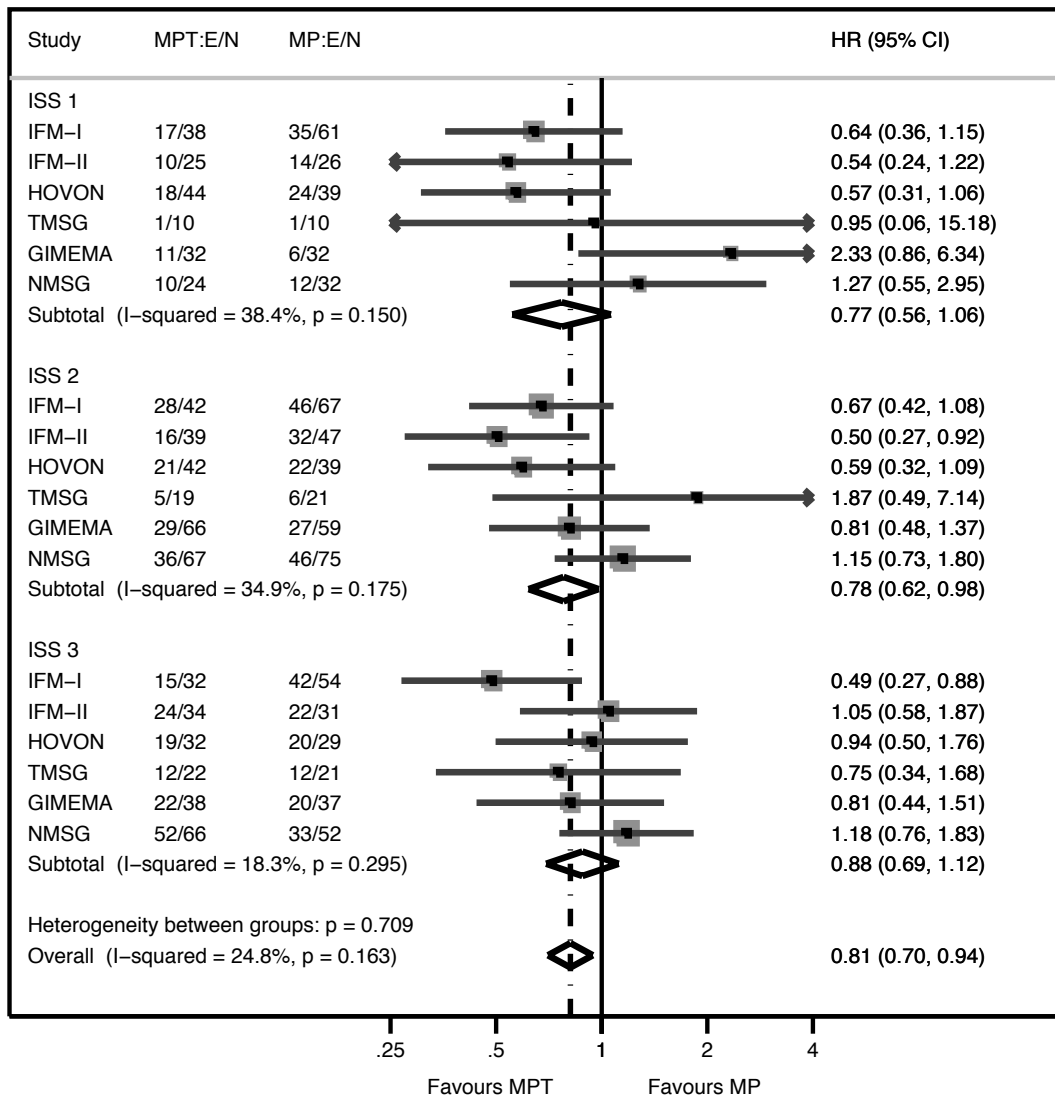
**PFS**

Figures 5 and 6 show corresponding results for PFS. The estimated HR was 0.68 in favor of MPT, with 95% CI 0.61 to 0.76 (Figure 5;  $P < .0001$ ). Median PFS time for MP was 14.9 months (95% CI, 14.0-16.6 months) and for MPT 20.3 months (95% CI, 18.8-21.5 months). Two-year PFS was 28.4% (MP) and 42.5% (MPT). As before, there was evidence of study-to-study



**Figure 2. Forest plot: OS HR of MPT versus MP, by study.** The HR and 95% CI are shown for each study, and the overall HR (where the dotted line bisects the diamond) and 95% CI (indicated by left and right corners of the diamond), assuming a random effects model. For each treatment arm of each study, the number of events (deaths) and number of patients (E/N) is given. There is evidence of heterogeneity ( $I^2 = 61\%$ ) that may jeopardize the  $P$  value and CI of the estimated overall effect (0.82).

### Overall survival by ISS stage



**Figure 3. Forest plot: OS HR of MPT versus MP, for each ISS stage subdivided by study.** The 95% CIs are shown. Diamonds represent the HR and 95% CI for each stage and finally for the overall estimate, assuming a fixed effects model. For each treatment arm of each study, the number of events (deaths) and number of patients (E/N) is given.

heterogeneity (Figure 6), and this could be accounted for by a combination of heterogeneity in the baseline characteristics (especially ISS). After allowing for this heterogeneity, the overall HR was 0.68 with 95% CI 0.56 to 0.81 ( $P < .0001$ ), confirming the Cox model. The interaction tests again showed no evidence that baseline characteristics might interact with treatment effects, apart from a significant treatment interaction ( $P = .006$ ) according to creatinine level  $>$  or  $< 176\mu\text{M}$ .

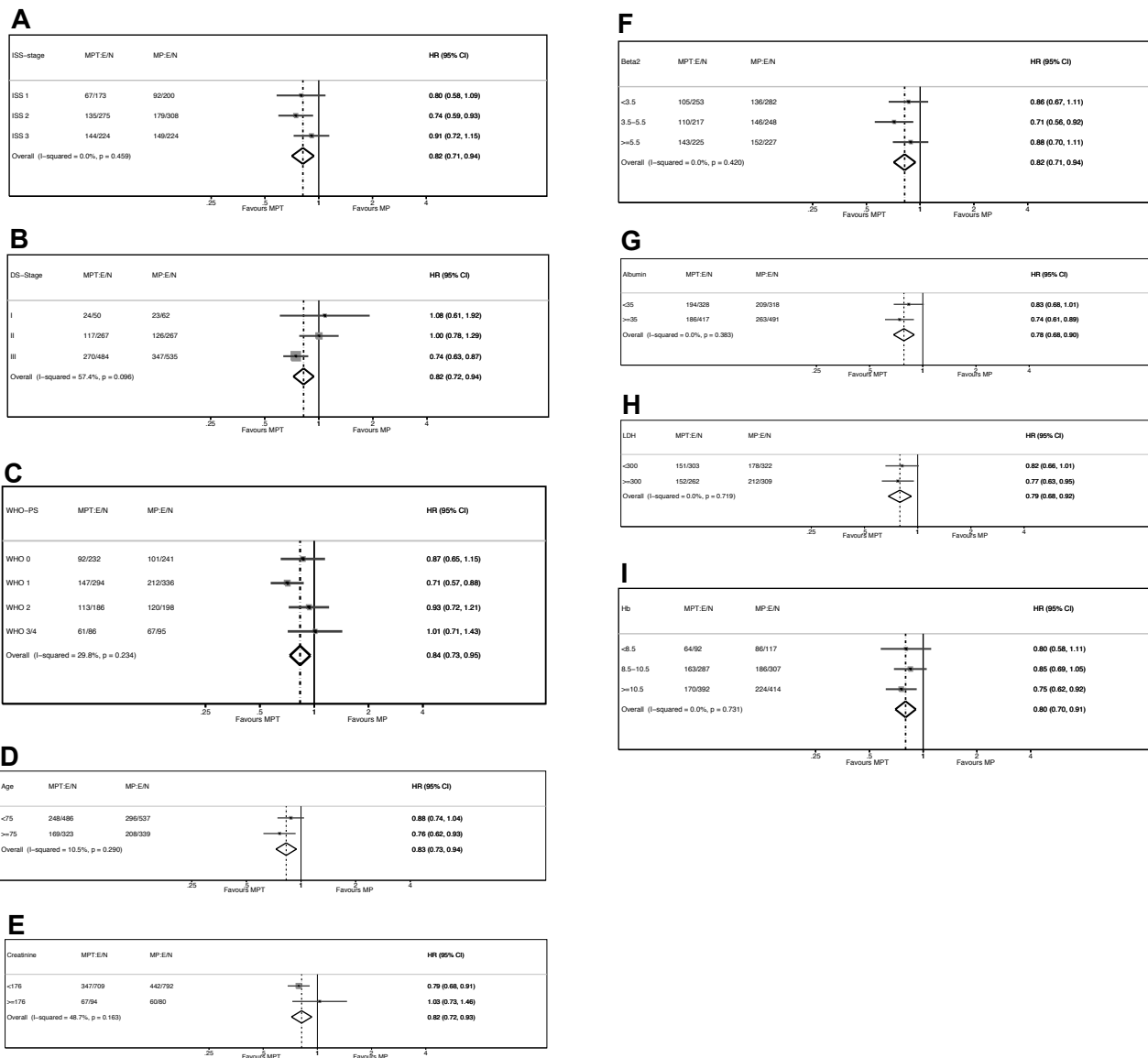
Survival from date of reported disease progression was also analyzed; the HR between treatment arms was 1.02, suggesting no shortening of survival after relapse.

#### Subgroup analyses

Two prespecified subgroup analyses were explored. The first hypothesis was that MPT only benefits patients with good WHO-PS of 0 to 2. As already noted, the test for interaction between OS and WHO-PS was nonsignificant, and there was no

evidence of variation of treatment effect across categories. However, the sample size for poor PS (3 or 4) was small, and only 2 trials exceeded 20 patients: NMSG, 107 patients; IFM-1, 23 patients. The HR for these patients was 1.01 (95% CI, 0.71-1.43), which, although not statistically significant, remains consistent with the prior hypothesis of no survival benefit from MPT if PS is poor.

The second hypothesis was that creatinine levels of  $176\mu\text{M}$  or greater at baseline would not affect any MP/MPT difference. Again, there was no significant interaction detectable for OS, and the forest plot showed that, for creatinine  $< 176$ , the HR = 0.79 (95% CI, 0.68-0.91,  $P = .001$ ); and for creatinine  $\geq 176$ , HR = 0.96 (95% CI, 0.73-1.46,  $P = .8$ ). However, for PFS, there was a significant interaction with creatinine as noted earlier:  $< 176$ , HR = 0.64 (95% CI, 0.55-0.74,  $P < .0001$ ), but at  $\geq 176$ , there was no evidence of PFS advantage with HR = 0.87 (95% CI, 0.54-1.39,  $P = .6$ ). Thus, the benefit of MPT in terms of PFS could be questioned for patients with baseline creatinine above this threshold.



**Figure 4.** Forest plot. OS HR of MPT versus MP for ISS (A), DS stage (B), WHO-PS (C), and age (D), creatinine (E),  $\beta_2$ -microglobulin (Beta2; F), albumin (G), LDH (H), and hemoglobin (Hb; I). E/N is the ratio of events (deaths) to the number of patients.

**Response in first year**

The best response reported in the first 12 months is shown as a percentage for VGPR and PR (Table 3). In all studies, there was a higher percentage of VGPR in the MPT group compared with MP (not shown), and overall VGPR was reported for 25% of MPT

patients versus 9% MP; PR or better was 59% MPT and 37% MP (overall  $\chi^2$  test,  $P < .0001$ ).

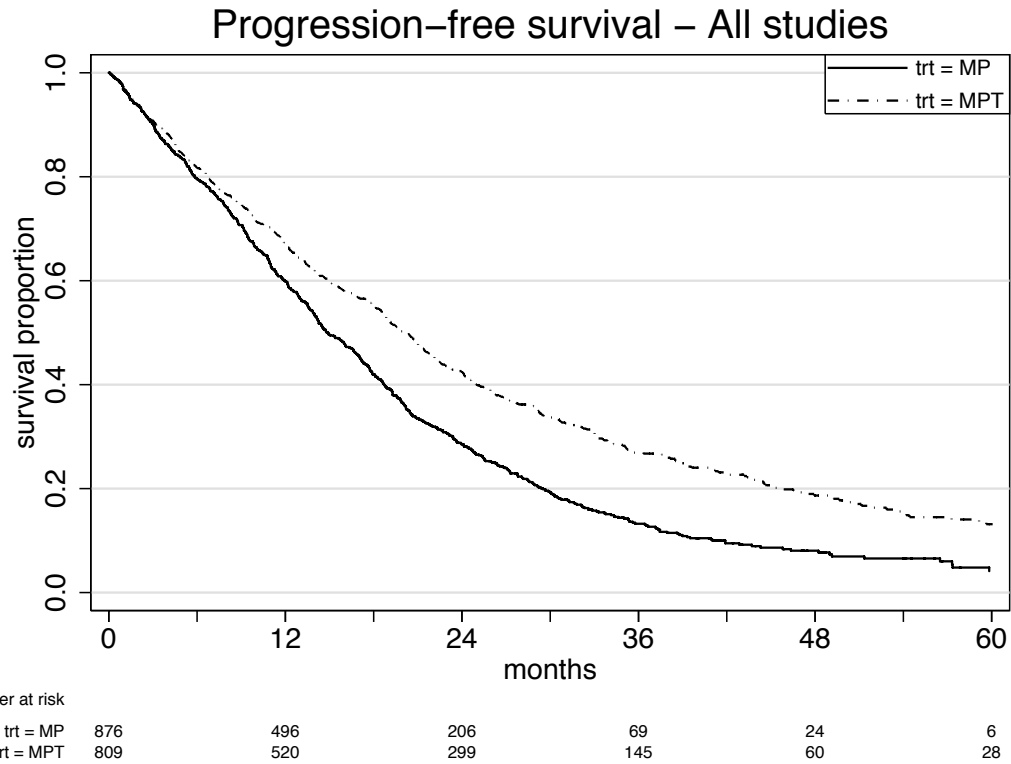
**Table 3. Best response recorded in first year, by treatment (percentage of patients)**

Study	MPT	MP	Total
No. of patients	815	870	1685
<b>Best response, %</b>			
VGPR or better	25	9	17
PR	34	28	31
No response	23	44	34
Death 1 year	17	19	18

Response was unknown for 46 (3%) of patients living at 1 year (27 MPT, 19 MP), and these were assumed to be “no response.” Overall  $\chi^2$  (3 degrees of freedom):  $P < .0001$ .

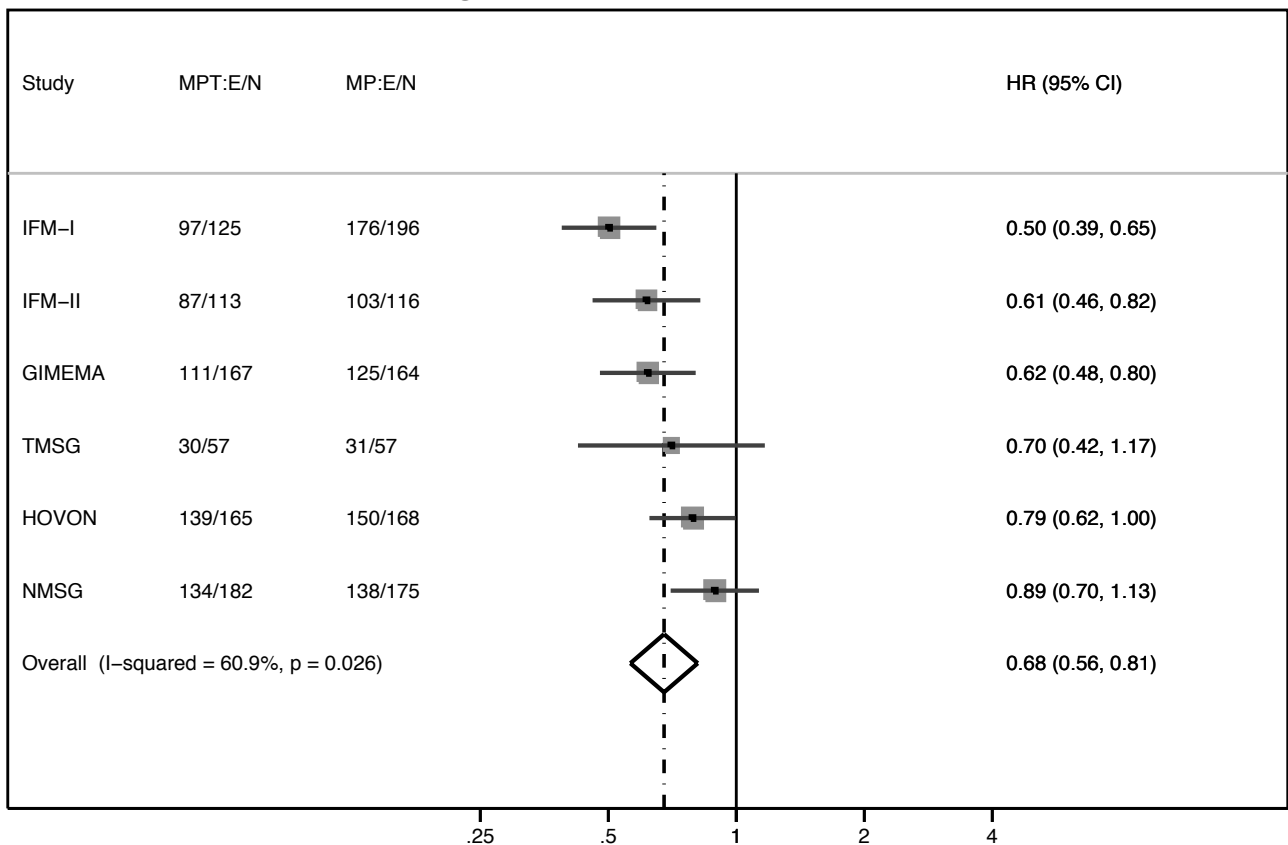
**Discussion**

Following this meta-analysis, the effect of thalidomide is now well established based on 6 studies. After due allowance was made for baseline characteristics, there was clear evidence of a benefit to patients receiving thalidomide, both in terms of OS and PFS. These results were also supported by the response to therapy that was observed in the first 12 months. The estimated HR for treatment effect on OS was 0.81, and highly significant. This represents an extra 6.6 months in median survival time, corresponding to an approximate 20% increase, which is a substantial and clinically important increase in survival. The variation in basic characteristics of the studies, as discussed later in “Discussion,” may have contributed to the substantial variation in results. This reflects



**Figure 5. PFS: all patients, by treatment.** Cox model for treatment, stratified random effects (frailty) model for study: HR = 0.68 (95% CI, 0.61-0.76);  $P < .0001$ . Median survival: MP 14.9 months (95% CI, 14.0-16.6 months); MPT 20.3 months (95% CI, 18.8-21.5 months).

### Progression-free survival



**Figure 6. Forest plot: PFS HR of MPT versus MP, by study.** The HR and 95% CI are shown for each study, and the overall HR (where the dotted line bisects the diamond) and 95% CI (indicated by left and right corners of the diamond), assuming a random effects model. For each treatment arm of each study, the number of events (deaths) and number of patients (E/N) is given. Overall test of HR = 1:  $P < .0001$ .

how thalidomide is used in real life and should be taken into consideration.

This meta-analysis highlights the value of having individual patient data, which enables a full exploration of the causes of heterogeneity. This in turn enables more precise estimation of HRs with narrower CIs for the estimates and more convincing *P* values. An analysis only using published summary results from the trials,<sup>10</sup> or “aggregate data,” cannot explore in full the impact of adding thalidomide to MP. We observed significant heterogeneity among the trials, but this could be largely accounted for as a consequence of major differences in baseline characteristics. Levels of the prognostic factors differed substantially from trial to trial; and in particular, both stage of disease and performance status showed major variations. Thus, ISS sufficed to account for much of the study-to-study heterogeneity. However, improvement in survival (OS and PFS) for the MPT arm appeared to be unaffected by level of ISS. LDH, which was not retained as an important prognostic factor in the multivariate model, varied substantially; in part, this variation in LDH (and other parameters) might be the result of variation across the many laboratories in the 6 countries.

We failed to find indication that poor performance status (WHO-PS > 2) contraindicates the use of MPT. However, the observed HRs were consistent with there being no advantage to MPT in these patients, although the sample size was limited and one trial dominated this subgroup analysis. Thus, despite the large combined sample of 1685 patients, the interaction test was underpowered, and the results regarding frail patients remain inconclusive. Similarly, indications against using MPT in patients with high creatinine levels were weak for OS, although the HRs remained consistent with there being less or no advantage. For PFS, however, there was a significant treatment interaction indicating less benefit in patients thought to have potential renal failure. Because patients presenting with renal failure do notably worse with many treatments, we recommend caution in using MPT for these patients.

An optimal dose of thalidomide with MP is not established. In the studies that introduced thalidomide in relapsing myeloma, the intended dose was 800 mg, which later was shown to be too high.<sup>1</sup> The 6 trials used a variety of dosage levels and cycles, both for MP and MPT. The French and Turkish trials used MP and thalidomide for 12 months, whereas the other trials continued thalidomide until relapse. However, none of the studies was designed to detect dose-effect relationships, and all used simple MP/MPT randomizations. In practice, clinicians would vary the actual dose according to patient's status, evidence of response or relapse, and occurrence of side effects or toxicity. A substantial proportion of patients in all studies either stopped thalidomide prematurely or had dosage reduced. Thus, it would be difficult to analyze the individual patient data for dose received and infeasible to eliminate bias from the consequent analysis. Therefore, we decided to rely on exploration of the forest plots (such as Figure 1) to determine whether studies with a policy of high dose differed from those with low dose. No such difference is obvious.

Similarly, inspection of the forest plots revealed no obvious association of effect with use of double-blind placebo.

Other novel agents, such as bortezomib and lenalidomide, have also been introduced in combination with MP in first-line treatment.<sup>24-26</sup> As with thalidomide, the combination of MP plus bortezomib increases remission rates, PFS, and OS.<sup>24</sup> MP plus lenalidomide increases remission rates and PFS, but an increase in OS has not yet been demonstrated.<sup>25</sup> However, the observation period is still too short to evaluate this fully. The MP combinations with thalidomide, bortezomib, or lenalidomide are all effective treatment options but have not been compared with each other, and a strict priority cannot be given.

The main limitation of this meta-analysis, as with any overview, is that the patient population as well as dose and duration of treatment varied across studies, resulting in heterogeneity as discussed. Sample size remained too small for reliable subgroup analyses. This paper focuses on the impact of MPT for OS, PFS, and response to therapy. Increased OS or PFS has to be matched against toxicity and side effects, which are being analyzed by another team within our collaborative group and will be reported separately. It has been suggested that MPT patients may have double the rate of grade 3 or 4 adverse events,<sup>27</sup> which may also explain the early deaths and lack of first-year OS benefit.

In conclusion, this meta-analysis of 6 trials demonstrated that thalidomide added to MP improves survival in previously untreated elderly patients with multiple myeloma, extending the median OS time by 20%. There was no evidence of benefit (nor of lack of benefit) to patients with poor PS. However, we advise against thalidomide for patients with potential renal failure. More MP-MPT trials are unlikely to be conducted, and it is important to reach a definitive conclusion based on all available trial data. These analyses show that MP alone should no longer be the reference and that MP plus thalidomide is an effective first-line treatment of multiple myeloma.

## Authorship

Contribution: P.M.F. was responsible for the data analysis and wrote the manuscript; and all authors were the principal investigators and major contributors or statisticians for respective trials, participated in the elaboration of the meta-analysis statistical plan, and made important contribution to the preparation and revision of the manuscript.

Conflict-of-interest disclosure: A.P. has honoraria and an advisory role for Celgene. S.B. has honoraria from Celgene. C.H., P. Moreau, and T.F. have honoraria or an advisory role for Celgene, Pharmion, and Janssen Cilag. A.W. and P.G. have an advisory role for Janssen Cilag, Celgene, and Mundipharma. M.B. has honoraria from and is a member of the Speakers Bureau for Celgene and Ortho Biotech. The remaining authors declare no competing financial interests.

Correspondence: Peter Fayers, Institute of Applied Health Sciences, University of Aberdeen, Aberdeen AB25 2ZD, United Kingdom; e-mail: p.fayers@abdn.ac.uk.

## References

1. Singhal S, Mehta J, Desikan R, et al. Antitumor activity of thalidomide in refractory multiple myeloma. *N Engl J Med*. 1999;341(21):1565-1571.
2. Glasmacher A, Hahn C, Hoffmann F, et al. A systematic review of phase-II trials of thalidomide monotherapy in patients with relapsed or refractory multiple myeloma. *Br J Haematol*. 2006; 132(5):584-593.
3. Palumbo A, Bringhen S, Caravita T, et al. Oral melphalan and prednisone chemotherapy plus thalidomide compared with melphalan and prednisone alone in elderly patients with multiple



- myeloma: randomised controlled trial. *Lancet*. 2006;367(9513):825-831.
4. 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.
  5. 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 randomized trial. *Lancet*. 2007;370(9594):1209-1218.
  6. 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.
  7. 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.
  8. Waage A, Gimsing P, Fayers PM, et al. Melphalan and prednisone plus thalidomide or placebo in elderly patients with multiple myeloma. *Blood*. 2010;116(9):1405-1411.
  9. 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.
  10. Kapoor P, Rajkumar SV, Dispenzieri A, et al. Melphalan and prednisone versus melphalan, prednisone and thalidomide for elderly and/or transplant ineligible patients with multiple myeloma: a meta-analysis. *Leukemia*. 2011;25:689-696.
  11. Greipp PR, San Miguel J, Durie BG, et al. International staging system for multiple myeloma. *J Clin Oncol*. 2005;23(15):3412-3420.
  12. Durie BG, Salmon SE. A clinical staging system for multiple myeloma: correlation of measured myeloma cell mass with presenting clinical features, response to treatment, and survival. *Cancer*. 1975;36(3):842-854.
  13. Oken MM, Creech RH, Tormey DC, et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am J Clin Oncol*. 1982;5(6):649-656.
  14. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958;53(282):457-478.
  15. Cox DR. Regression model and life-tables (with discussion). *J Roy Stat Soc B*. 1972;34(2):187-220.
  16. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539-1558.
  17. Tudor Smith C, Williamson PR, Marson AG. Investigating heterogeneity in an individual patient data meta-analysis of time to event outcomes. *Stat Med*. 2005;24(9):1307-1319.
  18. Katsahian S, Latouche A, Mary J-Y, Chevret S, Porcher R. Practical methodology of meta-analysis of individual patient data using a survival outcome. *Contemp Clin Trials*. 2008;29(2):220-230.
  19. Higgins JPT, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*, Version 5.1.0 [updated March 2011]. Cochrane Collaboration, 2011. Available from [www.cochrane-handbook.org](http://www.cochrane-handbook.org).
  20. Rucker G, Schwarzer G, Carpenter JR, Schumacher M. Undue reliance on I(2) in assessing heterogeneity may mislead. *BMC Med Res Methodol*. 2008;8:79.
  21. StataCorp. *Stata Statistical Software*, Release 11. College Station, TX: StataCorp LP; 2009.
  22. Gutierrez RG. Parametric frailty and shared frailty survival. *Stata J*. 2002;2:22-44.
  23. Harris R, Bradburn M, Deeks J, Harbord R, Altman D, Sterne J. Metan: fixed- and random-effects meta-analysis. *Stata J*. 2008;8:3-28.
  24. 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.
  25. Palumbo A, Dimopoulos M, Delforge M, et al. A phase 3 study to determine the efficacy and safety of lenalidomide combined with melphalan and prednisone in patients  $\geq$  65 years with newly diagnosed multiple myeloma (NDMM) [abstract]. *Haematologica* (EHA Annual Meeting Abstracts). 2010;95(suppl 2):234. Abstract 0566.
  26. Rajkumar SV, Jacobus S, Callander NS, et al. Lenalidomide plus high-dose dexamethasone versus lenalidomide plus low-dose dexamethasone as initial therapy for newly diagnosed multiple myeloma: an open-label randomised controlled trial. *Lancet Oncol*. 2010;11(1):29-37.
  27. Palumbo A, Facon T, Sonneveld P, et al. Thalidomide for treatment of multiple myeloma: 10 years later. *Blood*. 2008;111(8):3968-3977.