Quality of life during and following sequential treatment of previously untreated patients with multiple myeloma: findings of the Medical Research Council Myeloma IX randomised study

Kara-Louise Royle,[1] Walter M. Gregory,[1] David A. Cairns,[1] Sue E. Bell,[1] Gordon Cook,[2] Roger G. Owen,[3] Mark T. Drayson,[4] Faith E. Davies,[5,]* Graham H. Jackson,[6] Gareth J. Morgan[5,]* and J. Anthony Child[1]

[1]CTRU, University of Leeds, [2]LICAP, University of Leeds, [3]HMDS, St James's University Hospital, Leeds, [4]Clinical Immunology Service, University of Birmingham, Birmingham, [5]Institute of Cancer Research, Sutton and [6]Northern Cancer Centre, Newcastle University, Newcastle-upon-Tyne, UK

Summary

In the Medical Research Council (MRC) Myeloma IX trial (ISRCTN684564111) patients were randomised to sodium clodronate or zoledronic acid and induction treatment: cyclophosphamide, vincristine, doxorubicin and dexamethasone (CVAD) or cyclophosphamide, thalidomide and dexamethasone (CTD) followed by autologous stem cell transplant (ASCT) in the intensive pathway; attenuated CTD or melphalan and prednisolone (MP) in the non-intensive pathway. Subsequent randomisation allocated patients to either thalidomide or observation. The European Organisation for Research and Treatment of Cancer (EORTC) quality of life (QoL) questionnaires, QLQ-C30 and QLQ-MY24, were administered at baseline, 3, 6 and 12 months and annually thereafter, enabling the effect of sequential treatment on patient-reported health-related QoL (HR-QoL) to be investigated. The protocol specified four subscales of interest: Pain, Fatigue, Global Health Status/Quality of Life and Physical Functioning at 3, 6 and 12 months that were compared using linear models. The intensive pathway showed significant differences in favour of CTD for Fatigue at 3 months and Physical Functioning at 12 months. The non-intensive pathway and maintenance phase reported significant differences at 3 months; Pain (improved with attenuated CTD) and Global Health status/Quality of Life (improved with observation). The improved outcomes in MRC Myeloma IX were accompanied by some beneficial and few detrimental effects on HR-QoL.

Keywords: multiple myeloma, quality of life, EORTC QLQ-C30, EORTC MY-24, immunomodulatory agent.

The widespread adoption of intensive therapy and the subsequent introduction of novel agents in the treatment of multiple myeloma (MM) inevitably raises questions as to effects on health-related quality of life (HR-QoL). However, there is still relatively little published quality of life (QoL) data in systematically treated and assessed patients, notably in the first-line treatment setting.

The Medical Research Council (MRC) series of randomised controlled trials of treatment in previously untreated patients with MM up to MRC Myeloma IX provided only limited data on performance status. Myeloma IX included intensive and non-intensive pathways and enrolled patients of all ages. Patient-reported HR-QoL was a secondary endpoint in the trial protocol, with the aim of obtaining QoL data during the induction, consolidation and maintenance phases of treatment to assess the impact of the different regimens and components of treatment. Here, we present the results of the longitudinal analyses.

Methods

MRC myeloma IX

Myeloma IX recruited patients with newly-diagnosed MM aged ≥18 years from 120 centres in the United Kingdom between 2003 and 2007 and was registered with the appropriate organisation (ISRCTN68454111). All 120 centres reviewed the protocol and all recruited participants gave written informed consent in accordance with the Declaration of Helsinki, with an opt-in component for the QoL substudy. The study was an open-label randomised, phase III trial with a factorial design in which there were two

randomisations. The first included randomised allocation of patients to receive a bisphosphonate (sodium clodronate or zoledronic acid) and induction treatment, either intensive or non-intensive, determined by the treating clinicians' assessments in combination with patient performance status and preferences following informed discussion. In the intensive pathway, patients were randomised to receive cyclophosphamide, vincristine, doxorubicin and dexamethasone (CVAD) or oral cyclophosphamide, thalidomide and dexamethasone (CTD) followed by high dose melphalan plus autologous stem cell transplant (ASCT). In the non-intensive pathway, the randomisation was to attenuated CTD (CTDa) or melphalan and prednisolone (MP). A further randomisation allocated patients to either thalidomide maintenance therapy or observation only. Details of the trial, including the randomisation methods and sample size calculation, have been published elsewhere together with results of primary and secondary endpoints related to treatment effectiveness and safety (Morgan *et al*, 2010, 2012a,b,c, 2013).

Patient-reported outcome data collection

To collect patient-reported QoL information, two patient-completed European Organisation for Research and Treatment of Cancer (EORTC) questionnaires were used: the EORTC QLQ-C30 and QLQ-MY24. The QLQ-C30 is a fully validated tool used for assessing the QoL of cancer patients (Aaronson *et al*, 1993). When scored, the 30 questions produce 15 sub-scales relating to HR-QoL. The QLQ-MY24 is a specific questionnaire module for patients with MM (Stead *et al*, 1999). The QLQ-MY24 was validated throughout the original study period and consists of 24 questions, which when scored results in 5 subscales. As a result of the validation process the QLQ-MY24 has been refined to the QLQ-MY20 (removing the Social Support subscale) (Cocks *et al*, 2007). However, as it was administered, the original QLQ-MY24 is used here.

The over-arching QoL population was defined as any patient in the intention to treat (ITT) population, who returned a questionnaire at any time point. The protocol stated that consenting patients should complete questionnaires at baseline, 3, 6, 12 months and then annually up to 5 years post-randomisation, unless the patient participated in the maintenance randomisation when they should complete a pre-maintenance randomisation questionnaire and then at the same intervals as above. For the primary analysis three separate QoL populations were considered; intensive, non-intensive and maintenance. The intensive population was defined as any individual who returned a questionnaire during their induction treatment (pre-second randomisation) and was randomised to receive CTD or CVAD. Similarly, the non-intensive population was defined as any individual who returned a questionnaire during their induction treatment (pre-second randomisation) and was randomised to receive CTDa or MP. Finally, the maintenance population was defined as any individual who was randomised to receive maintenance therapy (thalidomide or observation only) and returned a questionnaire during their treatment. In all three populations, questionnaire compliance was calculated as the proportion of those in the population who returned a questionnaire out of those expected to return a questionnaire (alive on study treatment and not withdrawn). An alternative population to the non-intensive and intensive populations was defined for some of the exploratory analysis. This considered the two populations together as a first randomisation population; specifically, any individual who returned a questionnaire during their induction treatment (pre-second randomisation) was included in this population.

Data imputation and subscale scoring

Missing questionnaire data and relevant baseline factors were imputed using multiple imputation by chained equations (MICE) (White *et al*, 2011). The imputation was conducted separately for the three analysis populations described above. In addition, the imputation was conducted by randomised induction treatment in the intensive and non-intensive pathways and randomised treatment in the maintenance phase, therefore only individuals with the same randomisation allocation contributed to an individual's imputed value.

The total number of imputed datasets was determined as the maximum percentage of 'missingness' in a compulsory question at any time point, or baseline measurement [stratification factors, age, sex and International Staging System (ISS)] across the three populations. For each population, imputation was conducted in temporal order, therefore the baseline measurements were imputed first, followed by baseline questionnaire values, then the 3 months questionnaire values and so on. Within each time point, the variables were imputed in order of least to most missing. Note that only questionnaires from the treatment period being considered were used within each imputation phase: for example, only induction questionnaires from non-intensive pathway patients were used to impute data for the non-intensive pathway.

The imputed datasets were scored according to the questionnaire manuals and analysis performed in each imputed dataset. The parameter estimates (coefficient and standard errors) were combined according to Rubin (1987) and appropriate Wald tests obtained (White *et al*, 2011). All analyses was undertaken using SAS version 9.4 (SAS Institute, Cary, NC, USA) and data collected up to 8 May 2014.

Primary analysis

Analyses of patient-reported QoL were pre-specified in a separate statistical analysis plan. For the primary analysis the dependent variables were defined as the four subscales of interest (Pain, Fatigue, Global Health Status/Quality of Life and Physical Functioning) at three time points (3, 6 and 12 months) for each analysis population (intensive, non-intensive and maintenance).

Multiple linear regression was used to regress the subscale of interest at each time point on the variable of interest. For the first randomisation, this was the allocated induction treatment (intensive: CTD *versus* CVAD, non-intensive: CTDa *versus* MP) adjusting for the baseline value of the subscale, the bisphosphonate allocation (sodium clodronate or zoledronic acid), ISS (I-III), age, sex, haemoglobin at baseline (<115, $\geq$115 g/l for males; <95, $\geq$95 g/l for females), corrected serum calcium at baseline (<2·6, $\geq$2·6 mmol/l), serum creatinine at baseline (<140, $\geq$140 μmol/l) and platelet count at baseline (<150, $\geq$150 $\times$ 10^9/l). Similarly, for the maintenance randomisation, the variable of interest was the randomised treatment allocation (thalidomide *versus* observation only) adjusting for the baseline value of the subscale, the treatment group allocated by the first randomisation, ISS (I–III), age and sex. The assumptions of the linear regression models were investigated for each comparison. All statistical tests were two-sided and assessed at the 5% level of significance. As a guide for interpretation, an arbitrary difference of $\geq$10 points in a subscale was pre-defined as the minimal important difference (MID) required to suggest clinical relevance (Osoba *et al*, 1998).

Sensitivity and exploratory analysis

Sensitivity analysis assessed the impact of imputation by estimating models on complete case information. Exploratory analysis consisted of graphical displays, which summarized all questionnaire subscales with respect to the induction treatments and bisphosphonate using complete case information.

Results

Patient progress and baseline characteristics

Overall 1970 patients were recruited to Myeloma IX; 1822 (92·49%) of these were included in the QoL sub-study. The progress of patients through the sub-study is shown for the 1819 participants at first randomisation (N = 1061 intensive pathway, N = 758 non-intensive pathway) and for the 751 patients at maintenance randomisation in the CONSORT diagrams (Fig 1). In each population intensive and non-intensive pathways [Table I (key demographics and subscales of special interest), Table SI (haematological characteristics and remaining subscales)] and maintenance phase (Table SII) the distributions of age, gender, ISS and baseline QoL were similar by allocated treatment. The average ages of patients in the intensive pathway, non-intensive pathway and maintenance phase were 58, 73 and 63 years respectively, the latter figure reflecting the inclusion of patients from the two pathways.

Questionnaire return

The CONSORT diagrams show the questionnaire return rate by allocated treatment group (Fig 1). The return rate and completeness of the questionnaires were generally good with a lower return rate at baseline (20% missing baseline questionnaires in both the intensive and non-intensive pathways and 44% missing in the maintenance phase). In the intensive pathway, the CVAD group had a fairly constant rate of 'missingness', which contrasted with the CTD group where a large discrepancy was observed at 4 years (CVAD: 16% missing, CTD: 24% missing). In the non-intensive pathway, a similar proportion of questionnaires were returned in both groups at each time point with the exception of 12 months post-randomisation (MP: 13% missing, CTDa: 24% missing). In the maintenance phase, the return rate improved after the low return rate at baseline, with only approximately 10% of questionnaires missing up to 3 years post-randomisation. The highest percentage of total 'missingness' not relating to a conditional question was 63%, 71% and 48% for the intensive and non-intensive pathways and maintenance phase, respectively. Thus 71 imputed datasets were created for each analysis population.

Principal findings in the treatment pathways

In the intensive pathway there were no significant differences at any time point between treatment with CVAD and CTD with respect to Pain and Global Health Status/Quality of Life (Fig 2A). However, there were small significant differences in favour of CTD for Fatigue at 3 months with a persistent trend for benefit at 6 and 12 months (Fatigue lower with CTD, 3 months: −4·02 [95% confidence interval (CI) −7·43, −0·61], P = 0·02; 6 months: −2·06 [95% CI −6·72, 4·48], P = 0·39; 12 months: −3·63 [95% CI −8·51, 1·26], P = 0·15, Fig 2A) and Physical Functioning at 12 months after small non-significant benefits at 3 and 6 months (Physical Functioning higher with CTD, 3 months: 0·83 [95% CI −2·37, 4·04], P = 0·61; 6 months: 0·82 [95% CI −2·47, 4·10], P = 0·63; 12 months: 4·49 [95% CI 0·22, 8·76], P = 0·04, Fig 2A).

In the non-intensive pathway there was a small but highly significant difference in favour of CTDa for Pain at 3 months with a persistent trend for benefit at 6 and 12 months (Pain lower for CTDa, 3 months: −8·88 [95% CI −13·58, −4·18], P = 0·0002; 6 months: −4·82 [95% CI −9·91, 0·26], P = 0·06; 12 months: −4·57 [95% CI −11·96, 2·81], P = 0·22, Fig 2B). This difference was smaller than our pre-specified MID. There were no significant differences for Fatigue, Global Health Status/Quality of Life or Physical Functioning (Fig 2B).

In the maintenance phase there was a small significant difference in favour of observation only for Global Health Status/Quality of Life at 3 months with a persistent trend for detriment at 6 months and 12 months (Global Health Status/Quality of Life lower with thalidomide maintenance, 3 months: −3·39 [95% CI −6·33, −0·45], P = 0·02; 6 months: −1·28 [95% CI −4·86, 2·30], P = 0·48; 12 months: −2·15 [95% CI −5·98, 1·67], P = 0·27, Fig 2C). There were no significant differences for Pain, Fatigue or Physical Functioning (Fig 2C).

(A)

Intensive pathway

1114 patients randomised to receive first-line bisphosphonate treatment plus intensive induction chemotherapy

3 excluded
 1 duplicate patient
 1 no consent received
 1 misdiagnosed

1111 eligible for ITT analysis

278 assigned to receive clodronic acid plus CVAD

278 assigned to receive zoledronic acid plus CVAD

278 assigned to receive clodronic acid plus CTD

277 assigned to receive zoledronic acid plus CTD

540 started CVAD
16 did not start CVAD
 4 died prior to treatment
 8 other
 4 missing reason

548 started CTD
7 did not start CTD
 6 other
 1 missing reason

747 received 200 mg/m^2 melphalan and autologous stem-cell transplantation
 383 on zoledronic acid
 364 on clodronic acid
364 did not receive high-dose melphalan plus autologous stem-cell transplantation
 129 patient not fit/clinical decision
 93 died
 29 patients decision
 15 inadequate number of stem cells
 98 other

Quality of life

1111 ITT Population
 1054 gave QoL consent
 22 gave implicit QoL consent
 35 refused QoL consent

15 returned no questionnaires

1061 QoL population

520 assigned to CVAD

541 assigned to CTD

| | Baseline | |
520 participants
 403 questionnaires returned
 117 missing questionnaires

541 participants
 429 questionnaires returned
 112 missing questionnaires

29 died
1 withdrew

24 died

490 participants
 412 questionnaires returned
 78 missing questionnaires

3 months

517 participants
 456 questionnaires returned
 61 missing questionnaires

11 moved to maintenance therapy
13 died

19 moved to maintenance therapy
18 died

466 participants
 388 questionnaires returned
 78 missing questionnaires

6 months

480 participants
 409 questionnaires returned
 71 missing questionnaires

182 moved to maintenance therapy
22 died
1 withdrew

211 moved to maintenance therapy
21 died

261 participants
 218 questionnaires returned
 43 missing questionnaires

12 months

248 participants
 205 questionnaires returned
 43 missing questionnaires

37 moved to maintenance therapy
35 died
1 withdrew

17 moved to maintenance therapy
21 died

188 participants
 156 questionnaires returned
 32 missing questionnaires

2 years

210 participants
 164 questionnaires returned
 46 missing questionnaires

20 died

28 died

168 participants
 136 questionnaires returned
 32 missing questionnaires

3 years

182 participants
 143 questionnaires returned
 39 missing questionnaires

27 died

18 died

141 participants
 119 questionnaires returned
 22 missing questionnaires

4 years

164 participants
 125 questionnaires returned
 39 missing questionnaires

19 died

19 died

122 participants
 99 questionnaires returned
 23 missing questionnaires

5 years

145 participants
 112 questionnaires returned
 33 missing questionnaires
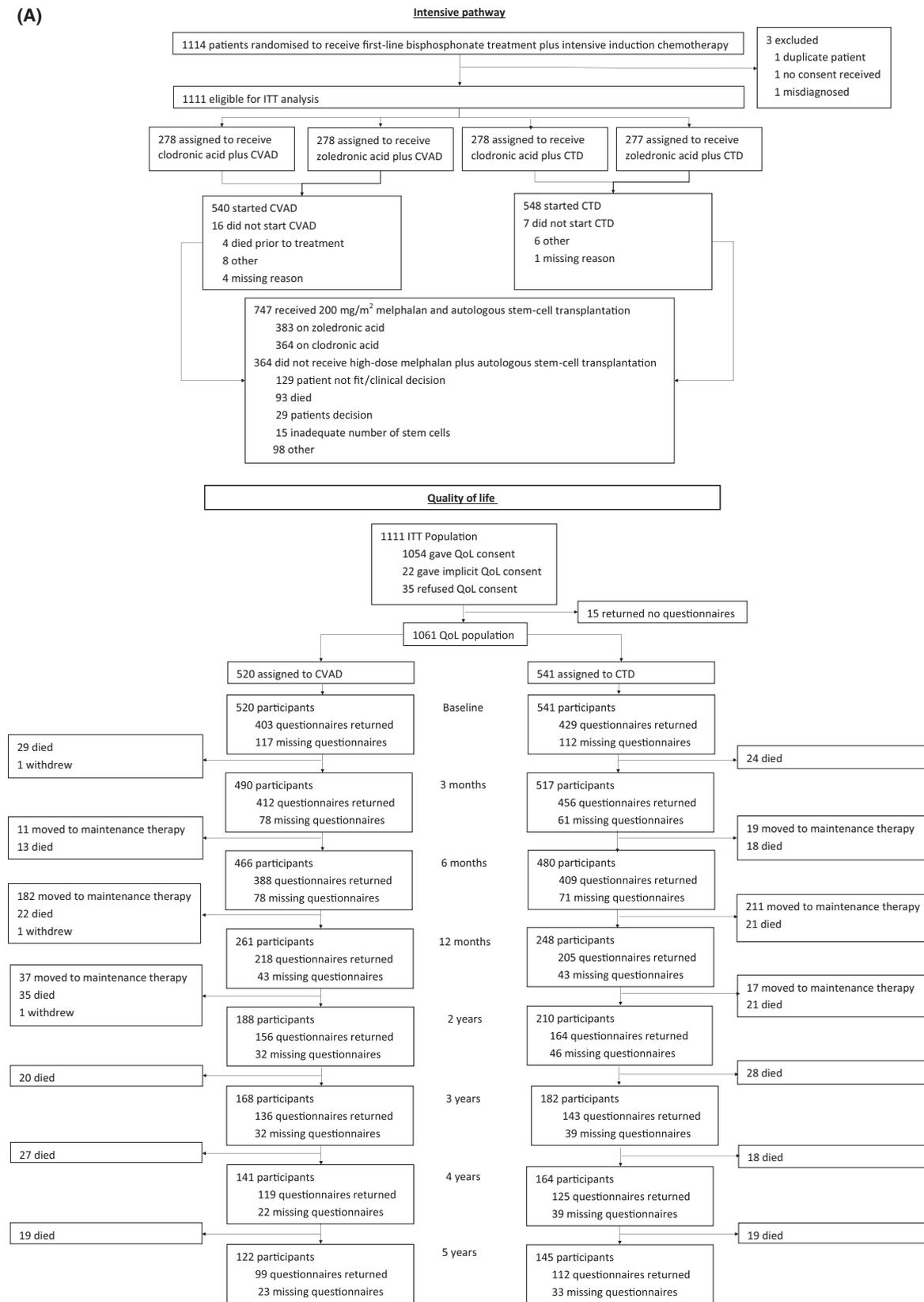
Fig 1. CONSORT diagrams. For the intensive pathway (A), non-intensive pathway (B) and the maintenance phase (C). CTD, cyclophosphamide, thalidomide and dexamethasone; CTDa, attenuated cyclophosphamide, thalidomide and dexamethasone; CVAD, cyclophosphamide, vincristine, doxorubicin and dexamethasone; ITT, intention to treat; MP, melphalan and prednisolone; QoL, quality of life.
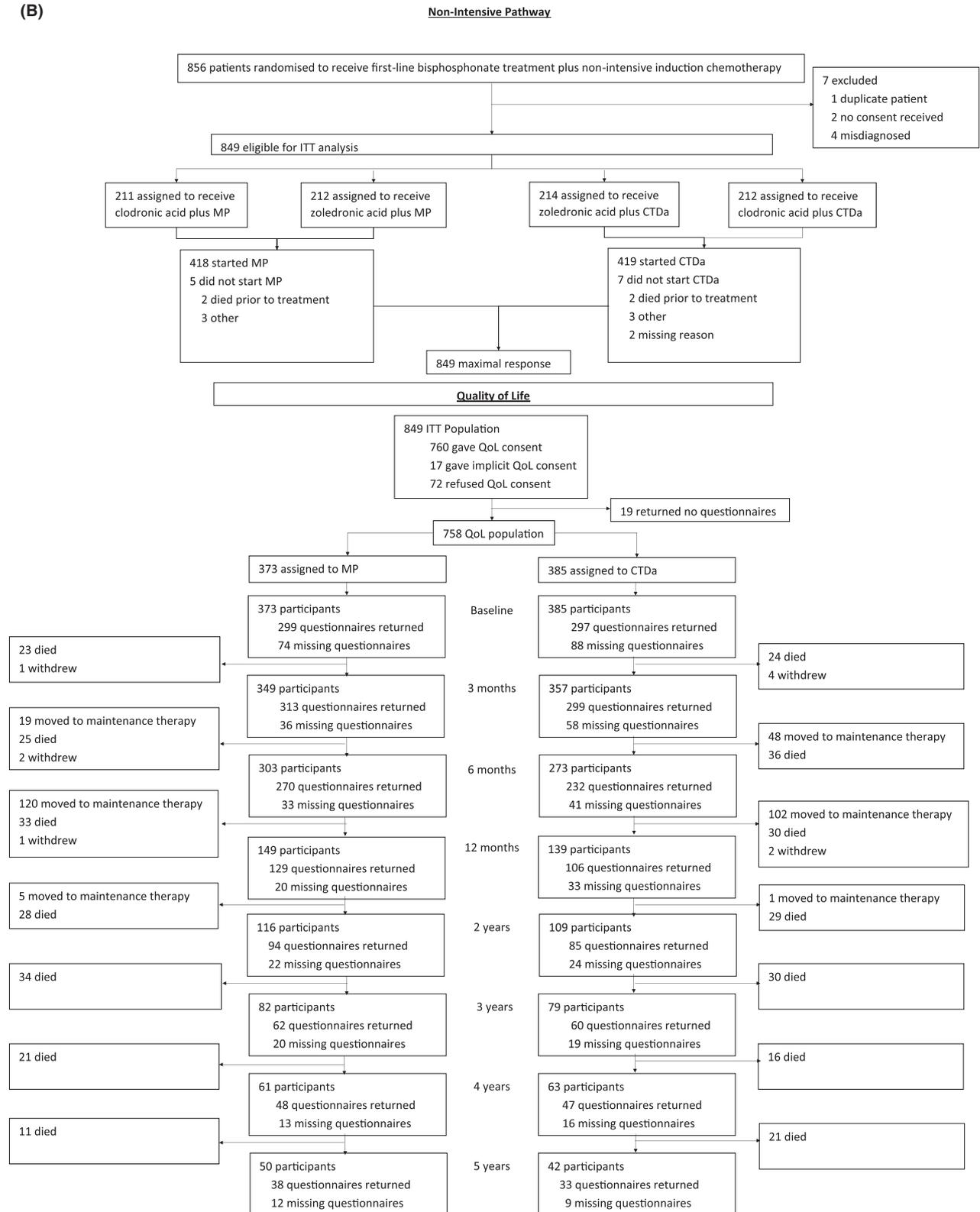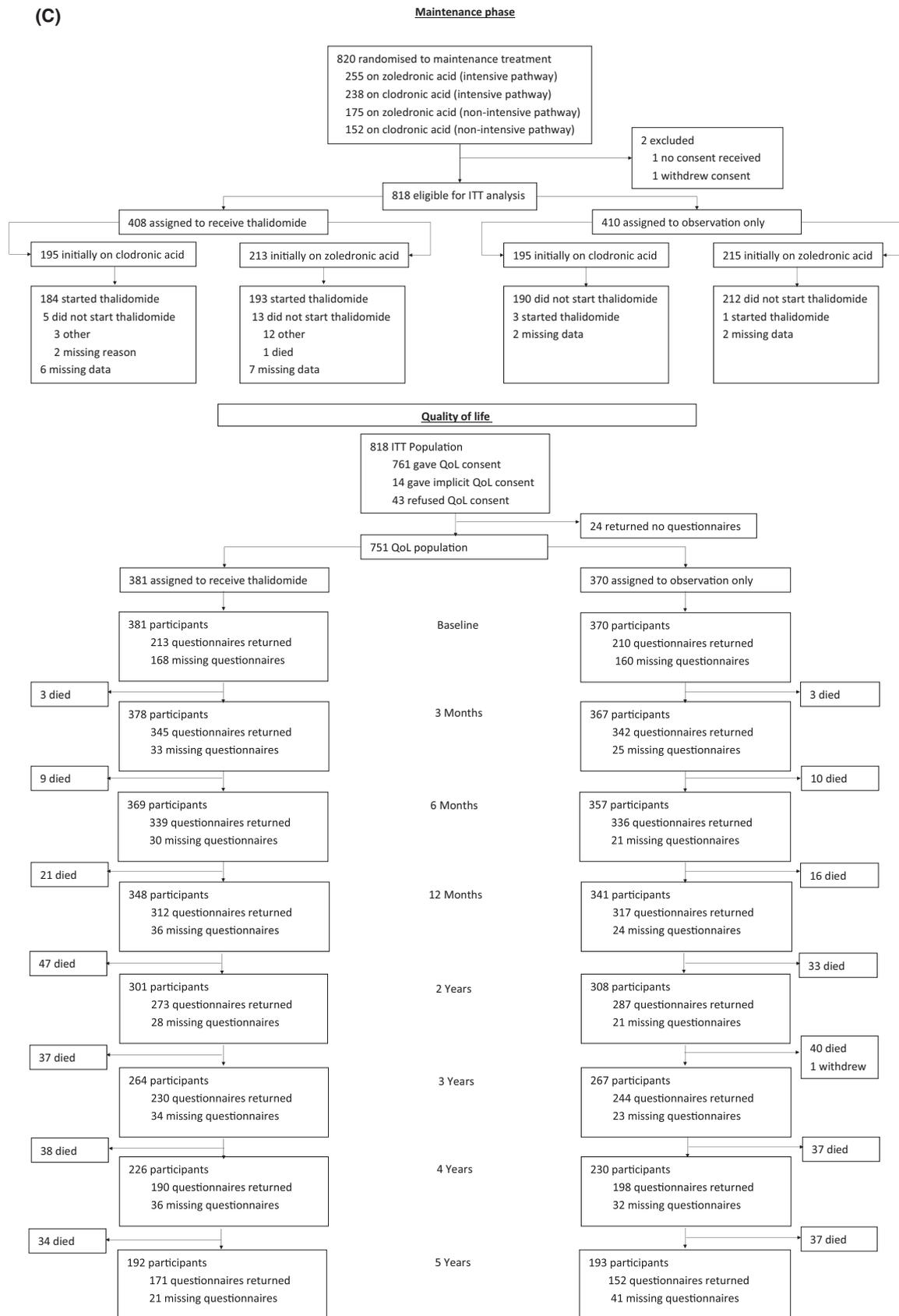
(B)

Non-Intensive Pathway

856 patients randomised to receive first-line bisphosphonate treatment plus non-intensive induction chemotherapy

7 excluded
  1 duplicate patient
  2 no consent received
  4 misdiagnosed

849 eligible for ITT analysis

211 assigned to receive clodronic acid plus MP

212 assigned to receive zoledronic acid plus MP

214 assigned to receive zoledronic acid plus CTDa

212 assigned to receive clodronic acid plus CTDa

418 started MP
5 did not start MP
  2 died prior to treatment
  3 other

419 started CTDa
7 did not start CTDa
  2 died prior to treatment
  3 other
  2 missing reason

849 maximal response

Quality of Life

849 ITT Population
  760 gave QoL consent
  17 gave implicit QoL consent
  72 refused QoL consent

19 returned no questionnaires

758 QoL population

373 assigned to MP

385 assigned to CTDa

Baseline

373 participants
  299 questionnaires returned
  74 missing questionnaires

385 participants
  297 questionnaires returned
  88 missing questionnaires

23 died
1 withdrew

24 died
4 withdrew

3 months

349 participants
  313 questionnaires returned
  36 missing questionnaires

357 participants
  299 questionnaires returned
  58 missing questionnaires

19 moved to maintenance therapy
25 died
2 withdrew

48 moved to maintenance therapy
36 died

6 months

303 participants
  270 questionnaires returned
  33 missing questionnaires

273 participants
  232 questionnaires returned
  41 missing questionnaires

120 moved to maintenance therapy
33 died
1 withdrew

102 moved to maintenance therapy
30 died
2 withdrew

12 months

149 participants
  129 questionnaires returned
  20 missing questionnaires

139 participants
  106 questionnaires returned
  33 missing questionnaires

5 moved to maintenance therapy
28 died

1 moved to maintenance therapy
29 died

2 years

116 participants
  94 questionnaires returned
  22 missing questionnaires

109 participants
  85 questionnaires returned
  24 missing questionnaires

34 died

30 died

3 years

82 participants
  62 questionnaires returned
  20 missing questionnaires

79 participants
  60 questionnaires returned
  19 missing questionnaires

21 died

16 died

4 years

61 participants
  48 questionnaires returned
  13 missing questionnaires

63 participants
  47 questionnaires returned
  16 missing questionnaires

11 died

21 died

5 years

50 participants
  38 questionnaires returned
  12 missing questionnaires

42 participants
  33 questionnaires returned
  9 missing questionnaires

Fig 1b. (Continued)

(C)

Maintenance phase

820 randomised to maintenance treatment
255 on zoledronic acid (intensive pathway)
238 on clodronic acid (intensive pathway)
175 on zoledronic acid (non-intensive pathway)
152 on clodronic acid (non-intensive pathway)

2 excluded
1 no consent received
1 withdrew consent

818 eligible for ITT analysis

408 assigned to receive thalidomide

410 assigned to observation only

195 initially on clodronic acid

213 initially on zoledronic acid

195 initially on clodronic acid

215 initially on zoledronic acid

184 started thalidomide
5 did not start thalidomide
3 other
2 missing reason
6 missing data

193 started thalidomide
13 did not start thalidomide
12 other
1 died
7 missing data

190 did not start thalidomide
3 started thalidomide
2 missing data

212 did not start thalidomide
1 started thalidomide
2 missing data

Quality of life

818 ITT Population
761 gave QoL consent
14 gave implicit QoL consent
43 refused QoL consent

24 returned no questionnaires

751 QoL population

381 assigned to receive thalidomide

370 assigned to observation only

	Baseline	
381 participants		370 participants
213 questionnaires returned		210 questionnaires returned
168 missing questionnaires		160 missing questionnaires

3 died   3 died

	3 Months	
378 participants		367 participants
345 questionnaires returned		342 questionnaires returned
33 missing questionnaires		25 missing questionnaires

9 died   10 died

	6 Months	
369 participants		357 participants
339 questionnaires returned		336 questionnaires returned
30 missing questionnaires		21 missing questionnaires

21 died   16 died

	12 Months	
348 participants		341 participants
312 questionnaires returned		317 questionnaires returned
36 missing questionnaires		24 missing questionnaires

47 died   33 died

	2 Years	
301 participants		308 participants
273 questionnaires returned		287 questionnaires returned
28 missing questionnaires		21 missing questionnaires

37 died   40 died
                           1 withdrew

	3 Years	
264 participants		267 participants
230 questionnaires returned		244 questionnaires returned
34 missing questionnaires		23 missing questionnaires

38 died   37 died

	4 Years	
226 participants		230 participants
190 questionnaires returned		198 questionnaires returned
36 missing questionnaires		32 missing questionnaires

34 died   37 died

	5 Years	
192 participants		193 participants
171 questionnaires returned		152 questionnaires returned
21 missing questionnaires		41 missing questionnaires

Fig 1c. (Continued)

Table I. Select baseline demographics of the QoL population for patients in the intensive and non-intensive pathway. The full haematological characteristics and remaining subscales can be found within Table SI.

	Intensive pathway			Non-intensive pathway		
	CVAD ($n = 520$)	CTD ($n = 541$)	Total ($n = 1061$)	MP ($n = 373$)	CTDa ($n = 385$)	Total ($n = 758$)
Age (years) at first randomisation, mean (SD)	57·7 (7·39)	57·9 (7·47)	57·8 (7·43)	73·3 (5·74)	73·1 (5·00)	73·2 (5·37)
Sex, *n* (%)						
Male	324 (62·3%)	336 (62·1%)	660 (62·2%)	204 (54·7%)	217 (56·4%)	421 (55·5%)
Female	196 (37·7%)	205 (37·9%)	401 (37·8%)	169 (45·3%)	168 (43·6%)	337 (44·5%)
Race, *n* (%)						
Caucasian	510 (98·1%)	514 (95·0%)	1024 (96·5%)	365 (97·9%)	373 (96·9%)	738 (97·4%)
Black African	2 (0·4%)	7 (1·3%)	9 (0·8%)	1 (0·3%)	2 (0·5%)	3 (0·4%)
Black Caribbean	2 (0·4%)	3 (0·6%)	5 (0·5%)	4 (1·1%)	3 (0·8%)	7 (0·9%)
Asian	4 (0·8%)	9 (1·7%)	13 (1·2%)	1 (0·3%)	1 (0·3%)	2 (0·3%)
Other	1 (0·2%)	7 (1·3%)	8 (0·8%)	1 (0·3%)	3 (0·8%)	4 (0·5%)
Missing data	1 (0·2%)	1 (0·2%)	2 (0·2%)	1 (0·3%)	3 (0·8%)	4 (0·5%)
ISS, *n* (%)						
I	122 (23·5%)	150 (27·7%)	272 (25·6%)	63 (16·9%)	43 (11·2%)	106 (14·0%)
II	177 (34·0%)	185 (34·2%)	362 (34·1%)	139 (37·3%)	146 (37·9%)	285 (37·6%)
III	166 (31·9%)	157 (29·0%)	323 (30·4%)	141 (37·8%)	145 (37·7%)	286 (37·7%)
Missing data	55 (10·6%)	49 (9·1%)	104 (9·8%)	30 (8·0%)	51 (13·2%)	81 (10·7%)
Pain						
Mean (SD)	53·3 (34·41)	53·6 (35·60)	53·4 (35·01)	52·3 (34·46)	50·0 (35·76)	51·2 (35·10)
Missing	118	112	230	74	89	163
Fatigue						
Mean (SD)	50·2 (29·36)	50·2 (29·36)	50·2 (29·36)	50·1 (27·91)	50·0 (29·88)	50·1 (28·89)
Missing	118	112	230	78	88	166
Physical Functioning						
Mean (SD)	61·5 (30·07)	59·6 (31·04)	60·5 (30·57)	56·5 (27·72)	58·1 (29·34)	57·3 (28·53)
Missing	117	112	229	76	89	165
Global Health Status /Quality of Life						
Mean (SD)	51·2 (27·58)	48·6 (27·48)	49·9 (27·54)	49·7 (24·49)	51·6 (27·10)	50·6 (25·82)
Missing	118	117	235	79	94	173

CTD, cyclophosphamide, thalidomide and dexamethasone; CTDa, attenuated cyclophosphamide, thalidomide and dexamethasone; CVAD, cyclophosphamide, vincristine, doxorubicin and dexamethasone; MP, melphalan and prednisolone; SD, standard deviation; ISS, international staging system.

There was no indication of strong violations of the modelling assumptions. Any observed violations were consistent within each imputed dataset as well as in the complete case analysis.

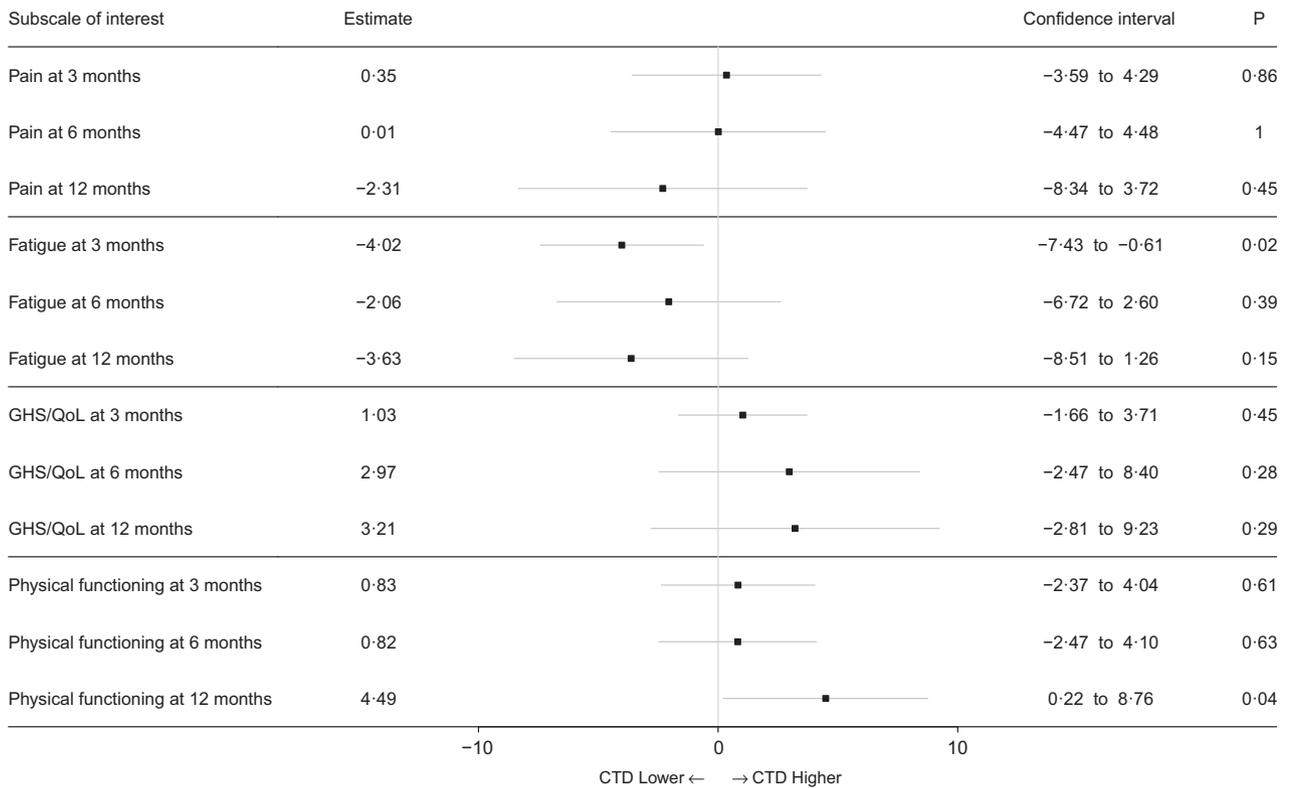*Long-term QoL in subscales of special interest*

In the longer-term, at 2, 3, 4 and 5 years post-randomisation, the pre-specified subscales of Pain, Fatigue, Global Health Status/Quality of Life and Physical Functioning show few differences between the allocated treatments in the intensive pathway, non-intensive pathway or maintenance phase. In the intensive pathway (Fig 3A) there is a relatively short-term negative effect on both Fatigue and Physical Functioning that could be attributable to high-dose therapy and ASCT. However, these negative effects do not persist beyond 1 year of follow-up and do not translate to Global Health Status/Quality of Life, which has a general increasing trend, or to any of the other subscales. There are no clear temporal trends in the subscales of special interest for either the non-intensive pathway or maintenance phase (Fig 3B and C respectively).
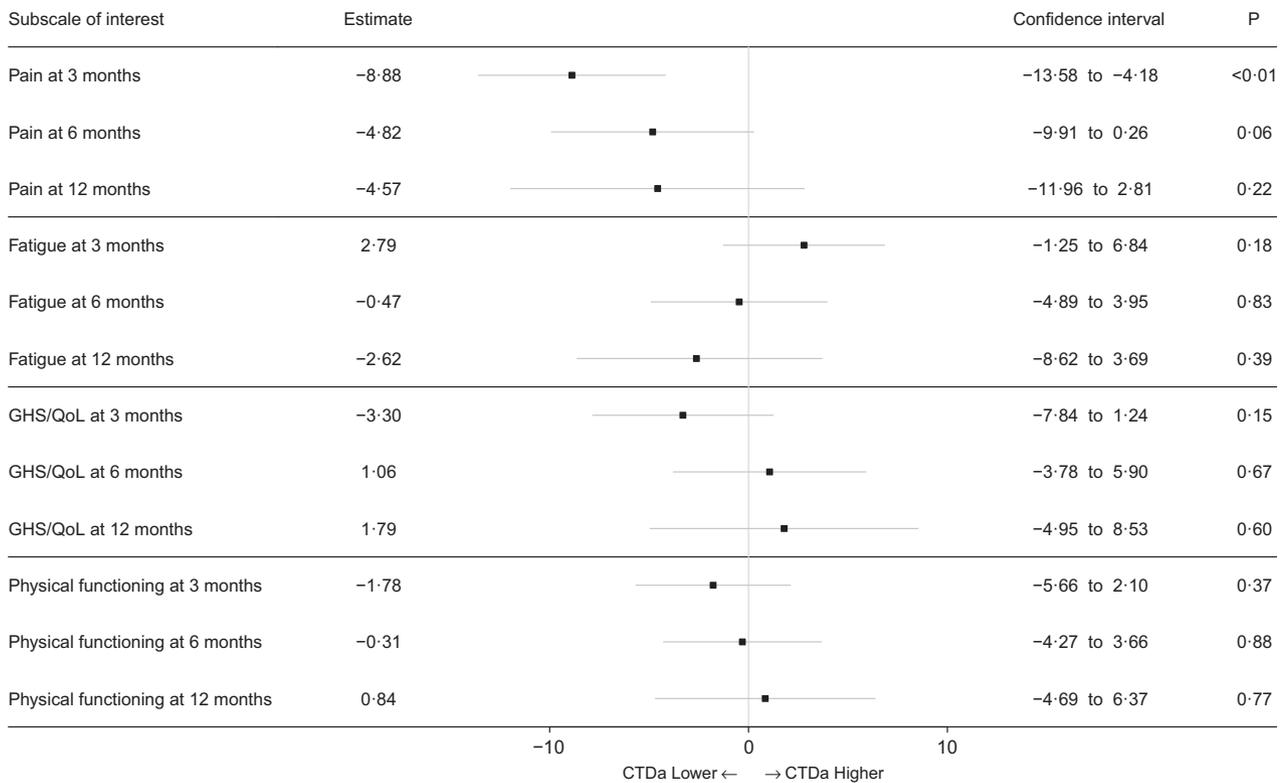
Short and long-term QoL in other subscales

Other EORTC QLQ-C30 and QLQ-MY24 subscales show largely similar trends and trajectories by allocated treatment. In the intensive pathway CVAD patients had worse Diarrhoea, Insomnia and Side Effects compared to CTD patients at 3 months. It appeared that patients in the CTD group had worse Constipation than patients on CVAD at 3 months (Figure S1A). In the non-intensive pathway, patients on CTDa had worse Constipation than patients on MP at 3 months (Figure S1B). In the maintenance phase, Constipation was worse in those on thalidomide maintenance therapy compared to observation only both at 3 and 6 months (Figure S1C).

(A)

Primary Analysis Results: Intensive Pathway
CTD vs CVAD

Subscale of interest	Estimate		Confidence interval	P
Pain at 3 months	0·35		−3·59 to 4·29	0·86
Pain at 6 months	0·01		−4·47 to 4·48	1
Pain at 12 months	−2·31		−8·34 to 3·72	0·45
Fatigue at 3 months	−4·02		−7·43 to −0·61	0·02
Fatigue at 6 months	−2·06		−6·72 to 2·60	0·39
Fatigue at 12 months	−3·63		−8·51 to 1·26	0·15
GHS/QoL at 3 months	1·03		−1·66 to 3·71	0·45
GHS/QoL at 6 months	2·97		−2·47 to 8·40	0·28
GHS/QoL at 12 months	3·21		−2·81 to 9·23	0·29
Physical functioning at 3 months	0·83		−2·37 to 4·04	0·61
Physical functioning at 6 months	0·82		−2·47 to 4·10	0·63
Physical functioning at 12 months	4·49		0·22 to 8·76	0·04

−10   0   10

CTD Lower ←   → CTD Higher

(B)

Primary Analysis Results: Non−Intensive Pathway
CTDa vs MP

Subscale of interest	Estimate		Confidence interval	P
Pain at 3 months	−8·88		−13·58 to −4·18	<0·01
Pain at 6 months	−4·82		−9·91 to 0·26	0·06
Pain at 12 months	−4·57		−11·96 to 2·81	0·22
Fatigue at 3 months	2·79		−1·25 to 6·84	0·18
Fatigue at 6 months	−0·47		−4·89 to 3·95	0·83
Fatigue at 12 months	−2·62		−8·62 to 3·69	0·39
GHS/QoL at 3 months	−3·30		−7·84 to 1·24	0·15
GHS/QoL at 6 months	1·06		−3·78 to 5·90	0·67
GHS/QoL at 12 months	1·79		−4·95 to 8·53	0·60
Physical functioning at 3 months	−1·78		−5·66 to 2·10	0·37
Physical functioning at 6 months	−0·31		−4·27 to 3·66	0·88
Physical functioning at 12 months	0·84		−4·69 to 6·37	0·77

−10   0   10

CTDa Lower ←   → CTDa Higher

Fig 2. Forest plots of the primary analysis results. For the intensive pathway (A) and non-intensive pathway (B) the models were adjusted for the baseline value of the subscale, the bisphosphonate allocation, International Staging System (ISS), age, sex and the stratification factors of the first randomisation. For the maintenance phase (C), the models were adjusted for the baseline value of the subscale, the treatment group allocated by the first randomisation, ISS, age and sex. Note that Pain and Fatigue are worse with higher scores whereas Global Health Status/Quality of Life and Physical Functioning are better with higher scores. The black squares and horizontal lines represent the treatment estimate and the associated 95% confidence interval, respectively. CTD, cyclophosphamide, thalidomide and dexamethasone; CTDa, attenuated cyclophosphamide, thalidomide and dexamethasone; CVAD, cyclophosphamide, vincristine, doxorubicin and dexamethasone; GHS/QoL, Global Health Status/Quality of Life; MP, melphalan and prednisolone.

(C)

Primary Analysis Results: Maintenance Phase
Thalidomide vs Observation

Subscale of interest	Estimate	Confidence interval	P
Pain at 3 months	−0·33	−4·58 to 3·92	0·88
Pain at 6 months	0·94	−3·72 to 5·61	0·69
Pain at 12 months	−2·24	−6·71 to 2·23	0·33
Fatigue at 3 months	2·61	−0·88 to 6·10	0·14
Fatigue at 6 months	1·40	−2·14 to 4·94	0·44
Fatigue at 12 months	−0·27	−4·08 to 3·53	0·89
GHS/QoL at 3 months	−3·39	−6·33 to −0·45	0·02
GHS/QoL at 6 months	−1·28	−4·86 to 2·30	0·48
GHS/QoL at 12 months	−2·15	−5·98 to 1·67	0·27
Physical functioning at 3 months	−0·53	−3·73 to 2·67	0·74
Physical functioning at 6 months	0·37	−2·83 to 3·67	0·82
Physical functioning at 12 months	−0·27	−3·64 to 3·10	0·88

−10   0   10

Thalidomide Lower ←   → Thalidomide Higher

Fig 2b. (Continued)

Short and long-term QoL with respect to allocated bisphosphonate treatment

There was no clear evidence of a difference in any subscale with respect to allocated bisphosphonate treatment, as all of the confidence intervals overlapped at each time point within each subscale (Figure S2).

Discussion

The Myeloma IX study evaluated the use of an immunomodulatory agent (as both induction and maintenance treatment), zoledronic acid (a newer bisphosphonate agent that required parenteral treatment) and high dose therapy with ASCT (without entry criteria explicitly connected to patient age) and showed a number of positive clinical outcomes with improved overall response, progression-free survival (PFS) and overall survival (OS). The analysis in this report has shown that these improvements have not come at the cost of clinically relevant decreases in patient-reported HR-QoL over the short-term (during initial treatment) or in the long-term (during maintenance treatment and subsequent follow-up).

For the primary analysis, the only factor to demonstrate a clear and highly statistically significant difference between treatments was the Pain result at 3 months, in favour of CTDa ($P = 0·0002$), this result being supported by a persistent trend in benefit at 6 and 12 months. Although there were statistically significant differences at the 5% level in Fatigue at 3 months and Physical Functioning at 12 months, in favour of CTD, and in Global Health Status/Quality of Life at 3 months in the maintenance phase in favour of observation only compared with thalidomide maintenance,

Intensive Pathway: CTD vs CVAD

Fig 3. Descriptive plots for the exploratory analysis showing the subscales; Pain, Fatigue, Global Health Status/Quality of Life and Physical Functioning. For the intensive pathway (A), non-intensive pathway (B) and maintenance phase (C). These graphs present the mean subscale scores and approximate 95% confidence intervals at each time point by the variable of interest for the four key subscales in complete case data. Note that Pain and Fatigue are worse with higher scores whereas Global Health Status/Quality of Life and Physical Functioning are better with higher scores. CTD, cyclophosphamide, thalidomide and dexamethasone; CTDa, attenuated cyclophosphamide, thalidomide and dexamethasone; CVAD, cyclophosphamide, vincristine, doxorubicin and dexamethasone; MP, melphalan and prednisolone.

these differences were all slight, and should be interpreted with caution given the number of tests performed.

It is interesting that pain was reduced with CTDa compared with MP, especially in the short term. It is apparent that the treatment that is more active against the myeloma is also the most effective at relieving pain, despite potentially having greater toxicity. Although the therapies used in Myeloma IX have been largely superseded by more active agents, this finding is still highly relevant to current studies, and is important to explain to the patient, given that, although they are receiving potentially more toxic drugs, the greater activity of these drugs will render them more effective, both in terms of treating the myeloma itself, but also potentially in terms of relieving pain. The same applies to the other HR-QoL measures, namely fatigue and physical functioning, which appeared to be improved with CTD/CTDa.

In terms of clinical relevance, according to the pre-defined MID, the primary analysis showed no 'clinically relevant'

differences. However, adopting alternative guidance on the interpretation of the individual elements of the EORTC QLQ-C30 questionnaire suggests that all but Pain at 3 months, which favoured CTDa in the non-intensive pathway, was unlikely to have much clinical relevance (Cocks et al, 2011). This same interpretation would have been the conclusion had we adopted alternative criteria where anything between a 6- and 17-point difference in any sub-scale is regarded as clinically relevant (Kvam et al, 2010).

However, if these results are considered temporally and in the long-term, most of the subscales are in favour of the experimental treatment. For example, Global Health Status/Quality of Life in the intensive pathway increases in favour of CTD as time passes, despite the difference not attaining statistical or clinical significance according to our pre-determined MID. This raises the question, what is more important: a large change in an individual scale, or multiple small changes across key subscales that are maintained over time?

(B)   Non−Intensive Pathway: CTDa vs MP
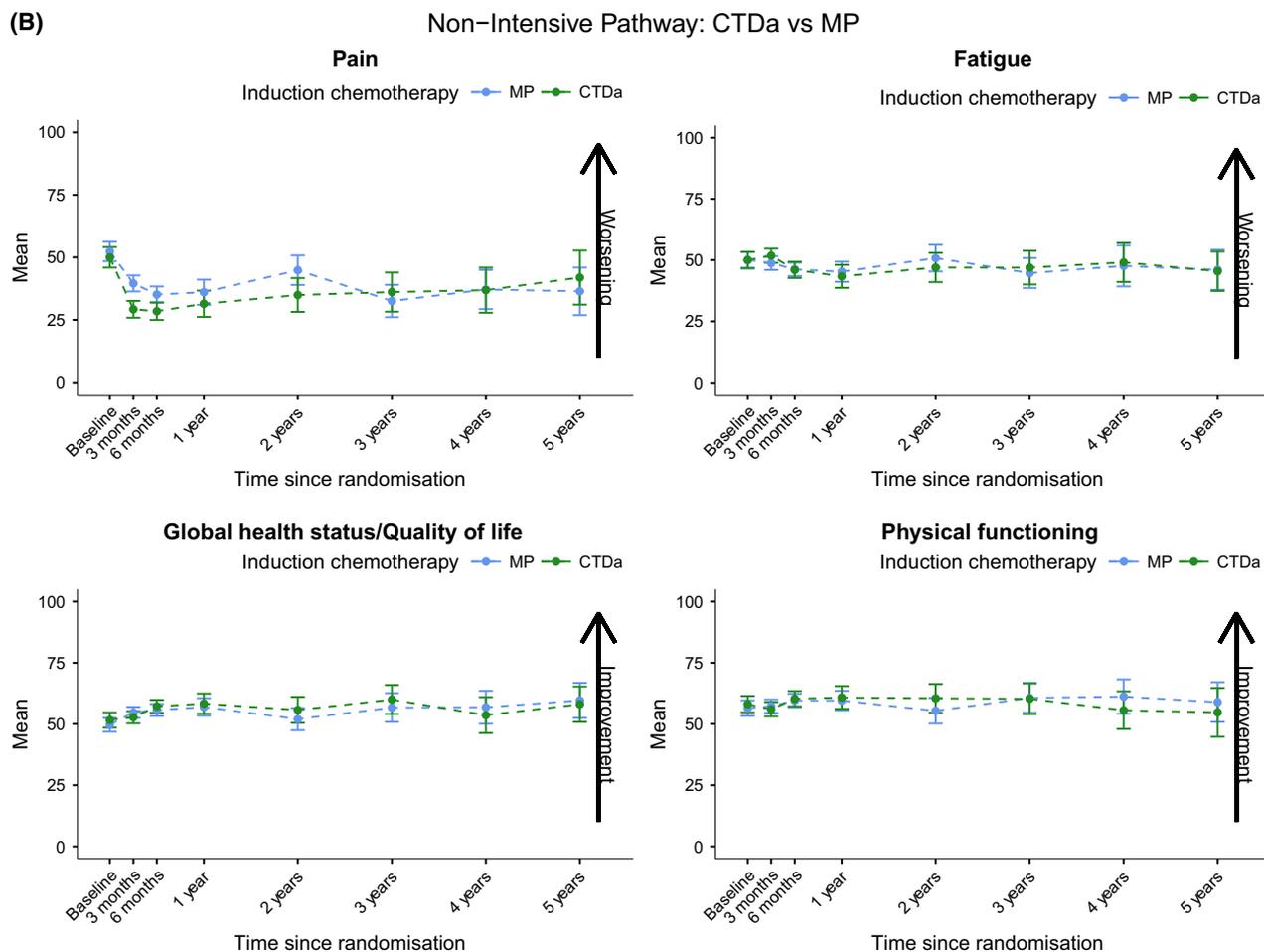
Fig 3b. (Continued)

Alternative definitions of a MID, whether it be ≥5 points (Kvam *et al*, 2010; Verelst *et al*, 2011; Delforge *et al*, 2012), subscale specific (Cocks *et al*, 2011; Dimopoulos *et al*, 2014; Delforge *et al*, 2015), or ≥10 points, as used in this work, makes comparison between studies very difficult and research less meaningful. A consensus on such measurements is clearly required.

In our study, absence of significant differences at each time point during follow up indicates that there is little perturbation of HR-QoL in the intensive pathway, when comparing the influence of CTD and CVAD, in the non-intensive pathway, when investigating the difference between CTDa and MP, or in the maintenance phase when examining the impact of thalidomide maintenance therapy compared to observation only. Also, so far as the two bisphosphonates are concerned, the informal comparisons suggest little difference in HR-QoL between patients taking clodronate and those receiving zoledronic acid.

These results are reassuring and supportive of our conclusions concerning the improved responses and survival outcomes following treatment in Myeloma IX. In the intensive

pathway, CTD improved the overall response rate compared to CVAD and CTD was shown to be non-inferior to CVAD in terms of PFS and OS (Morgan *et al*, 2012a). In the non-intensive pathway, CTDa improved response rate over MP and CTDa was shown to significantly improve PFS compared to MP (Morgan *et al*, 2011). In the maintenance phase, thalidomide was shown to improve PFS and was associated with improved OS compared to observation only although the median duration of maintenance therapy was only 9 months (Morgan *et al*, 2012c). Zoledronic acid reduced the risk of skeletal-related events (Morgan *et al*, 2012b) and significantly improved OS compared to sodium clodronate (Morgan *et al*, 2010). These gains in response and outcome were accompanied by some improvements rather than impairment of QoL, as determined by current HR-QoL measurement. However, the trial setting and the relatively short periods of maintenance thalidomide delivered, median treatment duration was 7 months (range, 0–50 months) overall and 9 (range, 0–50 months) and 6 (range, 0–46 months) months in the intensive and non-intensive pathways, respectively (Morgan *et al*, 2012c), preclude any firm conclusions

(C) Maintenance Phase: Thalidomide vs Observation Only
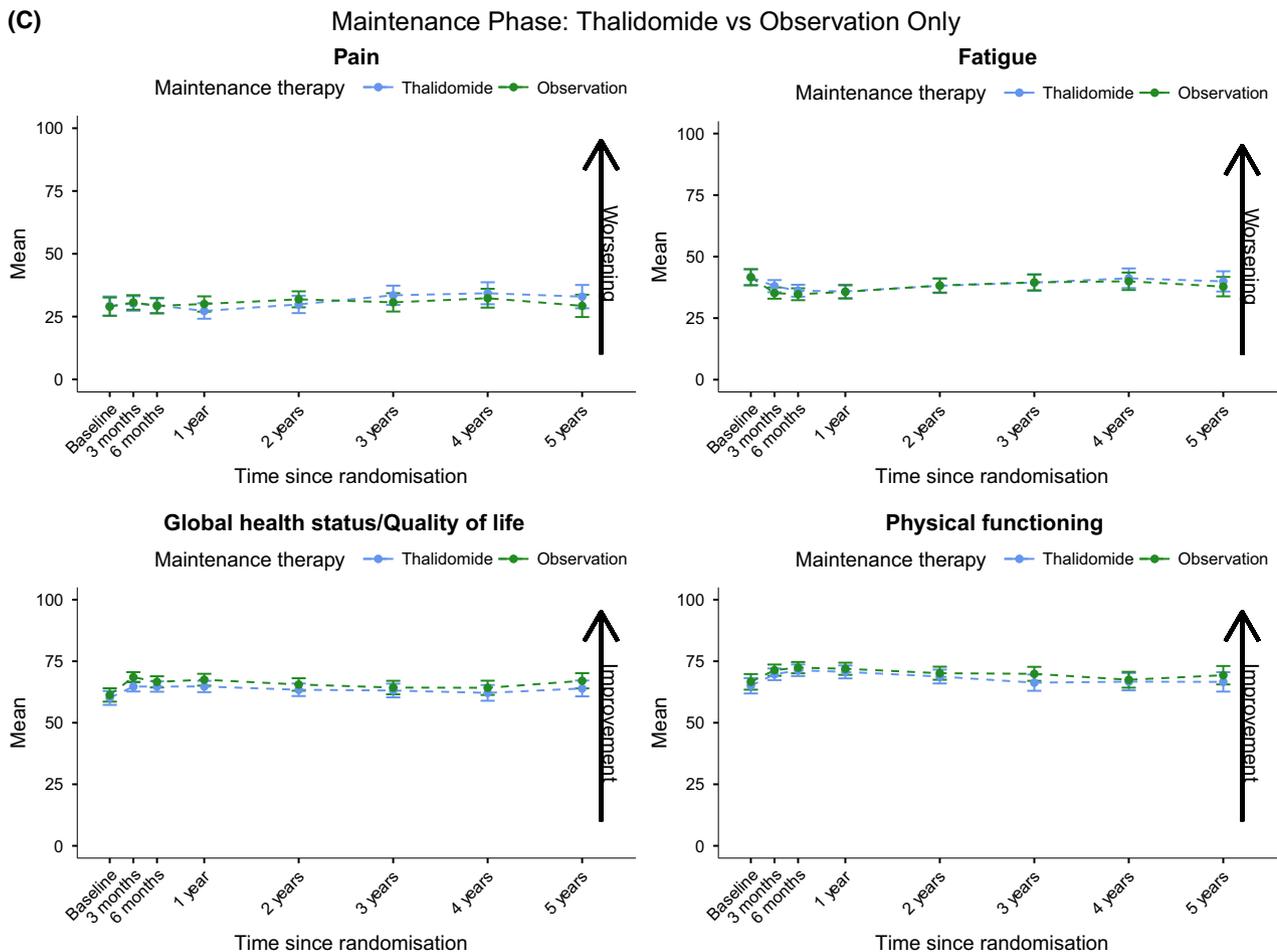
Fig 3c. (Continued)

about the longer term impact of cumulative neuropathic and other side effects on HR-QoL.

These findings are broadly similar to the few published reports of comparable approaches to treatment. The VISTA trial (Delforge *et al*, 2012) reported that QoL was comparable between the two arms (bortezomib and MP (VMP) *versus* MP) after approximately 1 year in a population similar to the non-intensive pathway. The finding that thalidomide maintenance therapy has little impact on QoL is consistent with an earlier study (Verelst *et al*, 2011) and the combination of lenalidomide and low-dose dexamethasone (RD) *versus* MPT has been reported to improve OS without having a negative effect to QoL (Delforge *et al*, 2015).

A limitation on all HR-QoL studies has been the high proportion of missing data. Although the returned questionnaires were largely complete, there was a substantial number of questionnaires missing at each time point, the baseline questionnaire having the worst return rate. To avoid the exclusion of a large proportion of patient data at later time points, MICE was implemented to impute any missing question answers or baseline measurements. MICE was chosen as it is a suitable stochastic method which is easy to implement

(https://stats.idre.ucla.edu/sas/seminars/multiple-imputation-in-sas/mi_new_1/), and has been applied in a number of recent clinical trials (Ali *et al*, 2013; Coates *et al*, 2015; Neoptolemos *et al*, 2017). MICE assumes that the missing data are missing at random (MAR). i.e., the data is missing conditional on the information that is observed. MAR is an untestable assumption; however, it was thought to be acceptable as the questionnaires were administered at definitive (randomisation-based) rather than subjective (treatment- or disease-based) time points. In addition, sensitivity analyses also showed that results obtained using MICE were consistent with those obtained in a complete case analysis. An alternative criticism could be the choice to exclude any corrections for multiplicity. As only a subset of the total number of endpoints (4 out of 20) and time points (3 out of 7) were pre-specified and modelled as of primary interest (generating 36 out of a possible 1020 tests) it was deemed not necessary. However, had we adjusted for the 36 simultaneous tests the results and conclusions would not have deviated from those discussed here given the small number of statistically significant results.

Patient-reported HR-QoL has been considered an important endpoint for patients with MM (Kvam & Waage, 2015),

giving the patient input on the effects of treatment (Delforge *et al*, 2012). The results of this study showed that improvements in clinical outcomes were not at the detriment of patient reported HR-QoL. The findings are reassuring in the context of continuing development of sequential treatment for induction, consolidation and maintenance. However, such large-scale studies, this being the largest to date, are a major undertaking and very unlikely to be given priority in future studies. Indeed, it could be that more sensitive QoL instruments, or potentially instruments focussed on specific domains, for example neurological, will be required to identify clinically relevant differences between treatment combinations. In the context of current trends in treatment with combinations of immunomodulatory agents, proteasome inhibitors, monoclonal antibodies and HDAC inhibitors, be it in modules of intensification as consolidation intended to achieve MRD negativity, sequential, multi-agent treatment as longer term "maintenance", or both, there will be a need to monitor the impact on patients' QoL. New standardised, ideally simplified, approaches that consider specific side effects of given agents as well as key QoL measures should be explored.

Acknowledgements

Authorship contribution

JAC, GJM and GHJ were the chief investigators, SEB, JAC, GJM, GHJ, FED and MTD designed the trial and developed the protocol. K-LR, WMG, DAC and JAC conducted the study, analysed the data and drafted the manuscript. All authors reviewed and approved the final manuscript.

Disclaimers

The views or opinions expressed in this manuscript are not a representation of the views or opinions of the study funders, sponsor or the authors' institutions.

Conflict of interest

K-LR: Research Funding – Amgen, Celgene, Merck. WMG: Research Funding – Amgen, Celgene, Merck. Honoraria – Janssen Corp; Consulting or Advisory Role – Celgene. DAC: Research Funding – Amgen, Celgene, Merck. SB: Research Funding – Amgen, Celgene, Merck. GC: Consulting or Advisory Role – Takeda, Celgene, Janssen Corp; Research Funding: Takeda, Janssen Corp. RGO: Honoraria – Takeda, Celgene, Janssen Corp; Consulting or Advisory Role – Celgene, Janssen Corp; Research Funding – Celgene; Travel, Accommodations, Expenses – Takeda, Janssen Corp. MTD: Stock and Other Ownership Interests – Abingdon Health; Consulting or Advisory Role – Abingdon Health. FED: Honoraria – Takeda, Janssen Corp, Celgene, Janssen Corp; Consulting or Advisory Role – Takeda, Celgene, Janssen Corp; Research Funding: Amgen, Celgene; Travel, Accommodations, Expenses: Takeda, Celgene. GJM: Honoraria – Takeda, Celgene; Consulting or Advisory Role – Celgene, Takeda, Bristol Meyers; Research Funding – Celgene, Janssen Corp. GJH: Honoraria – Takeda, Amgen, Celgene, Janssen Corp; Consulting or Advisory Role – Takeda, Amgen, Celgene, Janssen Corp; Research Funding: Amgen, Celgene; Travel, Accommodations, Expenses: Takeda, Celgene. JAC: Research Funding; Celgene.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table SI. Additional baseline demographics of the QoL population for the patients in the intensive and non-intensive pathways.

Table SII. Baseline demographics of the QoL population for patients in the maintenance phase.

Fig S1. Remaining descriptive plots for the three analysis populations. Intensive (A) non-intensive (B) and maintenance (C). These graphs present the mean subscale scores and approximate 95% confidence intervals at each time point by the variable of interest for complete case data.

References

Aaronson, N.K., Ahmedzai, S., Bergman, B., Bullinger, M., Cull, A., Duez, N.J., Filiberti, A., Flechtner, H., Fleishman, S.B., de Haes, J.C., Kaasa, S., Klee, M., Osoba, D., Razavi, D., Rofe, P.B., Schraub, S., Sneeuw, K., Sullivan, M. & Takeda, F. (1993) The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute*, **85**, 365–376.

Ali, A.M., Provenzano, E., Bartlett, J.M., Abraham, J., Driver, K., Munro, A.F., Twelves, C., Poole, C.J., Hiller, L., Dunn, J.A., Earl, H.M., Caldas, C. & Pharoah, P.D. (2013) Prognosis of early breast cancer by immunohistochemistry defined intrinsic sub-types in patients treated with adjuvant chemotherapy in the NEAT/BR9601 trial. *International Journal of Cancer*, **133**, 1470–1478.

Coates, L.C., Moverley, A.R., McParland, L., Brown, S., Navarro-Coy, N., O'Dwyer, J.L., Meads, D.M., Emery, P., Conaghan, P.G. & Helliwell, P.S. (2015) Effect of tight control of inflammation in

early psoriatic arthritis (TICOPA): a UK multicentre, open-label, randomised controlled trial. *Lancet*, **386**, 2489–2498.

Cocks, K., Cohen, D., Wisloff, F., Sezer, O., Lee, S., Hippe, E., Gimsing, P., Turesson, I., Hajek, R., Smith, A., Graham, L., Phillips, A., Stead, M., Velikova, G. & Brown, J. (2007) An international field study of the reliability and validity of a disease-specific questionnaire module (the QLQ-MY20) in assessing the quality of life of patients with multiple myeloma. *European Journal of Cancer*, **43**, 1670–1678.

Cocks, K., King, M.T., Velikova, G., Martyn St-James, M., Fayers, P.M. & Brown, J.M. (2011) Evidence-based guidelines for determination of sample size and interpretation of the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. *Journal of Clinical Oncology*, **29**, 89–96.

Delforge, M., Dhawan, R., Robinson, D. Jr, Meunier, J., Regnault, A., Esseltine, D.L., Cakana, A., van de Velde, H., Richardson, P.G. & San Miguel, J.F. (2012) Health-related quality of life in elderly, newly diagnosed multiple myeloma patients treated with VMP vs. MP: results from the VISTA trial. *European Journal of Haematology*, **89**, 16–27.

Delforge, M., Minuk, L., Eisenmann, J.C., Arnulf, B., Canepa, L., Fragasso, A., Leyvraz, S., Langer, C., Ezaydi, Y., Vogl, D.T., Giraldo-Castellano, P., Yoon, S.S., Zarnitsky, C., Escoffre-Barbe, M., Lemieux, B., Song, K., Bahlis, N.J., Guo, S., Monzini, M.S., Ervin-Haynes, A., Houck, V. & Facon, T. (2015) Health-related quality-of-life in patients with newly diagnosed multiple myeloma in the FIRST trial: lenalidomide plus low-dose dexamethasone versus melphalan, prednisone, thalidomide. *Haematologica*, **100**, 826–833.

Dimopoulos, M.A., Palumbo, A., Hajek, R., Kropff, M., Petrucci, M.T., Lewis, P., Millar, S., Zhang, J., Mei, J. & Delforge, M. (2014) Factors that influence health-related quality of life in newly diagnosed patients with multiple myeloma aged >/= 65 years treated with melphalan, prednisone and lenalidomide followed by lenalidomide maintenance: results of a randomized trial. *Leukaemia & Lymphoma*, **55**, 1489–1497.

Kvam, A.K. & Waage, A. (2015) Health-related quality of life in patients with multiple myeloma–does it matter? *Haematologica*, **100**, 704–705.

Kvam, A.K., Fayers, P. & Wisloff, F. (2010) What changes in health-related quality of life matter to

multiple myeloma patients? A prospective study. *European Journal of Haematology*, **84**, 345–353.

Morgan, G.J., Davies, F.E., Gregory, W.M., Cocks, K., Bell, S.E., Szubert, A.J., Navarro-Coy, N., Drayson, M.T., Owen, R.G., Feyler, S., Ashcroft, A.J., Ross, F., Byrne, J., Roddie, H., Rudin, C., Cook, G., Jackson, G.H. & Child, J.A.; National Cancer Research Institute Haematological Oncology Clinical Study Group. (2010) First-line treatment with zoledronic acid as compared with clodronic acid in multiple myeloma (MRC Myeloma IX): a randomised controlled trial. *Lancet*, **376**, 1989–1999.

Morgan, G.J., Davies, F.E., Gregory, W.M., Russell, N.H., Bell, S.E., Szubert, A.J., Navarro Coy, N., Cook, G., Feyler, S., Byrne, J.L., Roddie, H., Rudin, C., Drayson, M.T., Owen, R.G., Ross, F.M., Jackson, G.H. & Child, J.A. (2011) Cyclophosphamide, thalidomide, and dexamethasone (CTD) as initial therapy for patients with multiple myeloma unsuitable for autologous transplantation. *Blood*, **118**, 1231–1238.

Morgan, G.J., Davies, F.E., Gregory, W.M., Bell, S.E., Szubert, A.J., Navarro Coy, N., Cook, G., Feyler, S., Johnson, P.R., Rudin, C., Drayson, M.T., Owen, R.G., Ross, F.M., Russell, N.H., Jackson, G.H. & Child, J.A.; National Cancer Research Institute Haematological Oncology Clinical Studies Group. (2012a) Cyclophosphamide, thalidomide, and dexamethasone as induction therapy for newly diagnosed multiple myeloma patients destined for autologous stem-cell transplantation: MRC Myeloma IX randomized trial results. *Haematologica*, **97**, 442–450.

Morgan, G.J., Davies, F.E., Gregory, W.M., Szubert, A.J., Bell, S.E., Drayson, M.T., Owen, R.G., Ashcroft, A.J., Jackson, G.H. & Child, J.A.; National Cancer Research Institute Haematological Oncology Clinical Studies Group. (2012b) Effects of induction and maintenance plus long-term bisphosphonates on bone disease in patients with multiple myeloma: the Medical Research Council Myeloma IX Trial. *Blood*, **119**, 5374–5383.

Morgan, G.J., Gregory, W.M., Davies, F.E., Bell, S.E., Szubert, A.J., Brown, J.M., Coy, N.N., Cook, G., Russell, N.H., Rudin, C., Roddie, H., Drayson, M.T., Owen, R.G., Ross, F.M., Jackson, G.H. & Child, J.A.; National Cancer Research Institute Haematological Oncology Clinical Studies Group. (2012c) The role of maintenance thalidomide therapy in multiple myeloma: MRC Myeloma IX results and meta-analysis. *Blood*, **119**, 7–15.

Morgan, G.J., Davies, F.E., Gregory, W.M., Bell, S.E., Szubert, A.J., Cook, G., Drayson, M.T., Owen, R.G., Ross, F.M., Jackson, G.H. & Child, J.A. (2013) Long-term follow-up of MRC Myeloma IX trial: survival outcomes with bisphosphonate and thalidomide treatment. *Clinical Cancer Research*, **19**, 6030–6038.

Neoptolemos, J.P., Palmer, D.H., Ghaneh, P., Psarelli, E.E., Valle, J.W., Halloran, C.M., Faluyi, O., O'Reilly, D.A., Cunningham, D., Wadsley, J., Darby, S., Meyer, T., Gillmore, R., Anthoney, A., Lind, P., Glimelius, B., Falk, S., Izbicki, J.R., Middleton, G.W., Cummins, S., Ross, P.J., Wasan, H., McDonald, A., Crosby, T., Ma, Y.T., Patel, K., Sherriff, D., Soomal, R., Borg, D., Sothi, S., Hammel, P., Hackert, T., Jackson, R. & Buchler, M.W.; European Study Group for Pancreatic Cancer. (2017) Comparison of adjuvant gemcitabine and capecitabine with gemcitabine monotherapy in patients with resected pancreatic cancer (ESPAC-4): a multicentre, open-label, randomised, phase 3 trial. *Lancet*, **389**, 1011–1024.

Osoba, D., Rodrigues, G., Myles, J., Zee, B. & Pater, J. (1998) Interpreting the significance of changes in health-related quality-of-life scores. *Journal of Clinical Oncology*, **16**, 139–144.

Rubin, D.B. (1987) Multiple Imputation for Non-response in Surveys. Wiley, New York.

Stead, M.L., Brown, J.M., Velikova, G., Kaasa, S., Wisloff, F., Child, J.A., Hippe, E., Hjorth, M., Sezer, O. & Selby, P. (1999) Development of an EORTC questionnaire module to be used in health-related quality-of-life assessment for patients with multiple myeloma. *British Journal of Haematology*, **104**, 605–611.

Verelst, S.G., Termorshuizen, F., Uyl-de Groot, C.A., Schaafsma, M.R., Ammerlaan, A.H., Wittebol, S., Sinnige, H.A., Zweegman, S., van Marwijk Kooy, M., van der Griend, R., Lokhorst, H.M., Sonneveld, P. & Wijermans, P.W.; Dutch-Belgium Hemato-Oncology Cooperative Group. (2011) Effect of thalidomide with melphalan and prednisone on health-related quality of life (HRQoL) in elderly patients with newly diagnosed multiple myeloma: a prospective analysis in a randomized trial. *Annals of Hematology*, **90**, 1427–1439.

White, I.R., Royston, P. & Wood, A.M. (2011) Multiple imputation using chained equations: issues and guidance for practice. *Statistics in Medicine*, **30**, 377–399.