Comparing efficacy and survival of initial treatments for elderly patients with newly diagnosed multiple myeloma: a network meta-analysis of randomized controlled trials

Xiaoping Liu1,*, Jiaru Chen2,*, Yuncen A He3, Xianguy Meng1, Kaili Li4, Colin K He5, Shangqin Liu4

1 Center for Evidence-Based and Translational Medicine, Zhongnan Hospital of Wuhan University, 2 School of Basic Medical Science, Wuhan University, Wuhan, Hubei, People’s Republic of China; 3 School of Medicine, Case Western Reserve University, Cleveland, OH, USA; 4 Department of Hematology, Zhongnan Hospital of Wuhan University, Wuhan, Hubei, People’s Republic of China; 5 Orient Healthcare, New York, NY, USA

*These authors contributed equally to this work

Objective: The aim of this study was to evaluate the efficacy and clinical outcome of initial therapies for elderly patients with multiple myeloma (MM).

Methods: Randomized controlled trials (RCTs) were obtained through a comprehensive search. Response rate, progression-free survival (PFS) and overall survival (OS) were the interested outcome measures. Network meta-analysis (NMA) using graph theory methodology to construct an NMA model, and sensitivity analysis were performed.

Results: Nineteen RCTs containing 7,235 participants and 17 treatments were included in the NMA. As compared to the classic melphalan plus prednisone (MP) regimen, the majority of other initial regimens showed higher rates of complete response/near complete response, overall response rate (ORR) and better PFS as well as OS. These four outcomes favored the two lenalidomide plus dexamethasone regimens (continuous lenalidomide and 18 cycles of lenalidomide plus dexamethasone), especially continuous lenalidomide plus dexamethasone regimen, over the majority of other regimens including the two established standard treatments (MP plus thalidomide or bortezomib) for elderly patients with newly diagnosed MM.

Conclusion: Continuous lenalidomide plus dexamethasone ranked as the best regimen in terms of ORR and OS for the treatment of elderly patients with newly diagnosed MM.

Keywords: multiple myeloma, previously untreated, elderly patients, initial therapies, network meta-analysis

Introduction

Multiple myeloma (MM), a plasma cell neoplasm characterized by the production of monoclonal immunoglobulin, is the second most common hematological malignancy.1 With the increased life expectancy of the normal population, more than two-thirds of MM patients are first diagnosed at the age of >65 years.2 High-dose therapy (HDT) plus autologous stem cell transplantation (ASCT) and the introduction of novel agents, such as bortezomib, carfilzomib, thalidomide, lenalidomide and pomalidomide, have substantially improved response rate and prognosis for MM patients <65 years; however, the outcome of patients who are >65 years and generally considered unsuitable for HDT–ASCT remains poor. These patients usually need milder therapies.

Melphalan plus prednisone (MP) regimen has long been considered the standard care for elderly MM patients,3 and the application of the abovementioned novel agents has enriched the treatments for elderly MM patients. Nowadays, MP
plus thalidomide/lenalidomide, MP plus bortezomib and lenalidomide/thalidomide plus dexamethasone are the widely used regimens for elderly MM patients. Moreover, most of the Phase III randomized controlled trials (RCTs) compared MP regimen with MP-based regimens that include novel agents, yet few RCTs take regimens with novel agents into direct comparison. For this reason, network meta-analysis (NMA) of RCTs comparatively evaluating the response rate (complete response/near complete response [CR/nCR] and overall response rate [ORR]) and survival prognosis (progression-free survival [PFS] and overall survival [OS]) of the initial treatment for elderly MM patients is of necessity and interest. The NMA was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.

**Methods**

**Data sources**

We carried out a comprehensive literature search in existing databases such as PubMed, Embase, and the Cochrane Library and the Science Citation Index as well as relevant websites (eg, http://www.controlledtrials.com/ and https://www.ClinicalTrials.gov) aiming at identifying the potentially eligible RCTs for our analysis. Data from conference proceedings of the American Society of Hematology (2000–2015), the American Society of Clinical Oncology (2000–2015) and the European Hematology Association were also collected. Literature searches were last updated on April 20, 2016. The keywords utilized were “multiple myeloma OR plasmacytoma”, “newly diagnosed OR firstly diagnosed OR previously untreated”, “older patients OR elderly patients OR transplant ineligible OR unsuitable for transplantation” and “initial treatment OR initial therapy”. In addition, potentially relevant studies in the references of the trials were identified and other published systemic reviews and practice guidelines were examined as well. We selected potentially eligible studies through reviewing the titles and abstracts from the results of our search strategy. Furthermore, full-text articles were independently reviewed by two authors to ensure the studies meet the following inclusion criteria: 1) the studies were RCTs; 2) the participants were elderly patients with newly diagnosed MM who were unsuitable for HDT and 3) interventions were initial therapy for MM patients. Studies that did not focus on elderly patients or report the key end points (CR/nCR, ORR, PFS and OS) of our interest were excluded. According to the Jadad scale including the reporting of the randomization method, blinding scores and completeness of follow-up, the maximal score for an included study was 5 and studies were classified on the basis of quality as high (score: 3–5) versus low (score: 0–2).

**Outcome measures**

The goal of this NMA is to evaluate the effect of initial therapy for elderly MM patients. The key end points of this review are PFS (calculated from the time of randomization until the date of progression or death from any cause during treatment or until data censoring at the last date at which the patient was known to be progression free), OS (measured from the time of randomization until the date of death from any cause or until data censoring at the last date at which the patient was known to be alive), CR/nCR and ORR (evaluated with reference to the international uniform response criteria).

**Data extraction**

Relevant studies were investigated through full-text review, and only those that met all the inclusion criteria were included in the final analysis. A predesigned data extraction form involving baseline characteristics, outcomes and number of events was utilized in the data extraction process, which was conducted independently by Xiaoping Liu and Jiarui Chen. Any discrepancies between the two investigators at the screening or data extraction stage were resolved by discussion.

**Statistical analysis**

We utilized the R Project software, version 3.2.2 (www.r-project.org) and Stata 13.0 (StataCorp, College Station, TX, USA) to perform the statistical analysis. For the end points of CR/nCR and ORR, the natural log of relative risk (RR) and its standard error were applied to evaluate the effect of the initial treatments in elderly MM patients. Meanwhile, the natural log of hazard ratio (HR) and its standard error were applied to evaluate the survival data of these patients. In case the unavailability of any HR and its 95% confidence interval (CI) was detected in any studies, we calculated the HRs and 95% CIs with the method introduced by Tierney et al. We adopted I² statistic to quantify heterogeneity among the studies. When low heterogeneity was detected (the value of I² was <50%), a fixed-effects model was preferred for the meta-analysis. Conversely, if the value of I² was detected to be >50%, we applied the random-effects model instead. Meanwhile, a loop-specific approach was applied to evaluate the consistency assumption in each closed loop of the network. In each loop, we estimated the inconsistency factor...
Comparing treatments for elderly patients with MM

(IF) as the absolute difference between direct and indirect estimates for one of the comparisons in the loop. The IF is the log of two odds ratios (RoR) from direct and indirect evidence in the loop. RoR values close to 1 indicate that the two sources are in agreement.\textsuperscript{10,11} The graph-theoretical method that exploits the analogy between treatment networks and electrical networks was used to construct an NMA model accounting for the correlated treatment effects in multiarm studies.\textsuperscript{12} This has been found to be equivalent to the frequentist approach to NMA and can be implemented via the R package “netmeta” (version 0.8-0).\textsuperscript{13} In addition, using the $P$-value recommended by Rücker and Schwarzer,\textsuperscript{14} we estimated the probability that each treatment was the best regimen, the second best and so on.

Results

Literature search results

A total of 13,155 articles were obtained through our literature search, of which 10,932 were excluded since they were not clinical trials. After reviewing the titles and abstracts, 28 potentially relevant RCTs were considered well worth full-text reading. Finally, the following nine articles were excluded: four articles were updated analysis of three other trials, two articles did not focus on elderly populations, one article was a duplicate report, one article did not address the initial therapy for elderly patients with newly diagnosed MM and one article alternated the schemes of induction therapy in different ways. Hence, a total number of 19 studies were eventually included in this NMA (Figure 1).

Description of included trials

A total number of 7,235 participants were randomly assigned in these included trials with 36–547 participants per arm. The trials assessed a total of 17 interventions or control conditions, which were grouped into 23 pairwise comparisons (Figure 2). Most trials included two arms (n=14), but five trials included three arms. The most common pairwise comparison was MP plus thalidomide (MPT) versus MP (n=6). It is worth noting that patients who completed the planned MPT cycles received maintenance therapy with oral thalidomide until disease progression in four trials (MPT-T), whereas in other two trials, patients ended their planned MPT without any maintenance therapy administrated. Other regimens used in these included trials were as follows: cyclophosphamide plus thalidomide plus dexamethasone (CTD) regimen in one trial, cyclophosphamide plus prednisone plus lenalidomide (CPR)
in one trial, lenalidomide plus dexamethasone in 28-day cycles until disease progression (Ld continuous) and the same combination of lenalidomide plus dexamethasone for 18 cycles (Ld 18 cycles) in one trial, lenalidomide plus dexamethasone for 9 cycles (Ld 9 cycles) in one trial, reduced-intensity stem cell transplantation using melphalan 100 mg/m²; MP, melphalan plus prednisone; MPR, MP plus lenalidomide; MPR-R, MPR regimen followed by lenalidomide maintenance therapy until a relapse or disease progression occurred; MPT, MP plus thalidomide; MPT-T, MPT cycles received maintenance therapy with oral thalidomide until disease progression in four trials; RCTs, randomized controlled trials; TD, thalidomide plus dexamethasone; VD, bortezomib plus dexamethasone; VMP, bortezomib plus melphalan plus prednisone; VMPS, VMP plus siltuximab; VMPT-VT, VMP plus thalidomide that followed by maintenance with bortezomib and thalidomide; VTD, bortezomib plus TD.

Response to treatment
All 19 included RCTs reported the data of CR/nCR and ORR for the NMA. For the CR/nCR, no statistically significant evidence of inconsistency was found (Figure S1A) and, according to the value of associated \( I^2 = 18.6\% \), we selected the fixed-effects model to perform the analysis. As compared to MP, all the other regimens showed higher CR/nCR rates. Ld 18 cycles as well as Ld continuous were more efficacious than CPR, MEL100, MP, MPR, MPR-R, MPT, MPT-T and TD. Meanwhile, VMP, VMPS, VMPT and VTD had significantly higher rates of CR/nCR than those resulted by CPR, MP, MPR, MPR-R, MPT-T and TD. VMPS was associated with higher CR/nCR rates compared to VD and VMP. VTD was more efficacious compared to VD. No statistically significant difference could be found among VMPS, VMPT-VT, VTD, Ld 18 cycles and Ld continuous, and no significant difference could be found among MPR, MPR-R, MPT and TD (Figure 3A and Table S4).

VMPS, Ld 18 cycles and Ld continuous ranked as the best three regimens for this outcome, and VMPS was ranked the best regimen for this outcome with relatively higher probability (\( P\)-value = 0.8711, Table S5). With respective to the end point ORR, the fixed-effects model was preferred according to the value of \( I^2 = 0\), and there was no evident inconsistency found in the network for this outcome (Figure S1B). As with CR/nCR, all the regimens resulted in higher rates of ORR compared to MP. Ld 18 cycles and Ld continuous were more efficacious than CPR, Ld 9 cycles, MEL100, MPR, MPR-R, MPT, MPT-T and TD. VD, VMP, VMPS, VMPT and VTD had significantly higher rates of ORR than those resulted by CPR. Ld 9 cycles, MP, MPR, MPR-R, MPT-T and TD. VMPT-VT was associated with higher rates of ORR than VMP. As with CR/nCR, no evident difference was found in the other comparisons (Figure 3B and Table S6). Ld continuous, Ld 18 cycles, VMPT-VT and VMPS ranked as the best four regimens for this outcome, and Ld continuous was ranked the best regimen for this outcome with relatively higher probability (\( P\)-value = 0.9510, Table S5).

PFS
A total number of 18 RCTs reported PFS-associated data for this NMA. Owing to the absence of PFS-associated data in the trial conducted by San Miguel et al,17 the whole network framework was split into two separated networks. According to the respective \( I^2 \) values (\( I^2 = 73.6\% \) and 73.5%), the random-effects model was preferred for both subnetworks and no evident inconsistency was found (Figure S1C). Ld continuous, MPR-R, MPT and MPT-T resulted in better
Figure 3 Differences in CR/nCR, ORR, PFS and OS between ld continuous or VD regimen versus other initial regimens.

Notes: (A) Network analysis of CR/nCR, (B) network analysis of ORR, (C) subnetwork analysis of PFS, (D) network analysis of OS and (E) subnetwork analysis of PFS.

Abbreviations: CI, confidence interval; CPR, cyclophosphamide plus prednisone and lenalidomide; CR/nCR, complete response/near complete response; CTD, cyclophosphamide plus thalidomide and dexamethasone; HR, hazard ratio; LD 9 cycles, lenalidomide plus dexamethasone for 9 cycles; LD 18 cycles, lenalidomide plus dexamethasone for 18 cycles; LD continuous, lenalidomide plus dexamethasone in 28-day cycles until disease progression; MEL100, reduced-intensity stem cell transplantation using melphalan 100 mg/m^2; MP, melphalan plus prednisone; MPR, MP plus lenalidomide; MPR-R, MPR regimen followed by lenalidomide maintenance therapy until a relapse or disease progression occurred; MPR-T, MP plus thalidomide; MPT, MP plus thalidomide; MPT-T, MPT cycles received maintenance therapy with oral thalidomide until disease progression in four trials; ORR, overall response rate; OS, overall survival; PFS, progression-free survival; RR, relative risk; TD, thalidomide plus dexamethasone; VD, bortezomib plus dexamethasone; VD, bortezomib plus dexamethasone; VMPS, VMP plus siltuximab; VMPT-VT, VMP plus thalidomide that followed by maintenance with bortezomib and thalidomide; VTD, bortezomib plus thalidomide. PFS than MP. The PFS favored MPR-R and MPT over MP, TD and CTD. LD continuous resulted in better PFS than CPR, CTD, LD 9 cycles, MP, MPR and TD. There was no evident difference in the other comparisons in this subnetwork (Figure 3C and Table S7a). Meanwhile, in the other subnetworks, no significant difference could be found among regimens VD, VMP, VMPS, VMPT-VT and VTD in terms of PFS (Figure 3E and Table S7b). LD continuous...
(P-value = 0.9506) and VMPT-VT (P-value = 0.8759) ranked as the best for the two subnetworks (Table S5).

**OS**

Except for one trial conducted by San Miguel et al,15 OS was reported in 18 RCTs for this network. According to the value $F_I$ ($I=40%$), the fixed-effects model was preferred and there was no evident inconsistency found in the network (Figure S1D). Ld 18 cycles, Ld continuous, MP, MPR-R, MPT, MPT-T, VMP, VMPT-VT and VTD were associated with significant superiority when compared with TD in terms of OS. Ld 18 cycles resulted in better OS than CPR, CTD, Ld 9 cycles, MEL100, MP, MPR, MPR-R, MPT-T and TD. Ld continuous resulted in better OS than CPR, CTD, MEL100, MP, MPR, MPR-R, MPT, MPT-T, TD, VD and VTD. VMP was associated with better OS than TD, MPT-T, MPR, MP, MEL100 and CTD. There was no significant difference among VD, VMP, VMPT-VT and VTD for the outcome of OS (Figure 3D and Table S8). Ld 18 cycles and Ld continuous ranked as the best two regimens for OS, and Ld continuous ranked as the best regimen for this outcome with relatively higher probability (P-value = 0.9384, Table S5).

**Sensitivity analysis**

We performed sensitivity analysis of studies reporting on >100 patients in each arm and studies containing nontransplant therapy; thus, four trials were excluded for sensitivity analysis.15,21,25,27 The result indicated that Ld 18 cycles and Ld continuous had higher rates of CR/nCR than MP and MPT. VMPT-VT and VTD had higher rates of CR/nCR than CPR, PD, MPR, MPR-R, MPT-T and TD. VMPT-VT had higher rates of ORR than VMP. Ld 18 cycles and Ld continuous were more efficacious than CPR, Ld 9 cycles, MP, MPR, MPR-R, MPT-T and TD. VMPT-VT had higher rates of ORR than VMP. As compared to MP, all the other regimens demonstrated higher rates of ORR (Table S9). For the outcome of OS, OS favored Ld continuous over TD, MP and CTD. Ld continuous, MPR-R and MPT-T had better OS compared to MP (Table S11). Meanwhile, Ld continuous, Ld 18 cycles, VMP and MPT have better OS than MP. Furthermore, Ld continuous demonstrated better OS than CPR, CTD, Ld 9 cycles, MP, MPR, MPR-R, MPT, MPT-T and TD. OS favor VMP over CTD, MP, MPR, MPR-R, MPT-T and TD (Table S12). Ld continuous ranked as the best regimen for the outcome of ORR and OS (P-values 0.9189 and 0.9237, respectively).

**Discussion**

Treatments for elderly patients with newly diagnosed MM have changed dramatically with the introduction of novel agents including proteasome inhibitors (bortezomib) and immunomodulatory drugs (thalidomide and lenalidomide), and combined chemotherapies containing these novel agents varied.3 Although several trials and standard pairwise meta-analysis had been conducted to compare these combined chemotherapies for elderly MM patients,15,17,20,34 the most efficacious and tolerable regimens remain unclear. NMA, which can take both direct evidence within clinical trials and indirect evidence across clinical trials into a whole comparison and facilitate indirect comparisons of multiple inventions that have not been compared in head-to-head clinical trials, has been applied to characterize the optimal choice in many diseases including MM.35–38

Several clinical trials and meta-analysis suggested that MPT regimen had better outcome than traditional MP, which was confirmed in our NMA, and the majority of regimens studied in our NMA showed superiority compared to MP. For years, MPT and VMP were established as standard therapies for elderly MM patients unsuitable for HDTs. However, the introduction of lenalidomide plus dexamethasone might change this pattern.39 Benbouker et al39 demonstrated that response rates were higher in Ld continuous (75%) and Ld 18 cycles (73%) groups when compared with MPT group (62%), and Ld continuous resulted in a significant improvement in PFS when compared with MPT (HR = 0.72, P < 0.001) and Ld 18 cycles (HR = 0.7, P < 0.001). Similarly, our indirect comparisons suggested that continuous lenalidomide plus dexamethasone showed superiority compared to MPT (RRCR/nCR = 1.55, 95% CI: 1.31–1.82; RRORR = 1.21, 95% CI: 1.11–1.31; HR OS = 0.78, 95% CI: 0.64–0.96) and VMP (RRORR = 0.79, 95% CI: 0.61–1.01; HR OS = 1.41, 95% CI: 0.95–2.09), and the similar trend could be found when comparing Ld 18 cycles with MPT (RRCR/nCR = 0.66, 95% CI: 0.56–0.78; RRORR = 0.85, 95% CI: 0.58–0.72; HR OS = 1.14, 95% CI: 0.94–1.39) and VMP (RRCR/nCR = 1.39, 95% CI: 0.74–2.58; RRORR = 1.27, 95% CI: 0.99–1.63; HR OS = 0.71, 95% CI: 0.48–1.05). For the outcomes CR/nCR, ORR and OS, no significant difference could be found between Ld 18 cycles and 18 continuous, whereas Ld continuous had better PFS than Ld 18 cycles (HR = 0.70, 95% CI: 0.48–1.08), which was confirmed by a previous clinical trial.20 According to the P-values, bortezomib-based regimens (VD, VTD, VMP, VMPS and VMPT-VT) were also associated with promising antitumor activity, while these regimens had similar antitumor activities. Previous studies suggested...
that melphalan might be the best partner for bortezomib-based induction regimen in elderly patients with untreated MM when the efficacy, toxicities, and costs were taken into consideration.29,40

The NMA has limitations. 1) Owing to the definitions of adverse events in our included trials that differed significantly across the initial regimens, we could not perform a comprehensive analysis of adverse events, which prevented a balanced evaluation of both efficacy and tolerability for the initial treatments of elderly MM patients. Actually, in the FIRST trial20, the author demonstrated that both Ld continuous and Ld 18 cycles groups had lower rates of hematologic toxic events than the MPT group, peripheral sensory neuropathy was more common with MPT than with Ld continuous and Ld 18 cycles groups and the incidence of second primary cancers was low across treatment groups, but the incidence of grade 3 or 4 infection was increased with Ld continuous as compared with Ld 18 cycles or MPT (29% vs 22% and 17%) and Ld continuous was associated with an increase of 5 percentage points in the infection of any grade and an increase of 2 percentage points in thromboembolic events of grade 3 or 4. Glucocorticoids might contribute partly to the long-term adverse events with Ld continuous, and lower doses of dexamethasone or the use of prednisone in the future studies might be of interest.20 2) Many treatment pairs in the network were connected by a single study, and therefore, the results may be sensitive to changes in a single trial result. Therefore, the results should be interpreted with cautions and further studies are required to strengthen the NMA.

Conclusion
Our NMA demonstrated that the two lenalidomide plus dexamethasone initial treatments (18 cycles of lenalidomide plus dexamethasone and continuous lenalidomide plus dexamethasone), especially the continuous lenalidomide plus dexamethasone, resulted in better efficacy and prognosis for the elderly patients with MM.

Acknowledgments
This study was supported by the Research Fund from the National Natural Science Foundation of China (no. 81272627). The abstract of this article was presented at the 2016 ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop as a roundtable discussion.

Disclosure
The authors report no conflicts of interest in this work.

References


