Minimal Residual Disease Status as a Surrogate Endpoint for Progression-Free Survival in Newly Diagnosed Multiple Myeloma Studies: A Meta-analysis

Hervé Avet-Loiseau[1], **Heinz Ludwig**[2], **Ola Landgren**[3], **Bruno Paiva**[4], **Chris Morris**[5], **Hui Yang**[5], **Kefei Zhou**[5], **Sunhee Ro**[5], **Maria-Victoria Mateos**[6]

[1]University of Toulouse, Toulouse, France

[2]Wilhelminen Cancer Research Institute, Wilhelminenspital, Vienna, Austria

[3]Myeloma Service, Memorial Sloan Kettering Cancer Center, New York, NY, USA

[4]Clinica Universidad de Navarra, Centro de Investigacion Medica Aplicada (CIMA), Instituto de Investigacion Sanitaria de Navarra (IDISNA), CIBER-ONC number CB16/12/00369, Pamplona, Spain

[5]Amgen, Inc., Thousand Oaks, CA, USA

[6]University Hospital of Salamanca/IBSAL, Salamanca, Spain

Abstract

Background—Therapeutic advances have greatly extended survival times in patients with multiple myeloma, necessitating increasingly lengthy trials when using survival outcomes primary endpoints. A surrogate endpoint that can more rapidly predict survival could accelerate drug development. We conducted a meta-analysis to evaluate minimal residual disease (MRD) status as a valid PFS surrogate in newly diagnosed multiple myeloma (NDMM).

Materials and Methods—We searched abstracts in PubMed, The American Society of Hematology (ASH) and the European Hematology Association (EHA) for "myeloma", "minimal residual disease", and "clinical trial". Because of the need to evaluate the treatment effect on MRD response, only randomized studies for subjects with NDMM were included. Details on the MRD-tested populations were required. Meta-analysis was performed by principles outlined at the 2013 FDA workshop on MRD in AML.[1] For samples that were not measured for MRD and within the subset specified for MRD assessment, their MRD status was imputed from the samples that had known MRD status. Patients that were excluded from planned MRD assessment were considered MRD positive.

Address for correspondence: Hervé Avet-Loiseau, IUC-Oncopole, INSERM U1037, IUC, 31059 Toulouse, France, avet-loiseau.h@chu-toulouse.fr.

Results—Six randomized studies, representing 3283 patients and 2208 MRD samples, met analysis inclusion criteria. MRD-negativity rates ranged from 0.06 to 0.70. The treatment effect on the odds ratio (OR) for MRD-negative response strongly correlated with the hazard ratio (HR) for PFS with a coefficient of determination for the weighted regression line of 0.97. Our meta-analysis suggested that MRD status met both Prentice's criteria for PFS surrogacy.

Conclusions—These results support the claim that MRD status can be used as a surrogate for PFS in NDMM.

MicroAbstract

A surrogate endpoint that can support accelerated approval of a drug in NDMM is clearly needed. Here, we performed a meta-analysis of six randomized trials to evaluate the potential of MRD status in predicting clinical benefit. A strong correlation between the relative changes of experimental to control treatments in MRD-negative rates and PFS supported MRD status as a surrogate endpoint.

Keywords

Introduction

In recent years, new therapeutic options, such as proteasome inhibitors (bortezomib, carfilzomib, and ixazomib),[2–4] immunomodulators (lenalidomide and pomalidomide), [2, 3, 5, 6] and antibody therapies (daratumumab and elotuzumab),[7, 8] have greatly improved the prognosis for patients with multiple myeloma. Overall response rates now approach 100% in newly diagnosed multiple myeloma (NDMM), with up to 80% of patients achieving very good partial response (VGPR) or better.[9] A median progression-free survival (PFS) of approximately 3–4 years is frequently reported in clinical studies, and, today, in the routine clinical setting, it is not uncommon to see patients who have lived with multiple myeloma for 15–20 years.[4, 10, 11] These advances, however, could potentially be self-limiting, since ongoing and future trials in NDMM which use the traditional endpoint of PFS may require at least 5 years to collect mature data in young transplant-eligible patients, [12, 13] thereby discouraging and delaying the development of better therapies.

Surrogate endpoints have been explored as a potential alternative to lengthy studies in other oncology indications, such as follicular lymphoma, where recent improvements in survival have presented a similar conundrum.[14, 15] In multiple myeloma, depth of response has been shown to correlate with improved outcomes.[3, 8, 16, 17] Phase 3 studies[3, 18–21] and a meta-analysis of 21 studies[22] have identified a statistically significant association of complete response (CR) rate with PFS. Most patients initially achieve CR, however, eventually relapse, reflecting disease persistence via clonal evolution[23] and highlighting the need for quantification of disease burden beyond the CR category.[24]

In acute lymphoblastic leukemia (ALL),[25] acute promyelocytic leukemia,[26] and chronic myeloid leukemia[27] minimal residual disease (MRD) monitoring is a well-established standard of care. Measurement of MRD provides a more sensitive determination of disease

burden than CR. Next-generation flow (NGF) and next-generation sequencing (NGS) can detect as few as 1 in 10^6 clonal plasma cells. [28, 29] In patients with NDMM who achieve CR, most of the clinical benefit is associated with those who are MRD negative, while patients with MRD-positive CR have outcomes similar to those achieving a lesser response. [30, 31] Studies longitudinally examining PFS (and in some cases overall survival [OS]) stratifying subjects by MRD status have associated statistically superior survival outcomes with MRD-negative responses relative to MRD-positive responses,[29, 32–37] with deeper MRD response predicting longer survival.[36, 38] These results have raised the possibility of using MRD status in decision-making algorithms for subsequent treatments.[39]

The use of MRD status as a potential endpoint for clinical trials in multiple myeloma has been the subject of a US Food and Drug Administration (FDA) symposium.[40] The FDA, in collaboration with the Duke-Margolis Center for Health Policy, held a workshop on MRD as a surrogate endpoint in hematologic cancer trials, including multiple myeloma.[41] In order for an endpoint to be considered a surrogate "reasonably likely to predict clinical benefit" (that is, potentially acceptable for accelerated drug approval), it needs to meet two key criteria proposed by Prentice[42]: a surrogate endpoint must, 1) be correlated at a patient level with the clinical benefit endpoint independent of treatment, and, 2) fully capture the net effect of treatment on the clinical benefit endpoint. That is, the treatment effect on the surrogate endpoint must reliably predict the treatment effect on the clinical benefit endpoint, and not be merely a correlate of activity between two endpoint measurements.

The meta-analytic, multi-trial approach has become the preferred method for surrogate endpoint validation. However, in multiple myeloma, most published meta-analyses of MRD and PFS address only the first of the two Prentice criteria, i.e., only demonstrating correlation between MRD status and PFS, usually at the study level. Randomized studies with treatment effects on both MRD and PFS are required in order to assess the second criterion. Here we report the results of a meta-analysis of 6 randomized clinical trials for patients with NDMM that examined the potential of MRD status to fulfill both Prentice criteria and to qualify as a surrogate for PFS in this patient population.

Materials and Methods

Systematic Literature Review

To identify clinical trials in NDMM that included MRD data, we searched PubMed for the terms "myeloma," "minimal residual disease," and a publication type of "clinical trial." We also performed a similar search of abstracts from the American Society of Hematology (ASH) and the European Hematology Association (EHA) annual meetings. All studies included were for newly-diagnosed disease.

To be included in this meta-analysis, studies were required to meet four criteria: 1) a clinical trial must be randomized and only contain an experimental treatment and a control treatment, 2) MRD needed to be measured at a defined time point or period for the MRD study population, 3) the MRD subset must have been derived from a single study, and 4) it must have been a defined subset of the entire study population (e.g., CR) with details of the size of the MRD-tested population provided. Based on the principle that a surrogate

endpoint should capture the net effect of treatment on the clinical benefit endpoint, it is important to include a variety of treatments in the analysis of the surrogate endpoint to show the principle is met for any types of treatments. The studies with autologous stem cell transplantation (ASCT) following induction were also considered.

Statistical Analysis

We performed a meta-analysis of studies evaluating MRD and PFS in patients with NDMM in accordance with "Statistical Considerations for Surrogate Endpoints" presented at the FDA workshop on MRD in acute myeloid leukemia (AML). [1] The guidelines for establishing a validated surrogate endpoint are based on the two criteria suggested by Prentice.[42] Hypothesis testing for the second criterion (need to fully capture the net effect of treatment on the clinical benefit endpoint) requires full patient-level data and is also highly dependent on the power levels for both the surrogate endpoint and the clinical endpoint, which can be different across studies.[43] The approaches using proportion explained (or the proportion of treatment effect imparted by the surrogate)[44] or relative effect (or the ratio of the overall treatment effect of the clinical endpoint to that of the surrogate endpoint) [43] on the study-level correlation between the treatment effect on the surrogate endpoint and the clinical benefit endpoint were suggested to evaluate the second component of Prentice's criteria instead of the strict application of hypothesis testing which requires patient-level data. The study-level plot of treatment effect on PFS (a surrogate endpoint) versus treatment effect on OS (a clinical benefit endpoint) in stage 3 colorectal cancer was considered useful in evaluating whether the surrogate endpoint meets the second condition reasonably well when the strict evaluation of the second condition is not feasible.[45] In his paper, Fleming suggested 4 hierarchical levels of outcome measures. The highest level, or level 1, represents a bona fide clinical benefit endpoint. Level 2 designates a fully validated surrogate endpoint. While level 3 surrogate endpoint has yet to be validated, statistical analyses suggest that the net effect of treatment on the true clinical efficacy measure (e.g., PFS) is consistent with what would be predicted by the treatment effect on the outcome measure (e.g., MRD). Finally, the lowest level, or level 4, coincides with the first Prentice criterion and indicates an endpoint that only demonstrates the correlation with clinical benefit endpoint but fails to show the treatment effect level correlation. His paper notes that level 3 endpoints have been used as the basis for accelerated approval. In the absence of patient-level data and due to the small number of randomized studies (6) with treatment effects on both MRD and PFS, the goal of this meta-analysis was to examine MRD status as a level 3 endpoint using the study-level plot of two treatment effects to demonstrate whether MRD status meets the second Prentice criterion as described above.

MRD-negativity rates with exact 95% binomial confidence intervals (CIs) were calculated for each study and each treatment arm in randomized studies. Except for the CLARION study, MRD was evaluated in subsets of patients with the defined clinical responses. Imputation was not performed on the MRD status of the samples in the ALCYONE study since the MRD status for each arm was fully measured in the defined subset. To be consistent with the recently reported MRD-negativity rates in the CLARION study, the MRD-negativity rates were not further imputed for this study. In the four remaining studies, samples that were not measured for MRD and within the subset specified for MRD

assessment had their MRD status imputed from the samples with known MRD status. Among the three smaller studies with imputed MRD-negativity rates (IFM/DFCI 2009, GEM2005MAS65, and NCT00531453), the proportion of patients with known MRD status within the subsets specified for MRD assessment varied from 70% to 76%. As for the most extensive study, EMN02/HO95, the proportion of patients with known MRD status was 33%. Patients with responses that were excluded from planned MRD assessment were considered MRD positive. Odds ratios (ORs) and 95% CIs for randomized studies were calculated based on the MRD-negativity rate. The reported hazard ratios (HRs) for PFS from the 6 randomized studies were used. The scatter plot was drawn with the hazard and odds ratios in a logarithmic scale from the 6 randomized studies. Weighted least squares linear regression was implemented to predict HR from OR, where the weights were the sample sizes of the studies. The weighted linear regression line and the corresponding 95% confidence intervals (CIs) in the shaded area were added to help visualize the relationship between the hazard and odds ratios of the studies. Statistical analyses were performed using R software version 3.5.0.

Results

Study Selection

As depicted in Figure 1, 39 reports were initially identified in PubMed. Of these, 10 were eliminated for lack of relevance (6 included allogeneic transplant, 1 included ALL, 1 included a bone marrow graft, and 2 had no available MRD data), 7 were eliminated as duplicates, and 19 did not meet the criteria for inclusion. The 3 remaining studies were included in the meta-analysis (Table 1). [29, 34, 46, 47] A similar examination of ASH and EHA abstracts for the years 2014–2015 and 2017–2018 uncovered 2 additional studies, and the newly available CLARION study (NCT01818752; data on file) was also included. A recent publication[50] on CLARION reported the MRD results using a cutoff of 10^{-6}. However, in order to reduce the variation in MRD depth between the 6 randomized trials and to harmonize the MRD depth of response to the IMWG recommended cutoff[48], the MRD results were reported with a cutoff of 10^{-5} in this meta-analysis. In the end, 6 randomized studies[29, 34, 46, 47, 49–51] met all inclusion criteria and were included in this meta-analysis (Table 1). Collectively these encompass data for 3283 NDMM patients and 2208 MRD samples. Within these studies, the definitions of MRD varied from 1 in 10^4 to 1 in 10^5. Times at which MRD was assessed and subgroups included in the assessment also varied, the latter changing from patients that consented to MRD assessment to only patients with a CR (or suspected CR). MRD assessments by multiparameter flow cytometry (MFC), NGF and NGS were included.

Meta-analysis

Figure 2 depicts the study-level correlation of treatment effect on MRD status with the treatment effect on PFS in the 6 randomized trials using odds ratios (ORs) and hazard ratios (HRs). The MRD-negativity rate for each of the treatments ranged from 6% to 70% (Table 2). The MRD-negativity rates and the associated median PFS for the trials with MRD cutoffs from 10^{-4} to 10^{-5} were shown (Figure 2). Each treatment arm that had a higher MRD-negativity rate corresponded to a longer median PFS.

The two studies with significantly large treatment effects were IFM/DFCI 2009 and ALCYONE, which had ORs of 2.31 (95% CI, 1.69, 3.15) and 4.35 (95% CI, 2.64, 7.17), respectively, and also showed substantially lower risks of progression (HR, 0.69; 95% CI, 0.56, 0.84 and HR, 0.50; 95% CI, 0.38, 0.65, respectively). [47, 49] In other words, the large treatment effect on the experimental treatment arm resulted in a significantly higher proportion of MRD-negative patients when compared with the control treatment arm in IFM/DFCI 2009 and ALCYONE. The HR of 1.20 in Ludwig's study (NCT00531453) corresponded to the OR that was lower than one, indicating a higher MRD-negativity rate and a longer median PFS for the control treatment arm (Figure 2).[29] The treatment effects from the GEM2005MAS65 and EMN02/HO95 studies on MRD and PFS were modest. The CLARION study did not lead to a treatment effect for either measurement. Even though these studies span treatment effects of various magnitudes, a correlation was clearly observed between ORs and HRs as illustrated by a negatively trended and weighted regression line in logarithmic scale among the 6 randomized studies (Figure 3). All 6 randomized trials fell within the 95% CI boundary of the regression line with a coefficient of determination of 0.97. Within the limitations of the types of treatments conducted in these trials, this regression line predicts that for any two randomized patient populations, the treatment with the larger proportion of subjects with an MRD-negative response will experience a lower risk of disease progression. Together, these results show that the treatment effect on the MRD-negativity rate captures nearly the full treatment effect on PFS; thus suggesting the second Prentice criterion to the point where a level 3 endpoint is demonstrated.

Discussion

This meta-analysis examined the surrogacy of MRD negativity for PFS in patients with NDMM. We identified 6 randomized trials in which patients with previously untreated multiple myeloma were evaluated for MRD status following defined therapies and in which both the prevalence of MRD negativity and clinical outcomes for the entire patient population were fully reported. Our analysis mainly focused on PFS rather than OS for survival outcome because OS results were immature in most of these studies and may be unable to capture the full treatment effect on clinical outcome. The analysis included 3283 transplant eligible and ineligible patients treated with various combinations of IMiDs, proteasome inhibitors, daratumumab, with or without high-dose melphalan and stem cell transplant. The studies covered a range of MRD response rates with a variety of treatments. Limitations of this analysis include the use of study-level data, the limited number of studies available, differing cutoff levels for MRD (10^{-4} to 10^{-5}), low cutoff level of 10^{-4} in 3 of the studies, different MRD measurement approaches, the imputation of MRD status for missing measurements in a defined subset of patients, and the measurement of MRD in various response subgroups with dissimilar predefined time periods.

Meta-analysis of these data revealed a clear association between MRD negativity and improved PFS. Comparing all study arms, MRD-negativity rates, as measured by MFC, NGF, and NGS, increased as PFS increased. These results are consistent with published data supporting the correlation of MRD responses with clinical benefit in multiple myeloma. [31, 32, 36, 38] To date, however, published studies have yet to establish MRD as a surrogate

endpoint for PFS or OS according to the guidelines proposed by Prentice[42] and Fleming.[45] To address this, the present study investigated whether the effect of the intervention on MRD is predictive of the level of the effect on PFS, an important criterion for establishing MRD as at least a non-validated surrogate endpoint (level 3), rather than a simple correlate of two measurements (level 4 endpoint), so that it may be considered a basis for accelerated approval.[45] Among randomized studies the treatment effects on MRD-negativity rates were consistent in magnitude with the treatment effects on PFS (Figure 2 and 3), thus supporting the fulfillment of the second of the Prentice criteria.

The results of two additional meta-analyses of MRD status in patients with NDMM showed a statistically significant correlation between MRD-negative status and PFS. In the first, which analyzed four studies including one of those included here, [34] a random effects model, which weighted studies via the inverse-variance method, found that MRD-negativity status significantly correlated with PFS (HR, 0.35; $P<0.001$) and OS (HR, 0.48; $P<0.001$).[52] Another study analyzed data from 14 trials, one of which is included here, [46] using the Peto methodology,[53] and also found a statistically significant correlation between MRD negativity and PFS (HR, 0.41; $P<0.001$) as well as OS (HR, 0.57; $P<0.001$).[54] Although in line with our findings, some differences in the methodologies employed deserve comment. Both publications examined the effect of MRD-negative status among patients with available MRD status data on treatment effects in the trial arms. Their conclusions, therefore, apply to patients who were tested for MRD status, which is generally done only for those achieving VGPR or CR. Because this may represent a lower risk subset of patients than the overall population of the trial arm, and no assumptions were made regarding MRD status of patients not achieving VGPR or CR, the conclusions reached may not be extrapolated to the full trial arm populations, which included many patients not tested for MRD status.

The analysis that we used here went beyond evaluating the correlation of depth of MRD response with clinical outcome. Within each trial, we determined the full impact of each treatment on the probability of achieving an MRD-negative response for the entire patient population. This approach has the advantage of capturing the full treatment effect of MRD response on survival. Although the analysis obscures the impact of MRD response on survival, it unmasks the relationship between each treatment's ability to produce a deep response and the effect of each treatment on survival. This takes a major step closer toward establishing a causal link between the relative change in MRD status and the relative change in survival benefit. Furthermore, our approach is in line with the methodology outlined at the FDA workshop on MRD in AML. [1]

Irrespective of the method used for meta-analysis, a significant concern in examining MRD as a potential surrogate endpoint in NDMM trials has been the inconsistent definition of MRD negativity and the lack of standardization in MRD testing methods. [55, 56] To address this concern, the IMWG published guidelines defining the criteria for response and MRD assessment in multiple myeloma. [48] The use of these guidelines eliminates former inconsistencies and will allow cross-study comparisons of data and the development of treatment algorithms based on MRD status. Encouragingly, recent evidence suggests that the necessary standardization in MRD testing is, in fact, being implemented.[57]

Conclusions

In conclusion, results of this meta-analysis support the use of MRD negativity as a level 3 surrogate endpoint that likely predicts PFS in NDMM.

Acknowledgments

Abbreviations

ASCT	autologous stem cell transplant
CI	confidence interval
CR	complete response
FDA	Food and Drug Administration
HR	hazard ratio
IMWG	International Myeloma Working Group
KMP	carfilzomib, melphalan, and prednisone
MFC	multiparametric flow cytometry
MRD	minimal residual disease
NDMM	newly diagnosed multiple myeloma
NE	not estimable
NGF	next-generation flow
NGS	next-generation sequencing
OR	odds ratio
OS	overall survival
PFS	progression-free survival

RVD	lenalidomide, bortezomib, and dexamethasone
sCR	stringent complete response
VGPR	very good partial response
VMP	bortezomib, melphalan, and prednisone
VTD	bortezomib, thalidomide, and dexamethasone
VTDC	bortezomib, thalidomide, dexamethasone, and cyclophosphamide
VTP	bortezomib, thalidomide, and prednisone

References

1. Hokland P, Cotter F. Readying the minimal residual disease concept in acute myeloid leukaemia for prime time - the American way. Br J Haematol. 2013;162:429–430. [PubMed: 23772825]

2. Moreau P, Masszi T, Grzasko N, et al. Ixazomib, an Investigational Oral Proteasome Inhibitor (PI), in Combination with Lenalidomide and Dexamethasone (IRd), Significantly Extends Progression-Free Survival (PFS) for Patients (Pts) with Relapsed and/or Refractory Multiple Myeloma (RRMM): The Phase 3 Tourmaline-MM1 Study. Blood. 2015;126:727–727.

3. Stewart AK, Rajkumar SV, Dimopoulos MA, et al. Carfilzomib, lenalidomide, and dexamethasone for relapsed multiple myeloma. N Engl J Med. 2015;372:142–152. [PubMed: 25482145]

4. Durie BG, Hoering A, Abidi MH, et al. Bortezomib with lenalidomide and dexamethasone versus lenalidomide and dexamethasone alone in patients with newly diagnosed myeloma without intent for immediate autologous stem-cell transplant (SWOG S0777): a randomised, open-label, phase 3 trial. Lancet. 2017;389:519–527. [PubMed: 28017406]

5. Benboubker L, Dimopoulos MA, Dispenzieri A, et al. Lenalidomide and dexamethasone in transplant-ineligible patients with myeloma. N Engl J Med. 2014;371:906–917. [PubMed: 25184863]

6. San Miguel J, Weisel K, Moreau P, et al. Pomalidomide plus low-dose dexamethasone versus high-dose dexamethasone alone for patients with relapsed and refractory multiple myeloma (MM-003): a randomised, open-label, phase 3 trial. Lancet Oncol 2013;14:1055–1066. [PubMed: 24007748]

7. Lokhorst HM, Plesner T, Laubach JP, et al. Targeting CD38 with Daratumumab Monotherapy in Multiple Myeloma. N Engl J Med. 2015;373:1207–1219. [PubMed: 26308596]

8. Lonial S, Dimopoulos M, Palumbo A, et al. Elotuzumab Therapy for Relapsed or Refractory Multiple Myeloma. N Engl J Med. 2015;373:621–631. [PubMed: 26035255]

9. Landgren O, Owen RG. Better therapy requires better response evaluation: Paving the way for minimal residual disease testing for every myeloma patient. Cytometry B Clin Cytom 2016;90:14–20. [PubMed: 26147584]

10. Attal M, Lauwers-Cances V, Marit G, et al. Lenalidomide maintenance after stem-cell transplantation for multiple myeloma. N Engl J Med. 2012;366:1782–1791. [PubMed: 22571202]

11. McCarthy PL, Owzar K, Hofmeister CC, et al. Lenalidomide after stem-cell transplantation for multiple myeloma. N Engl J Med. 2012;366:1770–1781. [PubMed: 22571201]

12. Gandolfi S, Prada CP, Richardson PG. How I treat the young patient with multiple myeloma. Blood. 2018;132:1114–1124. [PubMed: 29967130]

13. Touzeau C, Moreau P, Dumontet C. Monoclonal antibody therapy in multiple myeloma. Leukemia. 2017;31:1039–1047. [PubMed: 28210004]

14. Sargent D, Shi Q, Yothers G, et al. Two or three year disease-free survival (DFS) as a primary end-point in stage III adjuvant colon cancer trials with fluoropyrimidines with or without oxaliplatin or irinotecan: data from 12,676 patients from MOSAIC, X-ACT, PETACC-3, C-06, C-07 and C89803. Eur J Cancer. 2011;47:990–996. [PubMed: 21257306]

15. Shi Q, Flowers CR, Hiddemann W, et al. Thirty-Month Complete Response as a Surrogate End Point in First-Line Follicular Lymphoma Therapy: An Individual Patient-Level Analysis of Multiple Randomized Trials. J Clin Oncol 2016:JCO2016708651.

16. Lonial S, Anderson KC. Association of response endpoints with survival outcomes in multiple myeloma. Leukemia. 2014;28:258–268. [PubMed: 23868105]

17. Palumbo A, Chanan-Khan A, Weisel K, et al. Daratumumab, Bortezomib, and Dexamethasone for Multiple Myeloma. N Engl J Med. 2016;375:754–766. [PubMed: 27557302]

18. Dimopoulos M, Spencer A, Attal M, et al. Lenalidomide plus dexamethasone for relapsed or refractory multiple myeloma. N Engl J Med. 2007;357:2123–2132. [PubMed: 18032762]

19. Harousseau JL, Palumbo A, Richardson PG, et al. Superior outcomes associated with complete response in newly diagnosed multiple myeloma patients treated with nonintensive therapy: analysis of the phase 3 VISTA study of bortezomib plus melphalan-prednisone versus melphalan-prednisone. Blood. 2010;116:3743–3750. [PubMed: 20628153]

20. Richardson PG, Sonneveld P, Schuster MW, et al. Bortezomib or high-dose dexamethasone for relapsed multiple myeloma. N Engl J Med. 2005;352:2487–2498. [PubMed: 15958804]

21. Weber DM, Chen C, Niesvizky R, et al. Lenalidomide plus dexamethasone for relapsed multiple myeloma in North America. N Engl J Med. 2007;357:2133–2142. [PubMed: 18032763]

22. van de Velde HJ, Liu X, Chen G, Cakana A, Deraedt W, Bayssas M. Complete response correlates with long-term survival and progression-free survival in high-dose therapy in multiple myeloma. Haematologica. 2007;92:1399–1406. [PubMed: 18024376]

23. Keats JJ, Chesi M, Egan JB, et al. Clonal competition with alternating dominance in multiple myeloma. Blood. 2012;120:1067–1076. [PubMed: 22498740]

24. Mailankody S, Korde N, Lesokhin AM, et al. Minimal residual disease in multiple myeloma: bringing the bench to the bedside. Nat Rev Clin Oncol. 2015;12:286–295. [PubMed: 25622976]

25. van Dongen JJ, van der Velden VH, Bruggemann M, Orfao A. Minimal residual disease diagnostics in acute lymphoblastic leukemia: need for sensitive, fast, and standardized technologies. Blood. 2015;125:3996–4009. [PubMed: 25999452]

26. De Angelis F, Breccia M. Molecular Monitoring as a Path to Cure Acute Promyelocytic Leukemia. Rare Cancers Ther. 2015;3:119–132. [PubMed: 27182481]

27. Savona MR. Molecular monitoring and minimal residual disease in the management of chronic myelogenous leukemia. J Community Support Oncol. 2014;12:171–178. [PubMed: 24971427]

28. Flores-Montero J, Sanoja-Flores L, Paiva B, et al. Next Generation Flow for highly sensitive and standardized detection of minimal residual disease in multiple myeloma. Leukemia. 2017;31:2094–2103. [PubMed: 28104919]

29. Ludwig H, Greil R, Masszi T, et al. Bortezomib, thalidomide and dexamethasone, with or without cyclophosphamide, for patients with previously untreated multiple myeloma: 5-year follow-up. Br J Haematol. 2015;171:344–354. [PubMed: 26153365]

30. Lahuerta JJ, Paiva B, Vidriales MB, et al. Depth of Response in Multiple Myeloma: A Pooled Analysis of Three PETHEMA/GEM Clinical Trials. J Clin Oncol 2017;35:2900–2910. [PubMed: 28498784]

31. Paiva B, Vidriales MB, Cervero J, et al. Multiparameter flow cytometric remission is the most relevant prognostic factor for multiple myeloma patients who undergo autologous stem cell transplantation. Blood. 2008;112:4017–4023. [PubMed: 18669875]

32. Korde N, Roschewski M, Zingone A, et al. Treatment With Carfilzomib-Lenalidomide-Dexamethasone With Lenalidomide Extension in Patients With Smoldering or Newly Diagnosed Multiple Myeloma. JAMA Oncol. 2015;1:746–754. [PubMed: 26181891]

33. Ladetto M, Ferrero S, Drandi D, et al. Prospective molecular monitoring of minimal residual disease after non-myeloablative allografting in newly diagnosed multiple myeloma. Leukemia. 2016;30:1211–1214. [PubMed: 26437784]

34. Mateos MV, Oriol A, Martinez-Lopez J, et al. GEM2005 trial update comparing VMP/VTP as induction in elderly multiple myeloma patients: do we still need alkylators? Blood. 2014;124:1887–1893. [PubMed: 25102853]

35. Paiva B, Gutierrez NC, Rosinol L, et al. High-risk cytogenetics and persistent minimal residual disease by multiparameter flow cytometry predict unsustained complete response after autologous stem cell transplantation in multiple myeloma. Blood. 2012;119:687–691. [PubMed: 22128143]

36. Rawstron AC, Child JA, de Tute RM, et al. Minimal residual disease assessed by multiparameter flow cytometry in multiple myeloma: impact on outcome in the Medical Research Council Myeloma IX Study. J Clin Oncol. 2013;31:2540–2547. [PubMed: 23733781]

37. Roussel M, Lauwers-Cances V, Robillard N, et al. Front-line transplantation program with lenalidomide, bortezomib, and dexamethasone combination as induction and consolidation followed by lenalidomide maintenance in patients with multiple myeloma: a phase II study by the Intergroupe Francophone du Myelome. J Clin Oncol 2014;32:2712–2717. [PubMed: 25024076]

38. Martinez-Lopez J, Lahuerta JJ, Pepin F, et al. Prognostic value of deep sequencing method for minimal residual disease detection in multiple myeloma. Blood. 2014;123:3073–3079. [PubMed: 24646471]

39. Landgren O, Giralt S. MRD-driven treatment paradigm for newly diagnosed transplant eligible multiple myeloma patients. Bone Marrow Transplant. 2016;51:913–914. [PubMed: 26926231]

40. Gormley NJ, Turley DM, Dickey JS, et al. Regulatory perspective on minimal residual disease flow cytometry testing in multiple myeloma. Cytometry B Clin Cytom. 2016;90:73–80. [PubMed: 26108351]

41. Administration USFaD. Minimal Residual Disease as a Surrogate Endpoint in Hematologic Cancer Trials. 2016.

42. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. Stat Med. 1989;8:431–440. [PubMed: 2727467]

43. Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H. The validation of surrogate endpoints in meta-analyses of randomized experiments. Biostatistics. 2000;1:49–67. [PubMed: 12933525]

44. Freedman LS, Graubard BI, Schatzkin A. Statistical validation of intermediate endpoints for chronic diseases. Stat Med. 1992;11:167–178. [PubMed: 1579756]

45. Fleming TR. Surrogate endpoints and FDA's accelerated approval process. Health Aff (Millwood). 2005;24:67–78. [PubMed: 15647217]

46. Ludwig H GR, Masszi T. Randomized phase 2 study of bortezomib, thalidomide, and dexamethasone with or without cyclophosphamide as induction therapy in previously untreated multiple myeloma (MM): Long-term follow-up results. Haematologica. 2014;99(s1):112–113.

47. Mateos MV, Dimopoulos MA, Cavo M, et al. Daratumumab plus Bortezomib, Melphalan, and Prednisone for Untreated Myeloma. N Engl J Med. 2018;378:518–528. [PubMed: 29231133]

48. Kumar S, Paiva B, Anderson KC, et al. International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. Lancet Oncol. 2016;17:e328–346. [PubMed: 27511158]

49. Avet-Loiseau H, Corre J, Lauwers-Cances V, et al. Evaluation of Minimal Residual Disease (MRD) By Next Generation Sequencing (NGS) Is Highly Predictive of Progression Free Survival in the IFM/DFCI 2009 Trial. Blood. 2015;126:191–191.

50. Cavo M, Hájek R, Pantani L, et al. Autologous Stem Cell Transplantation Versus Bortezomib-Melphalan-Prednisone for Newly Diagnosed Multiple Myeloma: Second Interim Analysis of the Phase 3 EMN02/HO95 Study. Blood. 2017;130:397–397. [PubMed: 28576879]

51. Oliva S, Hofste op Bruinink D, Íhová L, et al. Minimal residual disease (MRD) monitoring by multiparameter flow cytometry (MFC) in newly diagnosed transplant eligible multiple myeloma (MM) patients: Results from the EMN02/HO95 phase 3 trial. Journal of Clinical Oncology. 2017;35:8011–8011.

52. Landgren O, Devlin S, Boulad M, Mailankody S. Role of MRD status in relation to clinical outcomes in newly diagnosed multiple myeloma patients: a meta-analysis. Bone Marrow Transplant. 2016;51:1565–1568. [PubMed: 27595280]

53. Peto R Why do we need systematic overviews of randomized trials? Stat Med 1987;6:233–244. [PubMed: 3616281]

54. Munshi NC, Avet-Loiseau H, Rawstron AC, et al. Association of Minimal Residual Disease With Superior Survival Outcomes in Patients With Multiple Myeloma: A Meta-analysis. JAMA Oncol. 2016.

55. Flanders A, Stetler-Stevenson M, Landgren O. Minimal residual disease testing in multiple myeloma by flow cytometry: major heterogeneity. Blood. 2013;122:1088–1089. [PubMed: 23929839]

56. Roschewski M, Stetler-Stevenson M, Yuan C, Mailankody S, Korde N, Landgren O. Minimal residual disease: what are the minimum requirements? J Clin Oncol 2014;32:475–476. [PubMed: 24419126]

57. Salem D, Stetler-Stevenson M, Yuan C, Landgren O. Myeloma minimal residual disease testing in the United States: Evidence of improved standardization. Am J Hematol. 2016;91:E502–E503. [PubMed: 27556705]

Clinical Practice Points

- Using PFS as a surrogate endpoint for clinical trials may no longer be practical as recent therapeutic advances in newly diagnosed multiple myeloma (NDMM) have greatly extended survival times, resulting in increasingly lengthy trials; a new surrogate endpoint that likely predicts clinical benefit and accelerates drug approval is warranted

- Meta-analysis was performed on 6 randomized trials with a total of 3283 NDMM patients at study level to evaluate the potential of MRD status as a surrogate endpoint

- Treatment effect on MRD negativity has a strong positive correlation with the treatment effect on PFS for all randomized studies, demonstrating the ability of MRD status to predict clinical benefit

- The results of this study-level meta-analysis support the use of MRD status as a surrogate endpoint in clinical trials for patients with NDMM
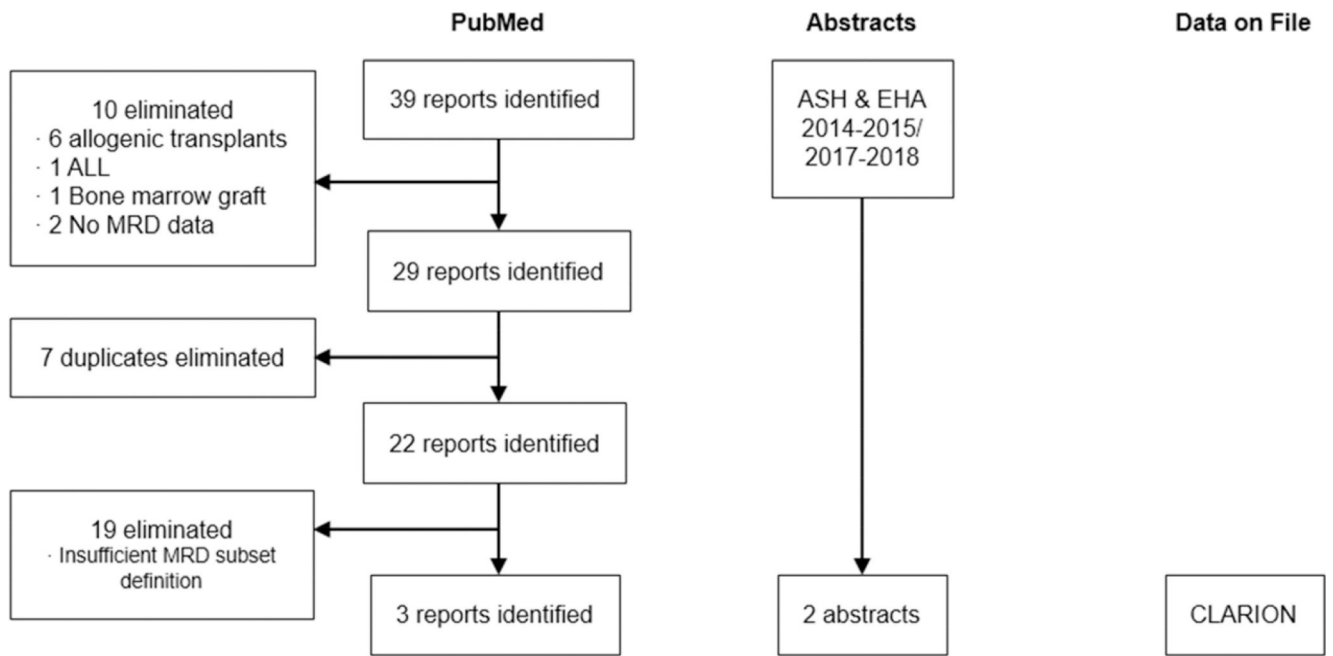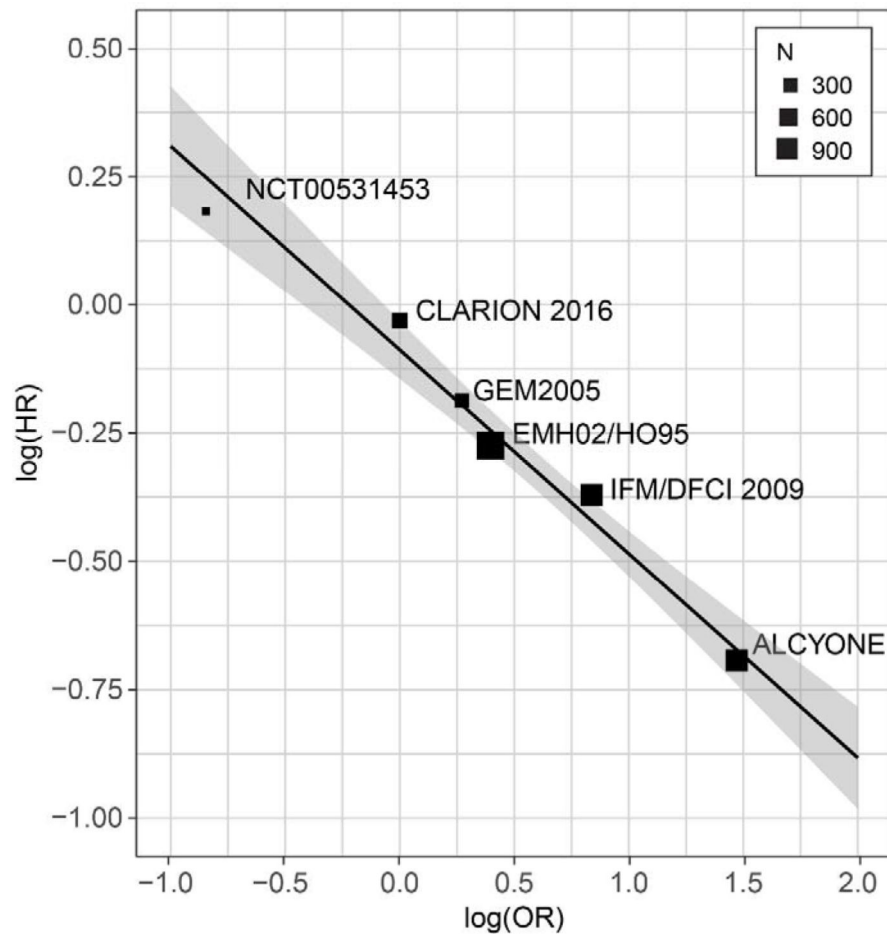
Figure 1.
Study selection flowchart. Results of literature search using PubMed and manual searches of the American Society of Hematology and the European Hematology Association annual meeting abstracts are shown. Additional data were included from the CLARION study. ALL, acute lymphoblastic leukemia; MRD, minimal residual disease.

	Sample Size (N)	MRD- (n, [%])	Median PFS (months)	Hazard Ratio [95% CI] (T/Co)		Odds Ratio [95% CI] (T/Co)	
IFM/DFCI 2009	700	423 (60)	—	0.69	0.69 [0.56, 0.84]	2.31	2.31 [1.69, 3.15]
RVD+ASCT (T)	350	246 (70)	43		—		—
RVD (Co)	350	177 (51)	34		—		—
GEM2005	260	34 (13)	31	0.83	0.83 [0.6, 1.1]	1.31	1.31 [0.64, 2.71]
VMP (T)	130	19 (15)	32		—		—
VTP (Co)	130	15 (12)	23		—		—
NCT00531453	98	42 (43)	—	1.20	1.20 [0.74, 1.97]	0.43	0.43 [0.19, 0.97]
VTDC (T)	49	16 (33)	36.3		—		—
VTD (Co)	49	26 (53)	56.3		—		—
ALCYONE	706	100 (14)	—	0.50	0.50 [0.38, 0.65]	4.35	4.35 [2.64, 7.17]
Dara+VMP (T)	350	78 (22)	NE		—		—
VMP (Co)	356	22 (6.2)	18.1		—		—
EMN02/HO95	1192	814 (68)	—	0.76	0.76 [0.64, 0.91]	1.49	1.49 [1.17, 1.88]
ASCT (T)	695	449 (65)	NE		—		—
VMP (Co)	497	274 (55)	44		—		—
CLARION	327	65 (20)	24	0.97	0.97 [0.71, 1.32]	1.00	1.00 [0.58, 1.72]
KMP (T)	166	33 (20)	24		—		—
VMP (Co)	161	32 (20)	24		—		—

0.5 1.0 1.5 2.0 2.5   0 1 2 3 4 5 6 7 8

Figure 2.
Treatment effect on progression-free survival and minimal residual disease for randomized studies using multiparameter flow cytometry, next-generation flow, and next-generation sequencing. MRD-negativity rates were calculated based on sample size and patients that were excluded from MRD assessment were MRD positive.

ASCT, autologous stem cell transplant; Co, control; CI, confidence interval; Dara, daratumumab; KMP, carfilzomib, melphalan, and prednisone; MRD, minimal residual disease; PFS, progression-free survival; RVD, lenalidomide, bortezomib, and dexamethasone; T, treatment; VMP, bortezomib, melphalan, and prednisone; VTD, bortezomib, thalidomide, and dexamethasone; VTDC, bortezomib, thalidomide, dexamethasone, and cyclophosphamide; VTP, bortezomib, thalidomide, and prednisone.

Figure 3.
Weighted regression plot comparing the MRD odds ratios with the PFS hazard ratios in a logarithmic scale among the 6 randomized trials (adjusted R-squared of 0.97). The equation for the weighted regression model is log (HR) = −0.4 * log (OR) − 0.09.

Table 1.

Summary of studies included in the meta-analysis.

Study	Treatment	Phase/Design	N[a]	MRD Method	MRD Cutoff	When Assessed	MRD Subgroup
IFM/DFCI 2009[49]	RVD + ASCT vs RVD	Ph 3 randomized	700/581	7-color FCM	10^{-4}	Pre-maintenance	At least VGPR
GEM2005MAS65MAS65[34]	VMP vs VTP	Ph3 randomized	260/153	4-color FCM	10^{-4}	After 6 induction cycles	At least VGPR
NCT00531453[29, 46]	VTDC vs VTD	Ph 2 randomized	98/58	MFC	10^{-4}[b]	At time of CR or suspected CR	Bone-marrow confirmed CR
ALCYONE[47]	Dara+ VMP vs VMP	Ph 3 randomized	706/236	NGS	10^{-5}	At time of CR or sCR	At CR or sCR
EMN02/HO95[50, 51]	ASCT vs VMP	Ph 3 randomized	1192/957	8-color FCM (EuroFlow)	10^{-5}	Pre-maintenance	At least VGPR
CLARION[c] (*Data on file*)	KMP vs VMP	Ph 3 randomized	327/223	NGF	10^{-5}	At end of treatment	Consented for MRD assessment

ASCT, autologous stem cell transplant; CR, complete response; FCM, flow cytometry; KMP, carfilzomib, melphalan, and prednisone; MFC, multiparametric flow cytometry; MRD, minimal residual disease; NGF, next generation flow; NGS, next generation sequencing; RVD, lenalidomide, bortezomib, and dexamethasone; sCR, stringent CR; VGPR, very good partial response; VMP, bortezomib, melphalan, and prednisone; VTD, bortezomib, thalidomide, and dexamethasone; VTDC, bortezomib, thalidomide, dexamethasone, and cyclophosphamide; VTP, bortezomib, thalidomide, and prednisone.

[a]Total/Available MRD samples.

[b]Not specified; cutoff was assumed.

[c]CLARION study data excludes patients in China.

Table 2.

Minimal residual disease negativity.

Study	Regimen	Sample Size	MRD Availability	MRD Negativity[a]	MRD Negativity[b] Rate (95% CI)
IFM/DFCI 2009[49]	— RVD + ASCT RVD	700 350 350	581 (at least VGPR) 308 273	423 246 177	0.60 (0.57, 0.64) 0.70 (0.65, 0.75) 0.51 (0.45, 0.56)
GEM2005MAS65[34]	—VMP VTP	260 130 130	153 (at least PR) 79 74	34 19 15	0.13 (0.09, 0.18) 0.15 (0.09, 0.22) 0.12 (0.07, 0.18)
NCT00531453[46]	—VTDC VTD	98 49 49	58 (suspected CR) 26 32	42 16 26	0.43 (0.33, 0.53) 0.33 (0.2, 0.48) 0.53 (0.38, 0.67)
ALCYONE[47]	— Dara+VMP VMP	706 350 356	236 (at CR or sCR) 149 87	100 78 22	0.14 (0.12, 0.17) 0.22 (0.18, 0.27) 0.06 (0.04, 0.09)
EMN02/HO95[50,51]	—ASCT VMP	1192 695 497	957 (at least VGPR) 584 373	814 449 274	0.68 (0.66, 0.71) 0.65 (0.61, 0.68) 0.55 (0.50, 0.59)
CLARION	—KMP VMP	327 166 161	223 (EOT) 113 110	65 33 32	0.20 (0.14, 0.27) 0.20 (0.14, 0.27) 0.20 (0.14, 0.27)

ASCT, autologous stem cell transplant; CI, confidence interval; CR, complete response; Dara, daratumumab; EOT, end of treatment; KMP, carfilzomib, melphalan, and prednisone; MRD, minimal residual disease; PFS, progression-free survival; PR, partial response; RVD, lenalidomide, bortezomib, and dexamethasone; VMP, bortezomib, melphalan, and prednisone; sCR, stringent complete response; VTD, bortezomib, thalidomide, and dexamethasone; VTDC, bortezomib, thalidomide, dexamethasone, and cyclophosphamide; VTP, bortezomib, thalidomide, and prednisone

[a] MRD negativity was calculated based on the sample size assuming those patients without an MRD status were MRD positive.

[b] MRD-negativity rate is presented as the number of MRD-negative patients divided by the total number of patients in the trial arm (sample size).