



Published in final edited form as:

*Leukemia*. 2019 July ; 33(7): 1713–1722. doi:10.1038/s41375-018-0329-0.

## Combination of Flow Cytometry and Functional Imaging for Monitoring of Residual Disease in Myeloma

L Rasche<sup>1,2</sup>, D Alapat<sup>3</sup>, M Kumar<sup>4</sup>, G Gershner<sup>1</sup>, J McDonald<sup>4</sup>, C.P Wardell<sup>1</sup>, R Samant<sup>4</sup>, R Van Hemert<sup>4</sup>, J Epstein<sup>1</sup>, A.F Williams<sup>3</sup>, S Thanendrarajan<sup>1</sup>, C Schinke<sup>1</sup>, M Bauer<sup>1</sup>, C Ashby<sup>1</sup>, R.G Tytarenko<sup>1</sup>, F van Rhee<sup>1</sup>, B.A Walker<sup>1</sup>, M Zangari<sup>1</sup>, B Barlogie<sup>1</sup>, F.E Davies<sup>1</sup>, G.J Morgan<sup>1</sup>, and N Weinhold<sup>1</sup>

<sup>1</sup>Myeloma Institute, University of Arkansas for Medical Sciences, Little Rock, AR, USA

<sup>2</sup>Department of Internal Medicine 2, University Hospital of Würzburg, Würzburg, Germany

<sup>3</sup>Pathology Department, University of Arkansas for Medical Sciences, Little Rock, AR, USA

<sup>4</sup>Radiology Department, University of Arkansas for Medical Sciences, Little Rock, AR, USA

### Abstract

The iliac crest is the sampling site for minimal residual disease (MRD) monitoring in Multiple Myeloma (MM). However, the disease distribution is often heterogeneous, and imaging can be used to complement MRD detection at a single site. We have investigated patients in complete remission (CR) during first-line or salvage therapy, for whom MRD flow-cytometry and the two imaging modalities positron-emission-tomography (PET) and diffusion-weighted magnetic resonance imaging (DW-MRI) were performed at the onset of CR. Residual focal lesions (FLs), detectable in 24% of first-line patients, were associated with short progression-free survival (PFS), with DW-MRI detecting disease in more patients. In some patients, FLs were only PET-positive, indicating that the two approaches are complementary. Combining MRD and imaging improved prediction of outcome, with double-negative and double-positive features defining groups with excellent and dismal PFS, respectively. FLs were a rare event (12%) in first-line MRD-negative CR patients. In contrast, patients achieving an MRD-negative CR during salvage therapy frequently had FLs (50%). Multi-region sequencing and imaging in an MRD-negative patient showed persistence of spatially separated clones. In conclusion, we show that DW-MRI is a

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:[http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

**Corresponding Author:** Niels Weinhold, PhD, Myeloma Institute, University of Arkansas for Medical Sciences, 4301 W. Markham, #816, Little Rock, AR 72205, [nweinhold@uams.edu](mailto:nweinhold@uams.edu), +1 (501) 916 0543.

#### AUTHORSHIP CONTRIBUTIONS

Conception and design: Weinhold N, Rasche L, Morgan GJ

Provision of study material or patients: Alapat D, Morgan GJ, Barlogie B, van Rhee F, Zangari M, Thanendrarajan S, Schinke C, Davies FE, Epstein, J, Williams, A.F.

Reporting imaging: Kumar, M, McDonald J, Samant R, Van Hemert R

Data analysis: Weinhold N, Rasche L, Gershner, G, Alapat, D, Ashby C, Bauer M, Wardell, CP, Walker BA

Wrote the paper: Rasche L, Weinhold N, Morgan GJ

#### Competing Interests

Bart Barlogie is a co-inventor on patents and patent applications related to use of GEP in cancer medicine that have been licensed to Quest diagnostics. The remaining authors declare no conflict of interest.

promising tool for monitoring residual disease that complements PET and should be combined with MRD.

## INTRODUCTION

Response rates in Multiple Myeloma (MM) have improved dramatically, and complete responses (CR) are observed in up to 80% of patients<sup>1</sup>. However, in CR patients minimal residual disease (MRD) frequently persists, which cannot be detected using the traditional combination of serum markers and bone marrow (BM) histology<sup>2</sup>. Flow-cytometry is an established tool for MRD monitoring, and it can identify a subgroup of patients with very favorable outcomes<sup>3-6</sup>. As a result, MRD-negativity has been defined as a response criterion by the International Myeloma Working Group (IMWG). Nevertheless, despite prolonged progression-free survival (PFS) in these patients, relapses are still observed, raising a question as to the origin of resistant tumor cells.

The iliac crest area is the usual sampling site for MRD testing, but the disease distribution is often heterogeneous, and spatial genomic heterogeneity is a common phenomenon in MM<sup>7-12</sup>. Clones from focal lesions (FL) or extramedullary disease can have an unfavorable genomic profile potentially leading to differential treatment response<sup>9, 11</sup>. In this context, a low concordance between MRD (with a sensitivity of  $1 \times 10^{-4}$ ) and <sup>18</sup>F-fluorodeoxyglucose (FDG) positron emission tomography (PET) has been reported, including patients with FDG-avid disease but negative MRD<sup>13</sup>. In the same study, concomitant MRD-negativity and normalized PET scans defined a group with very favorable outcome, supporting the conclusion that PET is complementary to MRD for assessment of residual disease<sup>13</sup>. We recently showed that diffusion-weighted magnetic resonance imaging (DW-MRI) can detect disease in a higher proportion of newly diagnosed MM (NDMM) patients than PET, as it is independent of the tumor metabolism<sup>14, 15</sup>. Yet, its performance for monitoring of residual disease has not been described.

The aims of this study were 1) to compare DW-MRI to PET for the detection of residual FLs in patients achieving CR, and 2) to test whether DW-MRI and PET could complement MRD with a sensitivity of  $1 \times 10^{-5}$ . To address these aims, we investigated MM patients who achieved CR during Total Therapies (TT), TT-like regimens or later during salvage therapy, for whom PET and DW-MRI were available together with MRD data (Fig. 1). To account for differences in response kinetics (e.g. time to best response is prolonged in patients of the CD-2 molecular subgroup<sup>16</sup>), we chose the onset of CR as an individual and flexible time point; a strategy which has recently been used for relapsed patients receiving Daratumumab<sup>17, 18</sup>. The study highlights the power of combining multiple diagnostic tests to address the negative impact of spatial tumor heterogeneity in MM on MRD diagnostics.

## METHODS

### Patients

The patient populations and the time points of assessment are shown in Figure 1. All NDMM patients were treated with novel agents and at least one autologous stem cell

transplantation. While 122 NDMM patients were enrolled in TTs, 46 patients were treated with TT-like regimens. The median time to CR was 0.62 years (range 0.04–2.96 years). Patients' characteristics are presented in Supplementary Table 1. All 33 patients, who achieved CR during salvage therapy, relapsed during/after TTs and had been treated with novel agents and at least one autologous stem cell transplantation. Conventional response was assessed according to the IMWG criteria, and CR was defined as negative immunofixation on the serum and urine, disappearance of any soft tissue plasmacytomas, and <5% plasma cells in BM aspirates<sup>19</sup>. The study was performed under University of Arkansas for Medical Sciences IRB approval (205415), and all patients signed written consent in accordance with the Declaration of Helsinki.

### Functional Imaging

DW-MRI examinations were performed on a 1.5-Tesla Philips Achieva scanner (Phillips, Koninklijke, Netherlands). The protocol included scanning from vertex to toes in 7–9 slabs depending on patient height. Each slab constituted 50 slices 5mm thick, field of view 450 mm, matrix 112×79, repetition time 7500 ms, time to echo 69.9 ms, number of acquisitions 2, with a “Q” body coil,  $b=0$  and 800 s/mm<sup>2</sup>. A coronal whole body T1 turbo-spin echo image was obtained as a localizer. Total imaging time for the study was approximately 26 min of which 5 min were spent for T1 image acquisition. Imaging according to this protocol has been performed at our institute since 2011. The following qualitative approach was used to detect residual FLs on DW-MRI: initially, two experienced radiologists visually assessed the signal intensity distribution on the inverted grey-scale b800 whole body DW-MRI maximum intensity projection (MIP) image in consensus read. Well delineated focal intraosseous intensities compared to the surrounding background (low signal on inverted image) that measured at least 0.5 cm in the largest diameter were subsequently visually inspected on the corresponding inverted grey-scale MIP exponential apparent diffusion coefficient (EADC) map and respective source images. As recently advocated, coronal 6mm T1-weighted whole body survey images were used to confirm intraosseous location<sup>20</sup>. Lesions were classified as restricting (low signal compared to the surrounding background) and non-restricting. An example for a restricting lesion is shown in Figure 1.

PET was performed on a Biograph, Reveal, or Discovery scanner as previously described<sup>15</sup>. For PET, a FL was defined as a circumscribed focus with increased FDG uptake compared to its surroundings<sup>15, 21, 22</sup>. Two independent radiologists and one nuclear medicine physician reported the scans and were blinded to the clinical information. Due to potential reactive changes in the bone marrow during treatment, we did not consider increased diffuse background signals as indicators of residual disease at the onset of CR.

### Multiparamter flow-cytometry with 8 colors

For the flow-cytometric assessment of MRD, BM samples were immunophenotyped on a FACSCanto II flow cytometer using an 8-color technique [CD138 (V-500), CD38 (FITC), CD19 (PE-Cy7), CD45 (V-450), CD27 (PerCPy5.5), CD81 (APC-H-7), CD56 (APC) and CD20 (PE)]<sup>4</sup>. Acquiring at least  $2.5 \times 10^6$  events, we defined MRD negativity by Flow as less than 20 events indicating phenotypically aberrant clonal plasma cells. The median limit of detection<sup>23</sup> for newly diagnosed and relapsed patients achieving MRD negative CR was

$0.88 \times 10^{-5}$  ( $0.57-1.68 \times 10^{-5}$ ) and  $0.84 \times 10^{-5}$  ( $0.59-1.04 \times 10^{-5}$ ), respectively. MRD assessment has been performed at our institute since 2013.

### **Risk stratification by gene expression profiling and the international staging system (ISS)**

CD138-enrichment of PCs was performed as published<sup>24, 25</sup>. GEP of CD138-enriched PCs using Affymetrix U133 2.0 plus arrays (Affymetrix, Santa Clara, CA, USA), GEP70-based risk designation and determination of the ISS were performed as described<sup>26, 27</sup>.

### **Whole exome sequencing, mutation calling, and copy number profiling**

Whole exome sequencing (WES) and mutation calling was performed as recently described<sup>9, 28</sup>. Briefly, tumor and control DNA were isolated from CD138-positive PCs and peripheral blood leukapheresis products collected after induction therapy, respectively. Exome enrichment was performed using the SureSelect Clinical Research Exome kit (Agilent, CA, USA) with additional baits covering the Ig and *MYC* loci. Paired-end sequencing was performed on a HiSeq2500 (Illumina, CA, USA).

Somatic single nucleotide variants (SNV) were called using MuTect v1.1.7<sup>29</sup> and Strelka v2.8.3<sup>30</sup> with default parameters. The intersection of SNVs identified by both variant callers was filtered using the ffilter.pl script (<https://github.com/ckandoth/variant-filter>) with default parameters. After exclusion of variants located in immunoglobulin loci, we determined read counts for all mutations per patient using the Rsamtools R package v1.24.0 and the following inclusion criteria: unique reads, coverage exceeding 20x in all samples of the patient, a mapping quality of at least 20 and base quality of at least 20 at the site of the variant. To avoid an overestimation of heterogeneity, we increased the threshold for heterogeneous mutations to a cancer clonal fraction of 0.2 (corresponding to a clonal proportion of 20%). Ig and *MYC* translocations were called using Manta v1.1.1<sup>31</sup> with a variant quality score threshold of 30. Mutations and translocations were manually inspected in IGV<sup>32</sup>, and annotated using ANNOVAR<sup>33</sup> and SNPeff v4.2<sup>34</sup>. Copy number aberrations (CNAs) were called using Sequenza<sup>35</sup>. The accuracy of CNA calls was verified by manual inspection of LogR and B allele frequency values.

### **Statistics**

The Kaplan-Meier method was used for survival analyses. PFS time was measured from onset of CR to relapse or death from any cause or censored at the date of last contact. Overall survival was not considered for this analysis due to limited follow-up and heterogeneous salvage therapies. For the same reason, we did not perform survival analyses for relapsed patients. Residual FLs and Flow results were tested in a multivariate Cox regression model including established risk factors. Fisher's exact test was used to compare the distribution of discrete variables across groups. Main analyses were undertaken using R (v3.3.1) software.

## RESULTS

### Residual FL detected by DW-MRI are associated with early disease progression

Recent studies have shown that residual PET-positive FLs in NDMM patients achieving CR are associated with adverse outcomes<sup>36, 37</sup>. Here, we addressed whether this holds true for residual FLs detected by DW-MRI. The functional imaging technique DW-MRI is based on measuring water diffusion *in vivo*. Diffusion is impeded in tissues with high cellularity such as tumor tissues<sup>38</sup>. We used 2 types of inverted grey-scale images to detect residual FLs: 1) a DWI map, and 2) an EADC map. Usually, hyper-intensities on DWI maps have to be confirmed by the corresponding EADC map, to exclude false-positive signals due to “T2 shine-through”; a phenomenon based on long T2 decay time in some tissues<sup>39</sup>. On DWI maps, we detected focal intensities in 82 of 168 first-line patients (49%) at the onset of CR. However, for more than half of these patients these regions did not exhibit abnormal impeded diffusion on corresponding EADC maps, indicating that the signals on DWI maps were not due to increased tumor cellularity. Indeed, focal intensities which were not detectable on the EADC map did not impact PFS after onset of CR (Fig. 2A). In contrast, restricting FLs defined a group of 35 patients (21%) with short PFS (median: 3.38 years, Hazard ratio: 2.01,  $p=0.01$ , Fig. 2A). Therefore, we only considered focal intensities with confirmed abnormal impeded diffusion for the following analyses.

Compared to PET, DW-MRI detected more patients with residual lesions (21% vs. 6%). Specifically, only 6 of the DW-MRI-positive patients also presented with FLs in PET. Yet, there were 5 patients with FLs in PET only, suggesting that the two techniques are complementary. In total, 40 of 168 (24%) patients had residual FLs at the onset of CR. The PFS of patients with FLs, which were only detectable using DW-MRI, was not significantly different from patients with PET-positive FLs (median PFS 3.4 years vs. 3.0 years,  $p=0.2$ , Fig. 2B).

Together, residual FLs at the onset of CR are associated with adverse PFS, with DW-MRI being the more sensitive method for detection of these lesions.

### Established risk classifiers improve the prognostic value of residual lesions

To address the question whether disease features, which were determined at baseline, impact the prognostic value of residual FLs, we stratified the patients by the gene expression based GEP70 risk classifier or the ISS. In order to increase the power of this analysis, we defined an “*Imaging+*” group, which included the patients with PET+ and/or DW-MRI+ FLs.

GEP70 high risk (HR) patients with residual FLs had a dismal outcome with a median PFS of 0.8 years after onset of CR (Fig. 2C). A similar outcome was seen in ISS III patients with residual FLs on imaging (median PFS of 0.9 years, Fig. 2D). In contrast, 30 patients with ISS I and no FLs had an excellent PFS, with only 2 events occurring in the study period, highlighting the power of combined functional imaging and risk status assessment. This is supported by results of a Cox model including the imaging status, ISS and GEP70, which demonstrate that all 3 parameters independently impacted outcome (Supplementary Table 2).

### Functional imaging complements MRD diagnostics using Flow

For 83 NDMM patients MRD, PET and DW-MRI data were available at the onset of CR. We detected MRD in 49 patients (59%), and MRD-positivity was associated with unfavorable outcome ( $p=0.03$ , Fig. 3). Combining MRD and functional imaging, residual disease was detectable in 53 patients (64%). The best outcome was seen for the 30 MRD-/Imaging- patients (3 events with a median follow-up of 3.6 years), the worst outcome was seen for the 10 MRD+/Imaging+ patients (median PFS: 2.1 years). Of note, 6 and 2 of the MRD+/Imaging+ patients were classified as GEP70 low risk and ISS I, respectively. Extending the multivariate Cox model for PFS by including MRD results in addition to Imaging, GEP70 and ISS, *Imaging+* remained an independent prognostic factor ( $p<0.001$ ). Only 4 of 83 patients were MRD-/Imaging+, associated with residual FLs being a rare event in MRD-negative NDMM patients at a sensitivity of  $1\times 10^{-5}$ . Yet, in this small group two patients experienced early relapse in the studied period, consistent with residual FLs in MRD-negative patients being prognostically relevant.

Together, the combination of MRD and functional imaging improves detection of residual disease leading to early relapse.

### Focal residual disease challenges MRD diagnostics in late-stage patients

Recently, we identified macrofocal disease as a frequent relapse pattern<sup>40</sup>. Macrofocal disease, in which malignant cells are predominantly located in FLs and barely present in random iliac crest biopsies, was seen in 25% of patients relapsing from TTs, showing that a heterogeneous disease distribution is a common feature of late-stage patients. To test whether this increased heterogeneity could confound MRD, we investigated a set of heavily pretreated patients relapsing from TTs (median lines of therapy: 3). These patients achieved CR during different types of salvage therapy which are listed in Supplementary Table 3. Combining MRD and Imaging data for 33 patients in CR during salvage therapy, we detected residual disease in 25 patients (76%). Of note, the proportion of patients, who were MRD-negative but had residual FLs was significantly higher compared to NDMM (8/16 vs 4/34 patients,  $p=0.01$ ). On the other hand, 10 patients (30%) were MRD+ but Imaging-, supporting the idea that a combined MRD/Imaging approach can improve detection of residual disease and should be used in late-stage patients.

### Multiple clones can survive extensive treatment

Our data supports the concept of differential response occurring spatially in MM with focal residual disease being the source of relapse. To further investigate this concept, we addressed the phylogenetic relationship between baseline and relapse clones for one patient who was comprehensively characterized longitudinally by multi-region sequencing, MRD, and functional imaging. This patient presented with more than 100 FLs and extensive spatial heterogeneity at baseline, defined by 4 different major clones at 4 BM sites (Fig 4). At relapse from MRD-negative CR, we obtained a sample from a large FL with paramedullary component in the left thoracic region. Using WES, we identified the baseline “Clone 1” as the ancestor of the clone dominating at this site. The clone, which we termed “Clone 1.2”, shared a “KRAS<sup>Gly12Val</sup>” with “Clone 1”, but had 65 new missense mutations (4 clonal, Supplementary Table 4), which were not detectable at the site of “Clone 1” at baseline

(median sequencing depth of 260x, Supplementary Table 4). This observation is consistent with the selection of “Clone 1.2” as a consequence of treatment. A second sample obtained from the right iliac crest showed an additional resistant clone at relapse, which was not detectable at baseline despite multi-region sequencing (Supplemental table 4). This clone, which we called “Clone 5”, had 15 private missense mutations (6 clonal) including another site-unique clonal *KRAS* mutation, indicating parallel evolution of *KRAS*-mutated clones, and highlighting the survival of multiple, spatially separated, clones even in a patient who was MRD-negative.

The MRD and Imaging follow-up data of this patient also recapitulated the aforementioned impact of spatial heterogeneity on MRD diagnostics. The patient achieved an MRD-negative CR but still presented with residual FLs (Fig 4). After relapse, the MRD level remained low or even negative for several subsequent tests. Imaging showed dynamic changes in the left thoracic region, with rapid resolution and recurrence of masses, while other FLs remained stable, such as the one in the left upper femur. This observation illustrated the presence of pronounced focally heterogeneous disease in late-stage patients.

Together, the case highlights the persistence and progression of multiple spatially separated clones in the BM irrespective of being in an MRD-negative CR.

## DISCUSSION

MRD-negativity is a strong predictor for outcome. It is anticipated that it will be incorporated as an endpoint and for response adapted therapy<sup>41, 42</sup>. Recently, we demonstrated that MM is characterized by spatial tumor heterogeneity with limited exchange between clones at spatially distinct sites<sup>9</sup>. This type of heterogeneity is a possible explanation for relapse in CR patients with an MRD-negative test based on a biopsy from the iliac crest.

Whole body functional imaging has been proposed as a complementary approach for monitoring of residual disease, because it can potentially detect focal residual disease at alternate sites<sup>43, 44</sup>. In this study we show for the first time that DW-MRI allows for detection of prognostically relevant residual FLs, complementing the already demonstrated value of this imaging technique as a tool for detection of FLs, disease staging, and prognosis at baseline, as well as response assessment<sup>14, 15, 45, 46</sup>. Yet, only restricting FLs were associated with short PFS, suggesting that non-restricting focal hyper-intensities in CR patients do not necessarily contain MM cells. However, we would like to emphasize that the biological basis underlying restriction within FLs in MM is poorly understood<sup>20, 47, 48</sup>. In this context, we recently showed that large FLs with paramedullary components can lack restriction in treatment naïve NDMM<sup>46</sup>. Furthermore, some residual FLs in this study were FDG-avid but lacked restriction, highlighting the need for a comprehensive analysis of FL characteristics. As we move forward to integrate DW-MRI into the diagnostics of MM, there is also a need to standardize the interpretation and reporting of this technique. Furthermore, since DW-MRI and PET provide non-redundant information, their concurrent acquisition using a PET-MRI scanner deserves at least further evaluation. These attempts are on-going.

Overall, DW-MRI detected more residual FLs than PET. Yet, we show that a dual approach using DW-MRI and PET is the more sensitive approach, allowing for detection of focal residual disease in 24% of NDMM patients achieving CR. The sensitivity for residual disease detection could be further increased to 64%, when MRD was included as an additional test. Confirming prior data, the best outcome in our set was seen for patients becoming negative by both, functional imaging and MRD<sup>13</sup>. This double-negativity could serve as a new endpoint for future trials. In contrast, patients with double-positivity had such a dismal outcome that a switch to alternative treatments would be beneficial.

We barely found residual FLs in MRD-negative NDMM patients in CR. In contrast, in late stage patients with MRD-negative CR, imaging was frequently positive, highlighting the pronounced heterogeneous distribution of disease. An important question for future studies will be, whether increasing the sensitivity by using next generation MRD flow cytometry (NGF) can improve detection of MRD in these patients. However, for NGF acquisition of up to 10 million cells is required which may limit the feasibility of this approach<sup>23, 49, 50</sup>. Non-optimal sensitivity is also an issue of current functional imaging techniques. It remains to be seen whether PET with alternative tracers, such as <sup>11</sup>C Methionine, or immunoPET approaches improve detection of focal residual disease<sup>44, 51, 52</sup>.

The patchy type of disease distribution in MM is perplexing and poorly understood. Our multi-region sequencing data illustrates that multiple clones with distinct biological features can persist at different sites in the BM cavity despite intense treatment and achievement of deep responses. It also indicates selection of rare subclones, which are not detectable at diagnosis despite high-depth sequencing of the identical site. Dissemination of resistant clones within the BM seemed to be substantially limited in the studied patient, impacting MRD diagnostics. Future studies are warranted to show whether this patient is representative of cases relapsing from CR. However, frequent relapses from MRD-negative CR support our idea that residual disease can persist at sites other than the pelvis.

Our observation of multiple resisting clones challenges the single iceberg as a metaphor for disease burden in MM, where the tip of the iceberg is immediately visible but its foundations go deep beneath the surface. Using this picture, MRD diagnostics aim “to reveal the previously unseen disease”<sup>53</sup>. Adjusting the metaphor by our observation, we think that multiple icebergs better represent the situation in a MM patient (Fig. 5). The tip of these icebergs (=FLs) may be visualized in medical imaging, depending on the size of disease foci and the imaging method used. The foundation of the iceberg can be assessed using highly sensitive approaches such as flow cytometry but false negative results will result during deep responses, if the iliac crest region is iceberg (disease)-free.

In conclusion, DW-MRI is a promising tool for detection of residual disease and complements PET. The combination of MRD diagnostics and functional imaging improves prediction of outcome, with double-negativity and double positivity defining groups with excellent and dismal PFS, respectively. Prospective trials using this information to tailor therapy are warranted. From a biological perspective, this study highlights the confounding effects of spatial heterogeneity and limited dissemination of clones within the BM on MRD



diagnostics. This may especially be true for patients achieving deep responses during salvage therapies.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## ACKNOWLEDGEMENTS

We thank the patients and staff of the Myeloma Institute, UAMS. We also thank the Department of Radiology, UAMS. This work was supported by P01 CA 55819 from the National Cancer Institute. Leo Rasche was supported by the Deutsche Forschungsgemeinschaft (DFG). Niels Weinhold was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number P20GM125503.

## REFERENCES

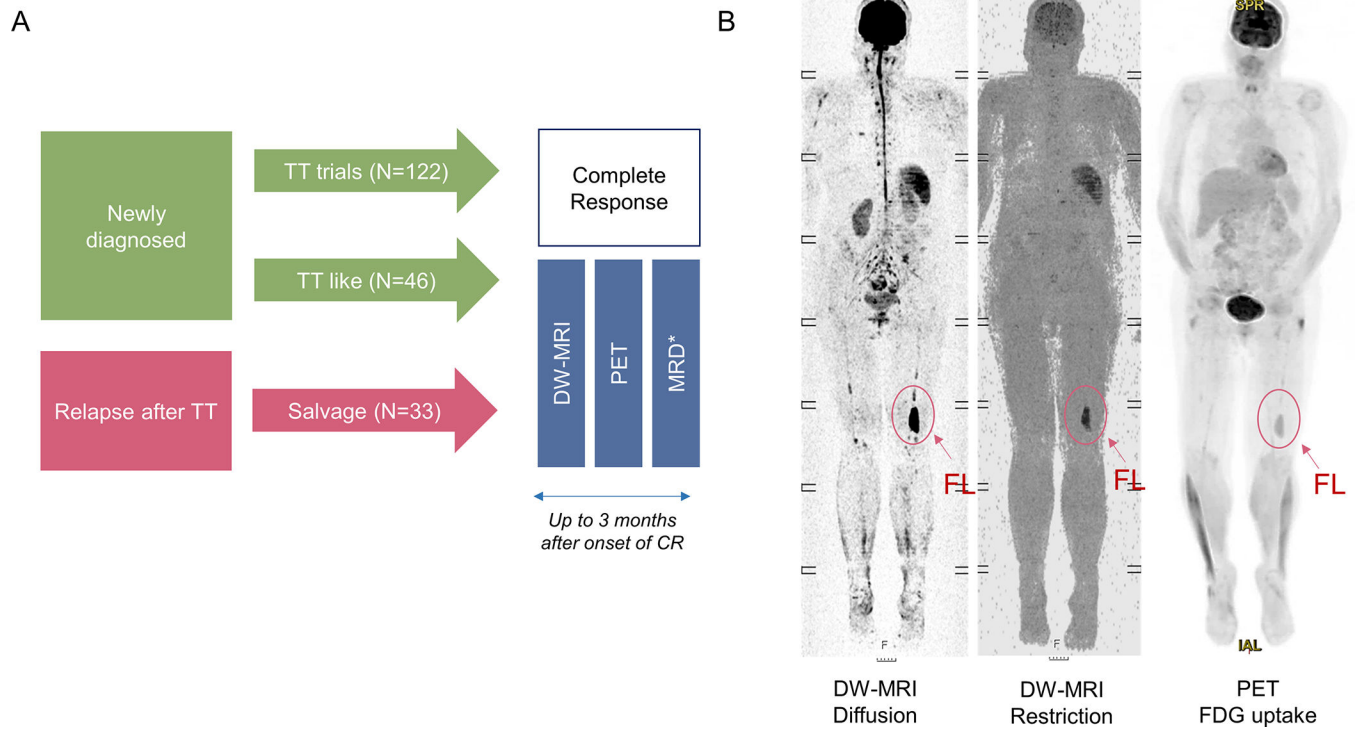
1. Avet-Loiseau H, Lauwers-Cances V, Corre J, Moreau P, Attal M, Munshi N. Minimal Residual Disease in Multiple Myeloma: Final Analysis of the IFM2009 Trial, Abstract 435. ASH 59th Annual Meeting & Exposition, December 2017 2017.
2. Anderson KC. Should minimal residual disease negativity be the end point of myeloma therapy? *Blood Adv* 2017 3 14; 1(8): 517–521. [PubMed: 29296970]
3. Munshi NC, Avet-Loiseau H, Rawstron AC, Owen RG, Child JA, Thakurta A, et al. Association of Minimal Residual Disease With Superior Survival Outcomes in Patients With Multiple Myeloma A Meta-analysis. *JAMA oncology* 2017 1; 3(1): 28–35. [PubMed: 27632282]
4. Schinke C, Hoering A, Wang H, Carlton V, Thanandarajan S, Deshpande S, et al. The prognostic value of the depth of response in multiple myeloma depends on the time of assessment, risk status and molecular subtype. *Haematologica* 2017; 102(8): e313–e316. [PubMed: 28522572]
5. Lahuerta JJ, Paiva B, Vidriales MB, Cordon L, Cedena MT, Puig N, et al. Depth of Response in Multiple Myeloma: A Pooled Analysis of Three PETHEMA/GEM Clinical Trials. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2017 9 1; 35(25): 2900–2910. [PubMed: 28498784]
6. 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 transplantation* 2016 12; 51(12): 1565–1568. [PubMed: 27595280]
7. Kis O, Kaedbey R, Chow S, Danesh A, Dowar M, Li T, et al. Circulating tumour DNA sequence analysis as an alternative to multiple myeloma bone marrow aspirates. *Nat Commun* 2017 5 11; 8: 15086. [PubMed: 28492226]
8. Lohr JG, Kim S, Gould J, Knoechel B, Drier Y, Cotton MJ, et al. Genetic interrogation of circulating multiple myeloma cells at single-cell resolution. *Sci Transl Med* 2016 11 02; 8(363): 363ra147.
9. Rasche L, Chavan SS, Stephens OW, Patel PH, Tytarenko R, Ashby C, et al. Spatial genomic heterogeneity in multiple myeloma revealed by multi-region sequencing. *Nat Commun* 2017 8 16; 8(1): 268. [PubMed: 28814763]
10. Raab MS, Lehnert N, Xu J, Ho AD, Schirmacher P, Goldschmidt H, et al. Spatially divergent clonal evolution in multiple myeloma: overcoming resistance to BRAF inhibition. *Blood* 2016 4 28; 127(17): 2155–2157. [PubMed: 26884375]
11. Lopez-Anglada L, Gutierrez NC, Garcia JL, Mateos MV, Flores T, San Miguel JF. P53 deletion may drive the clinical evolution and treatment response in multiple myeloma. *Eur J Haematol* 2010 4; 84(4): 359–361. [PubMed: 20002730]
12. Mithraprabhu S, Khong T, Ramachandran M, Chow A, Klarica D, Mai L, et al. Circulating tumour DNA analysis demonstrates spatial mutational heterogeneity that coincides with disease relapse in myeloma. *Leukemia* 2017 1 03.
13. Moreau P, Attal M, Caillot D, Macro M, Karlin L, Garderet L, et al. Prospective Evaluation of Magnetic Resonance Imaging and [(18)F]Fluorodeoxyglucose Positron Emission Tomography-Computed Tomography at Diagnosis and Before Maintenance Therapy in Symptomatic Patients

With Multiple Myeloma Included in the IFM/DFCI 2009 Trial: Results of the IMAJEM Study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2017 9 1; 35(25): 2911–2918. [PubMed: 28686535]

14. Pawlyn C, Fowkes L, Otero S, Jones JR, Boyd KD, Davies FE, et al. Whole-body diffusion-weighted MRI: a new gold standard for assessing disease burden in patients with multiple myeloma? *Leukemia* 2016 6; 30(6): 1446–1448. [PubMed: 26648535]
15. Rasche L, Angtuaco E, McDonald JE, Buros A, Stein C, Pawlyn C, et al. Low expression of hexokinase-2 is associated with false-negative FDG-positron emission tomography in multiple myeloma. *Blood* 2017 7 6; 130(1): 30–34. [PubMed: 28432222]
16. Weinhold N, Heuck CJ, Rosenthal A, Thanendrarajan S, Stein CK, Van Rhee F, et al. Clinical value of molecular subtyping multiple myeloma using gene expression profiling. *Leukemia* 2016 2; 30(2): 423–430. [PubMed: 26526987]
17. Dimopoulos MA, Oriol A, Nahi H, San-Miguel J, Bahlis NJ, Usmani SZ, et al. Daratumumab, Lenalidomide, and Dexamethasone for Multiple Myeloma. *The New England journal of medicine* 2016 10 6; 375(14): 1319–1331. [PubMed: 27705267]
18. Chari A, Suvannasankha A, Fay JW, Arnulf B, Kaufman JL, Ifthikharuddin JJ, et al. Daratumumab plus pomalidomide and dexamethasone in relapsed and/or refractory multiple myeloma. *Blood* 2017 8 24; 130(8): 974–981. [PubMed: 28637662]
19. Kumar S, Paiva B, Anderson KC, Durie B, Landgren O, Moreau P, et al. International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma. *Lancet Oncol* 2016 8; 17(8): e328–346. [PubMed: 27511158]
20. Lecouvet FE. Whole-Body MR Imaging: Musculoskeletal Applications. *Radiology* 2016 5; 279(2): 345–365. [PubMed: 27089188]
21. Bartel TB, Haessler J, Brown TL, Shaughnessy JD, van Rhee F Jr., Anaissie E, et al. F18-fluorodeoxyglucose positron emission tomography in the context of other imaging techniques and prognostic factors in multiple myeloma. *Blood* 2009 9 3; 114(10): 2068–2076. [PubMed: 19443657]
22. Waheed S, Mitchell A, Usmani S, Epstein J, Yaccoby S, Nair B, et al. Standard and novel imaging methods for multiple myeloma: correlates with prognostic laboratory variables including gene expression profiling data. *Haematologica* 2013 1; 98(1): 71–78. [PubMed: 22733020]
23. Burgos L, Paiva B EuroFlow-Based Next-Generation Flow Cytometry for Detection of Circulating Tumor Cells and Minimal Residual Disease in Multiple Myeloma In: Heuck C, Weinhold N (ed). *Multiple Myeloma Methods and Protocols*. Springer2018.
24. Zhan F, Huang Y, Colla S, Stewart JP, Hanamura I, Gupta S, et al. The molecular classification of multiple myeloma. *Blood* 2006 9 15; 108(6): 2020–2028. [PubMed: 16728703]
25. Shaughnessy J, Tian E, Sawyer J, Bumm K, Landes R, Badros A, et al. High incidence of chromosome 13 deletion in multiple myeloma detected by multiprobe interphase FISH. *Blood* 2000 8 15; 96(4): 1505–1511. [PubMed: 10942398]
26. Greipp PR, San Miguel J, Durie BG, Crowley JJ, Barlogie B, Blade J, et al. International staging system for multiple myeloma. *J Clin Oncol* 2005 5 20; 23(15): 3412–3420. [PubMed: 15809451]
27. Shaughnessy JD, Zhan F Jr., Burington BE, Huang Y, Colla S, Hanamura I, et al. A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. *Blood* 2007 3 15; 109(6): 2276–2284. [PubMed: 17105813]
28. Weinhold N, Ashby C, Rasche L, Chavan SS, Stein C, Stephens OW, et al. Clonal selection and double-hit events involving tumor suppressor genes underlie relapse in myeloma. *Blood* 2016; 128(13): 1735–1744. [PubMed: 27516441]
29. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* 2013 3; 31(3): 213–219. [PubMed: 23396013]
30. Saunders CT, Wong WS, Swamy S, Becq J, Murray LJ, Cheetham RK. Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* 2012 7 15; 28(14): 1811–1817. [PubMed: 22581179]

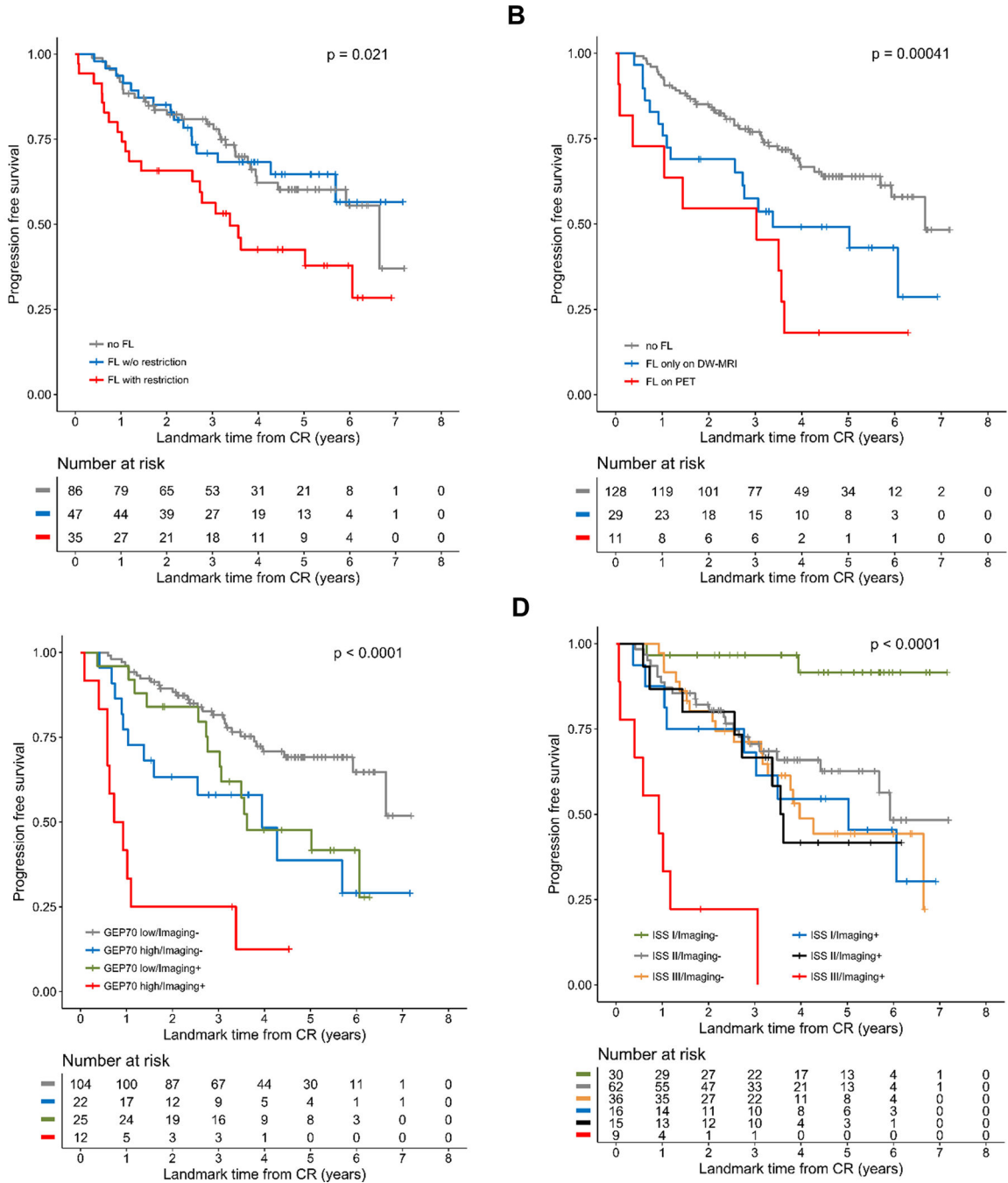
31. Chen X, Schulz-Trieglaff O, Shaw R, Barnes B, Schlesinger F, Kallberg M, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. *Bioinformatics* 2016 4 15; 32(8): 1220–1222. [PubMed: 26647377]
32. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol* 2011 1; 29(1): 24–26. [PubMed: 21221095]
33. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010 9; 38(16): e164. [PubMed: 20601685]
34. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* 2012 Apr-Jun; 6(2): 80–92. [PubMed: 22728672]
35. Favero F, Joshi T, Marquard AM, Birkbak NJ, Krzystanek M, Li Q, et al. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann Oncol* 2015 1; 26(1): 64–70. [PubMed: 25319062]
36. Zamagni E, Nanni C, Mancuso K, Tacchetti P, Pezzi A, Pantani L, et al. PET/CT Improves the Definition of Complete Response and Allows to Detect Otherwise Unidentifiable Skeletal Progression in Multiple Myeloma. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2015 10 01; 21(19): 4384–4390. [PubMed: 26078390]
37. Davies FE, Rosenthal A, Rasche L, Petty NM, McDonald JE, Ntambi JA, et al. Treatment to suppression of focal lesions on positron emission tomography-computed tomography is a therapeutic goal in newly diagnosed multiple myeloma. *Haematologica* 2018 3 22.
38. Koh DM, Collins DJ. Diffusion-weighted MRI in the body: applications and challenges in oncology. *AJR Am J Roentgenol* 2007 6; 188(6): 1622–1635. [PubMed: 17515386]
39. Provenzale JM, Engelter ST, Petrella JR, Smith JS, MacFall JR. Use of MR exponential diffusion-weighted images to eradicate T2 “shine-through” effect. *AJR Am J Roentgenol* 1999 2; 172(2): 537–539. [PubMed: 9930819]
40. Rasche L, Buros A, Weinhold N, Stein CK, McDonald JE, Chavan SS, et al. The Clinical Impact of Macrofocal Disease in Multiple Myeloma Differs Between Presentation and Relapse. *Blood* 2016 12 2; 128(22).
41. Landgren O, Lu SX, Hultcrantz M. MRD Testing in Multiple Myeloma: The Main Future Driver for Modern Tailored Treatment. *Semin Hematol* 2018 1; 55(1): 44–50. [PubMed: 29759154]
42. Anderson KC, Auclair D, Kelloff GJ, Sigman CC, Avet-Loiseau H, Farrell AT, et al. The Role of Minimal Residual Disease Testing in Myeloma Treatment Selection and Drug Development: Current Value and Future Applications. *Clinical cancer research : an official journal of the American Association for Cancer Research* 2017 8 1; 23(15): 3980–3993. [PubMed: 28428191]
43. Hillengass J, Merz M, Delorme S. Minimal residual disease in multiple myeloma: use of magnetic resonance imaging. *Semin Hematol* 2018 1; 55(1): 19–21. [PubMed: 29759148]
44. Pandit-Taskar N Functional Imaging Methods for Assessment of Minimal Residual Disease in Multiple Myeloma: Current Status and Novel ImmunoPET Based Methods. *Semin Hematol* 2018 1; 55(1): 22–32. [PubMed: 29759149]
45. Messiou C, Giles S, Collins DJ, West S, Davies FE, Morgan GJ, et al. Assessing response of myeloma bone disease with diffusion-weighted MRI. *Br J Radiol* 2012 12; 85(1020): e1198–1203. [PubMed: 23175485]
46. Rasche L, Angtuaco EJ, Alpe TL, Gershner GH, McDonald JE, Samant RS, et al. The presence of large focal lesions is a strong independent prognostic factor in multiple myeloma. *Blood* 2018 7 5; 132(1): 59–66. [PubMed: 29784643]
47. Padhani AR, van Ree K, Collins DJ, D’Sa S, Makris A. Assessing the relation between bone marrow signal intensity and apparent diffusion coefficient in diffusion-weighted MRI. *AJR Am J Roentgenol* 2013 1; 200(1): 163–170. [PubMed: 23255758]
48. Messiou C, Collins DJ, Morgan VA, Desouza NM. Optimising diffusion weighted MRI for imaging metastatic and myeloma bone disease and assessing reproducibility. *Eur Radiol* 2011 8; 21(8): 1713–1718. [PubMed: 21472473]
49. Roshal M Minimal Residual Disease Detection by Flow Cytometry in Multiple Myeloma: Why and How? *Semin Hematol* 2018 1; 55(1): 4–12. [PubMed: 29759152]

50. Flores-Montero J, Sanoja-Flores L, Paiva B, Puig N, Garcia-Sanchez O, Bottcher S, et al. Next Generation Flow for highly sensitive and standardized detection of minimal residual disease in multiple myeloma. *Leukemia* 2017 10; 31(10): 2094–2103. [PubMed: 28104919]
51. Lapa C, Garcia-Velloso MJ, Luckeath K, Samnick S, Schreder M, Otero PR, et al. (11)C-Methionine-PET in Multiple Myeloma: A Combined Study from Two Different Institutions. *Theranostics* 2017; 7(11): 2956–2964. [PubMed: 28824728]
52. Caserta E, Chea J, Minnix M, Viola D, Vonderfecht S, Yazaki P, et al. Copper 64-labeled daratumumab as a PET/CT imaging tracer for multiple myeloma. *Blood* 2018 2 15; 131(7): 741–745. [PubMed: 29301755]
53. <https://www.adaptivebiotech.com/clonoseq/clonoseq-assay>. ClonoSEQ Revealing Previously Unseen Disease.



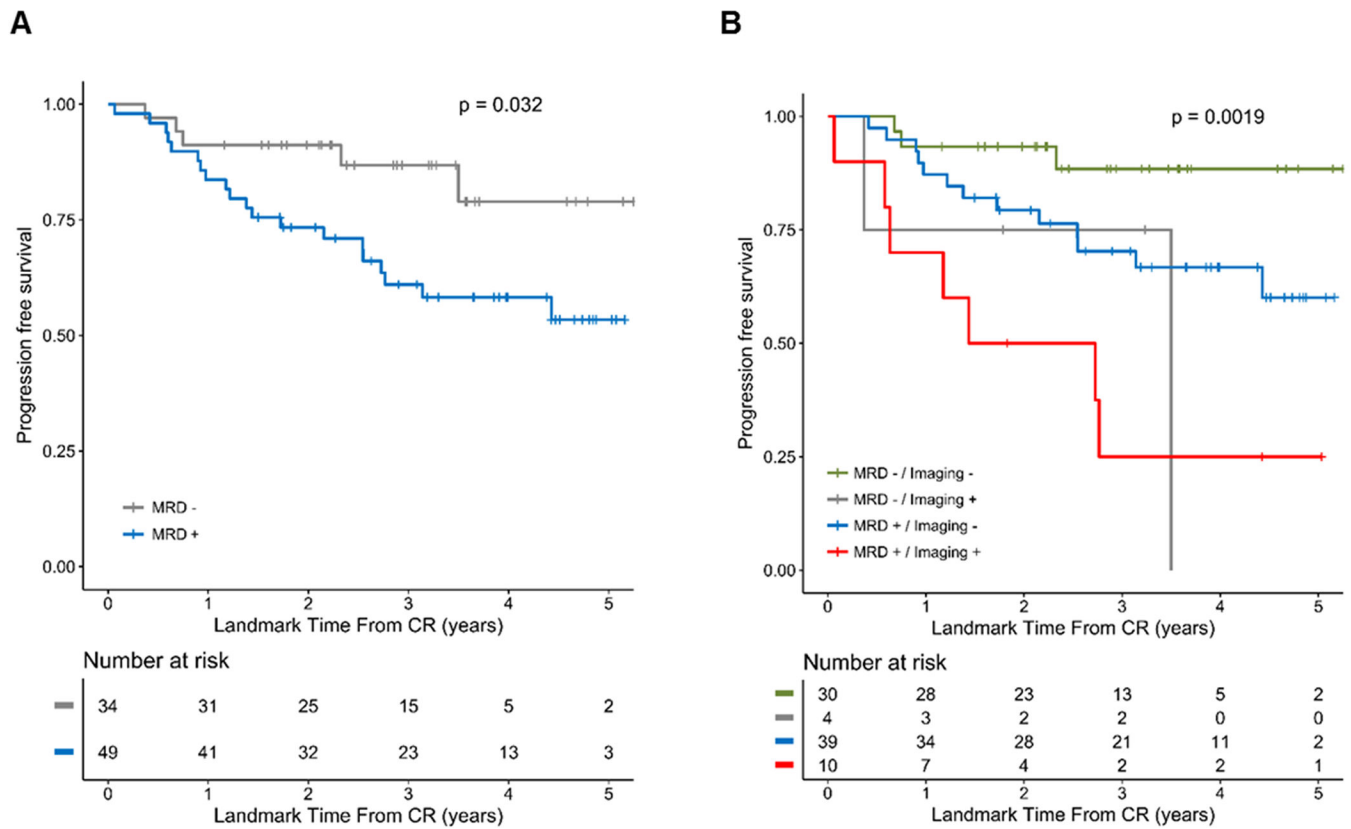
**Figure 1: Patients and study design.**

In (A) the study design is shown. (B) Simultaneous DW-MRI and PET scans are shown for a representative patient. A focal lesion in the left distal femur is indicated by arrows. TT, total therapy; DW-MRI, Diffusion-weighted magnetic resonance imaging; PET, positron-emission-tomography; Flow, 8-color flow cytometry with a limit of detection of  $\sim 1 \times 10^{-5}$ ; FL, focal lesion; \* MRD data was available for 83 first-line and all relapsed patients.

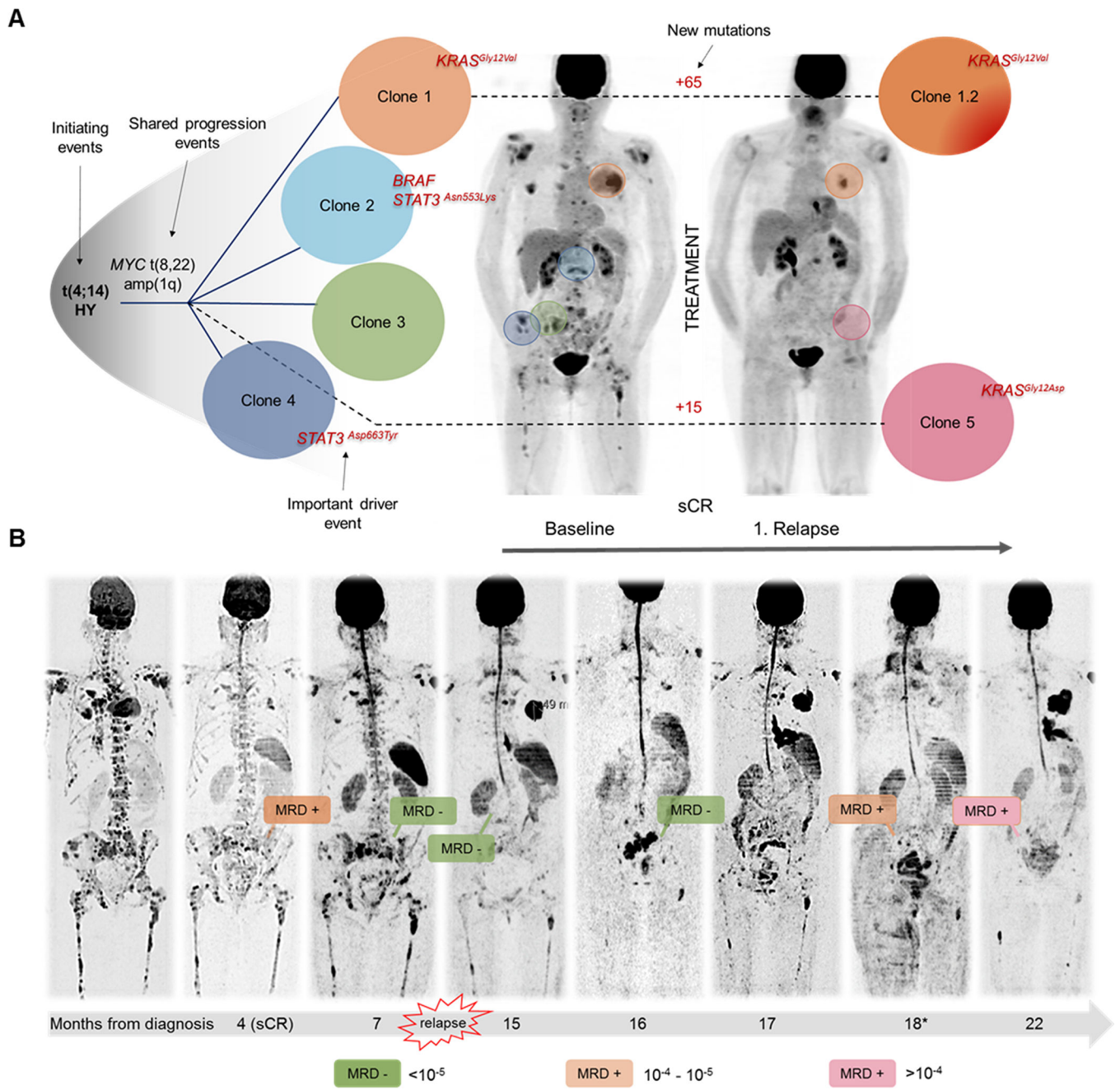


**Figure 2. The impact of residual focal lesions on outcome.**

In (A) the impact of residual FL on DW-MRI is depicted stratified by restriction status of these lesions on EADC maps. (B) Prognostic value of residual FLs detectable using DW-MRI only and PET, respectively. The prognostic impact of residual FLs on outcome stratified by the GEP70 risk classifier or the International staging system is shown in (C) and (D). The log-rank test was used to perform the group comparison.



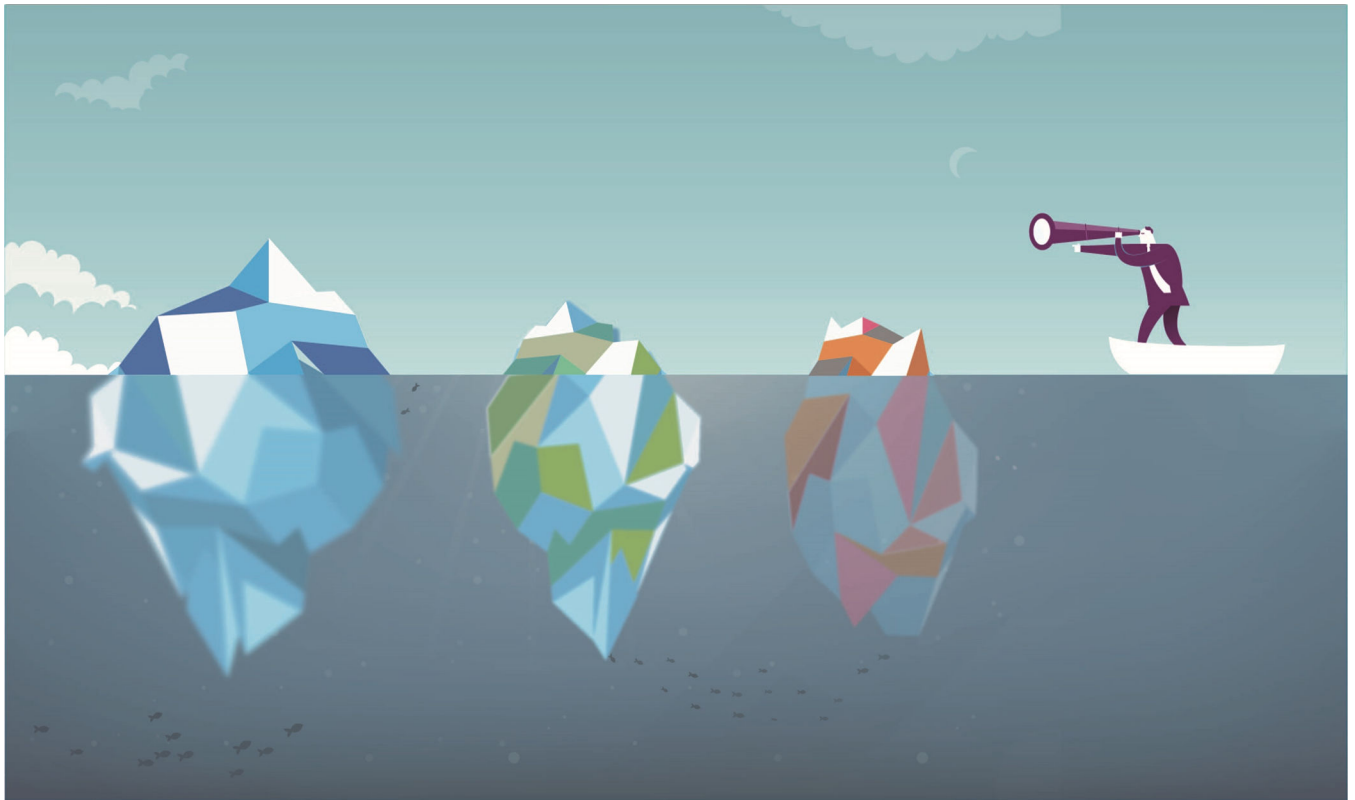
**Figure 3: Prognostic impact of MRD and its combination with functional imaging.** Progression free survival from onset of CR for newly diagnosed MM patients stratified by (A) MRD results, and the combination of MRD and functional imaging (B). The log-rank test was used to calculate p values.



**Figure 4: The impact of spatial heterogeneity on MRD diagnostics.**

In (A) the phylogenetic relationship between clones from different regions at baseline and relapse and the corresponding PET scans are presented for one patient. The location of samples is marked in the PET scans using the color code that was assigned to clones. (B) Follow-up DW-MRI scans of this patient together with MRD results from the iliac crest. The MRD level is shown by the indicated color code.





**Figure 5: The iceberg metaphor for disease burden in multiple myeloma adjusted for spatial heterogeneity.**

The tip of these icebergs, which correspond to FLs, may be visualized in medical imaging, depending on the size of disease foci and the imaging method used. The foundation of the iceberg can be assessed using highly sensitive approaches such as flow cytometry but false negative results will result during deep responses, if the iliac crest region is iceberg (disease)-free.