Am J Cancer Res 2022;12(4):1919-1933
www.ajcr.us /ISSN:2156-6976/ajcr0140875

Original Article

Characterizing the mutational landscape of MM and its precursor MGUS

Akanksha Farswan[1], Anubha Gupta[1], Lingaraja Jena[2], Vivek Ruhela[3], Gurvinder Kaur[2], Ritu Gupta[2]

[1]SBILab, Department of ECE, Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi), New Delhi 110020, India; [2]Laboratory Oncology Unit, Dr. B.R.A. IRCH, AIIMS, New Delhi 110029, India; [3]Department of Computational Biology, IIIT-Delhi, New Delhi 110020, India

Received November 30, 2021; Accepted March 2, 2022; Epub April 15, 2022; Published April 30, 2022

Abstract: Mutational Signatures and Tumor mutational burden (TMB) have emerged as prognostic biomarkers in cancer genomics. However, the association of TMB with overall survival (OS) is still unknown in newly diagnosed multiple myeloma (NDMM) patients. Further, the change in the mutational spectrum involving both synonymous and non-synonymous mutations as MGUS progresses to MM is unexplored. This study addresses both these aspects via extensive evaluation of the mutations in MGUS and NDMM. WES data of 1018 NDMM patients and 61 MGUS patients collected from three different global regions were analyzed in this study. Single base substitutions, mutational signatures and TMB were inferred from the variants identified in MGUS and MM patients. The cutoff value for TMB was estimated to divide patients into low TMB and high TMB (hypermutators) groups. This study finds a change in the mutational spectrum with a statistically significant increase from MGUS to MM. There was a statistically significant increase in the frequency of all the three categories of variants, non-synonymous (NS), synonymous (SYN), and others (OTH), from MGUS to MM ($P<0.05$). However, there was a statistically significant rise in the TMB values for TMB_NS and TMB_SYN only. We also observed that 3' and 5'UTR mutations were more frequent in MM and might be responsible for driving MGUS to MM via regulatory binding sites. NDMM patients were also examined separately along with their survival outcomes. The frequency of hypermutators was low in MM with poor OS and PFS outcome. We observed a statistically significant rise in the frequency of C>A and C>T substitutions and a statistically significant decline in T>G substitutions in the MM patients with poor outcomes. Additionally, there was a statistically significant increase in the TMB of the patients with poor outcome compared to patients with a superior outcome. A statistically significant association between the APOBEC activity and poor overall survival in MM was discovered. These findings have potential clinical relevance and can assist in designing risk-adapted therapies to inhibit the progression of MGUS to MM and prolong the overall survival in high-risk MM patients.

Keywords: Multiple myeloma, monoclonal gammopathy of undetermined significance, NGS, exome sequencing, tumor mutation burden, progression, mutational landscape

Introduction

Multiple Myeloma (MM) is a malignancy of abnormal plasma cells in the bone marrow where the progression of the disease is driven by numerous factors, including immune surveillance, microenvironment, and therapeutic agents. Monoclonal gammopathy of undetermined significance (MGUS) is a benign precursor state of MM characterized by lack of end-organ damage [1] and less than 10% of plasma cells in the bone marrow. MGUS may progress to asymptomatic or symptomatic multiple myeloma with a rate of nearly 1% per year [2], where MM is characterized by severe clinical problems such as bone fractures, anaemia, renal failure, and hypercalcemia. With the advent of Next Generation Sequencing technology, it has become easier to study the DNA of a patient and unearth the genetic causes of the disease. Multiple studies involving exome and genome data of MM have been performed to understand the genomic abnormalities driving tumor progression in MM. It is well established that the primary events in MM are either hyperdiploidy, i.e., trisomy of chromosomes 3, 5, 7, 9, 11, 15, 19 and/or 21 or non-hyperdiploidy involving translocations affecting the genes

encoding immunoglobulin (Ig) heavy chains (IGH)-mainly t(4;14), t(6;14), t(11;14), t(14;16) and t(14;20) [3]. Primary events are then followed by multiple secondary events promoting tumor progression. However, it has also been observed and validated that the genetic aberrations peculiar to MM are also present during the premalignant state of MGUS, where they do not show any clinical symptoms related to MM [4, 5]. It is, therefore, worthwhile to thoroughly investigate the mutational landscape of the genomic alterations affecting MGUS as well as MM. Though multiple studies have been performed to study the MGUS to MM progression [6-8], the landscape of the mutational patterns of the MGUS and MM largely remains unexplored. The study of the changing mutational spectrum of the MGUS as it advances to MM will provide more insight into the disease biology. Further, it will help identify the clinically relevant vital biomarkers that can assist in controlling the progression of MGUS to MM.

Mutational signatures have emerged as critical biomarkers in cancer genomics, with profound pathogenic, prognostic, and therapeutic implications. Multiple mutational events occur in a tumor, while only a few of these mutations are actual drivers of cancer. However, exploring the entire landscape of coding and non-coding mutations helps reveal the mutational signatures characteristics of the specific cancer types. For example, CG>AT transversion is associated with lung cancer [9], and CG>TA is associated with skin cancer [10, 11]. Various mutational signatures have been discovered based on the 96 possible combinations of the single base substitutions and their trinucleotide contexts. These signatures are linked with the defects of DNA repair mechanisms, ageing, UV exposure, and others, thereby validating the role of the mutational processes in shaping the genomic continuum of each cancer type [12-15]. Further, tumor mutational burden (TMB) has become a prominent biomarker of response to immunotherapy and is being explored for its association with overall survival, particularly in solid tumors. TMB is determined as the number of mutations identified per megabase. It has been observed that cancers with a high TMB load of greater than 10 mut/Mb have a better chance of responding to drugs called immune checkpoint inhibitors (ICIs). The primary function of ICIs is to activate the immune system better to recognize cancer cells [16] and act upon them. As a result, a high tumor mutational burden (TMB) has been increasingly associated with superior overall survival in ICI-treated patients. Multiple studies are now being conducted to discover the cancers with high TMB that respond best to ICIs and, thus, prolong the survival of cancer patients. In addition, the association of TMB with survival in non-ICI-treated patients has also been explored. It has been observed that high TMB was associated with poor prognosis and overall survival in the absence of immunotherapy, as opposed to ICI-treated patients in whom high TMB was associated with prolonged survival [17].

Synonymous mutations, earlier designated as silent mutations, were mostly ignored in cancer genomics due to their inability to alter the amino acid of the resultant protein [18]. However, they have the capability of changing the protein expression and function owing to their impact on RNA stability, RNA folding [19] or splicing [20], translation [21], or co-translational protein folding. Multiple studies have corroborated that natural selection is present in synonymous mutations [21-23], contrary to earlier studies that denied the role of selective pressure in synonymous mutations [24]. Various genome-wide association studies conducted in recent times have also confirmed the association of synonymous SNPs to human disease risk and other complex traits. Therefore, the role of synonymous mutations in the disease biology of MGUS and MM should be examined as it could lead to significant prognostic and clinical implications.

Motivated by the above discussion, an exhaustive investigation of the mutations altered in MGUS and MM was carried out in the present study. We explored the change in the mutational landscape as the disease progressed from the MGUS to MM. We found that the difference in the frequency of the single base substitution is significantly different in MGUS and MM. We have also analyzed the frequency of the different types of variants across MGUS and MM and found that few have changed significantly as the disease progressed from MGUS to MM. Further, we categorized MM patients into low TMB and high TMB (hypermutators) based on their overall survival data. We explored the impact of TMB on the frequency

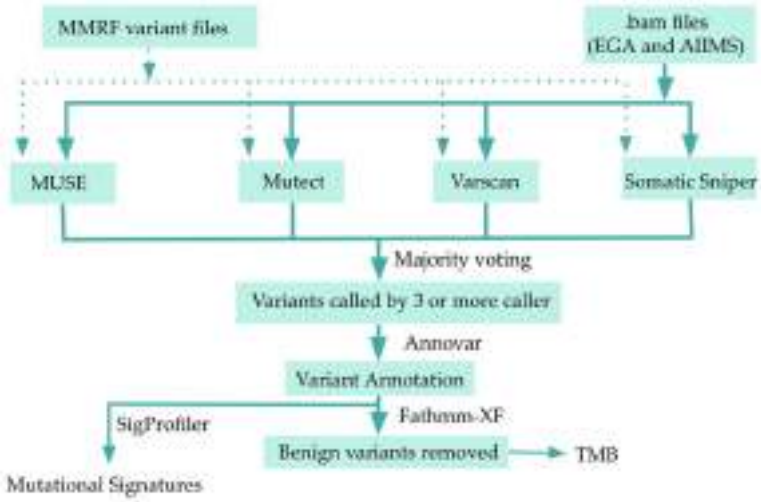


Figure 1. Workflow of the study and data analysis. Four different variant callers were used to identify variants in the MM and MGUS patients. Variants were finalized using the majority voting scheme. Variants were then annotated with Annovar for deducing TMB. Mutational signatures were inferred using Sigprofiler tool.

of single base substitutions and the different variant types across the low and high TMB groups of MM patients. The association of TMB with overall survival is still unknown in newly diagnosed multiple myeloma (NDMM) patients; therefore, we have correlated TMB with survival data and found that high TMB is linked with poor overall survival in NDMM patients.

Methods and materials

Datasets used in the study

The present study is based on the data of 1018 NDMM patients and 61 MGUS patients. Variant files generated from the exome data of 936 NDMM patients out of the total 1018 patients were obtained from the GDC portal via dbGaP authorized access (phs000748; phs000348). This data is a part of the MMRF CoMMpass study. Exome data of the remaining 82 NDMM patients were obtained from AIIMS, Delhi. In addition, exome data of 33 MGUS patients out of 61 patients was obtained from EGA (EGAD00001001901), and exome data of the remaining 28 patients was obtained from AIIMS, Delhi. Four variant callers, namely, MuSE [25], Mutect2 [26], VarScan2 [27], and Somatic-Sniper [28], was used for finding variants in patients from the MMRF CoMMpass study. Therefore, there were four vcf files corresponding to each variant caller for each patient. The workflow of the complete analysis is shown in **Figure 1**.

Analysis of exome data and the variants identified using the exome data

Exome data obtained from AIIMS and EGA was processed with a standard exome sequencing pipeline, and single nucleotide variants (SNVs) were extracted using MuSE, Mutect2, VarScan2, and Somatic-Sniper variant callers. SNVs were annotated using ANNOVAR [29] to gather the genomic information of the mutations, such as their variant type and the deleteriousness of the mutation, etc. FATHMM-XF [30] was used to remove the benign variants. The rest of the filtered variants were categorized into nonsynonymous (NS) variants, synonymous (SYN) variants, and other (OTH) variants. Exonic, nonsynonymous single nucleotide variants (snvs), ncRNA_exonic, stop gain, stop loss, start loss, splicing, frameshift insertion, and frameshift deletion were grouped in nonsynonymous variants. UTR3, synonymous single nucleotide variants (snvs), and UTR5 were grouped in synonymous variants. Non-frameshift insertion, non-frameshift deletion, non-frameshift substitution, intronic, intergenic ncRNA_intronic, upstream, downstream, unknown, and ncRNA_splicing were grouped in other variants.

Assessment of single base substitution, mutational signatures, and TMB

Variants identified by three or more callers were further processed to extract information on single base substitution and identify the mutational signatures present in the data. SigProfilerExtractor [31] was used to discover the single base substitutions and the mutational signatures in the MGUS and MM data. The etiology of the deduced signatures were found via the COSMIC v3.2 mutational signature database [32]. A total of six single base substitutions C>A, C>G, C>T, T>A, T>C, and T>G were identified. Tumor mutational burden (TMB) was calculated using the three different catego-

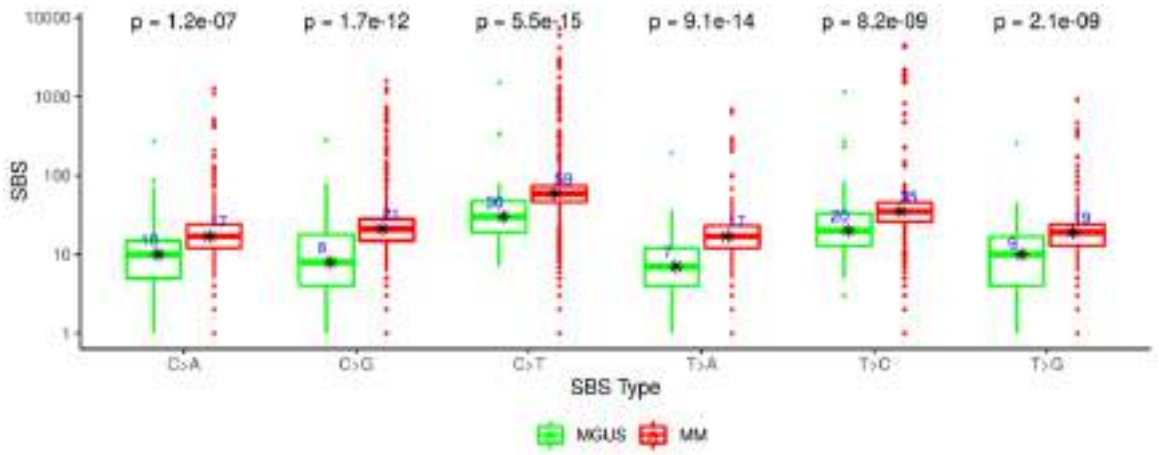


Figure 2. Boxplot shows the difference in the frequency of the single base substitutions between MGUS and MM patients. Wilcoxon rank-sum test was applied to determine if the change is statistically significant or not. For all the substitutions, there is significant variation in the frequency with *p*-values less than 0.05 between the two groups.

ries of variants-nonsynonymous (NS) variants, synonymous (SYN) variants, and other (OTH) variants. TMB was determined as described in [33]. TMB_NS, TMB_SYN, and TMB_OTH were estimated using nonsynonymous (NS) variants, synonymous (SYN) variants, and other (OTH) variants, respectively. Survival data were available for 832 (753+79) patients out of a total of 1018 NDMM patients, which were utilized to obtain the threshold values for TMB_NS, TMB_SYN, and TMB_OTH using the K-adaptive partitioning (KAP) algorithm [34] and Cutoff Finder [35].

Statistical analysis

Wilcoxon rank-sum test was used to determine if the change in the frequencies of the single base substitutions and the different types of variants is statistically significant between the MGUS and MM. Unpaired Wilcoxon rank-sum was applied because the data did not follow the normality distribution and was unpaired.

Results

Frequency of single base substitutions (SBS) increases significantly from MGUS to MM

There was an increase in the median and mean frequency of the single base substitutions from MGUS to MM. The change in the frequency was statistically significant with *p*-values less than 0.05 for all the substitutions according to the Wilcoxon rank-sum test (**Figure 2**). C>T substitution was observed with the highest frequency in MGUS and MM, increasing the median value from 30 to 59. T>C substitution was next, with an increase in the median value from 20 (MGUS) to 35 (MM). T>A was observed with the lowest frequency in MGUS and MM, increasing the median value from 7 to 17.

Calculation of threshold values for the SBS and comparison between the high and low-frequency MM groups

Due to the availability of survival data for 832 MM patients, threshold values for the substitutions were inferred. K-adaptive partitioning (KAP) algorithm and Cutoff Finder were used to deduce the thresholds. **Table 1**, Supplementary Table 3 and Supplementary Figures 5 and 6 show the cut-off values estimated for the different types of substitutions for PFS and OS. Similar cut-offs were deduced by the two tools, i.e. KAP and Cutoff Finder. The higher of the two cut-offs obtained via KAP were selected for C>T, T>C, C>G, C>A, T>G, and T>A substitutions and were 99, 12, 37, 28, 6, and 32, respectively. The patients were then organized into two groups, one with SBS values less than the selected cut-offs and the other one with SBS values greater than the chosen cut-offs. Kaplan Meier (KM) curves corresponding to the two groups revealed that there was a significant difference in the survival patterns of the two groups of patients for the substitutions, C>T, C>G, C>A, and T>A. However, cut-offs obtained for T>C and T>G substitutions yielded a significantly poor outcome for the group with values less than the selected cut-offs.

Table 1. The table shows the cut-offs obtained for the six different types of substitutions via KAP

SBS	Min	Median	Max	PFS cutoff	OS cutoff	Manual cut-off	Frequency (≤, >)	PFS *p*-value	OS *p*-value
C>A	0	17	1251	26	28*	-	712, 120	0.00025	5.13E-06
C>G	0	21	1575	37*	34	-	763, 69	0.026	2.20E-04
C>T	1	59	7315	79	99*	-	750, 82	0.001	4.80E-06
T>A	0	17	684	5	32*	-	784, 48	0.01	0.005
T>C	0	35	4498	12	11	80*	816, 16	0.19	0.01
T>G	0	19	915	6	6	41*	804, 28	0.018	0.007

Two cut-offs were obtained for each SBS, one using PFS and the other using OS. The higher of the two cut-offs and the patients were then organized into two groups, one with SBS values less than the selected cut-offs and the other one with SBS values greater than the selected cut-offs. KM analysis showed that there was a significant difference in the survival patterns of the two groups of patients for the substitutions, C>T, C>G, C>A, and T>A. However, cut-offs obtained for T>C and T>G substitutions did not yield a significant difference in the survival curves. Therefore, cutoffs were manually deduced for the two substitutions where the KM curve has the maximum separability. *Denotes selected cutoffs.

Therefore, cut-offs were manually deduced for T>C and T>G substitutions where the KM curve has the maximum separability and was found to be 80 and 41, respectively. Univariate and multivariate hazard analysis was also done using the selected cut-offs via KAP, as shown in the Supplementary Table 5. The hazard ratio for all the substitutions was greater than 1 in the univariate analysis, demonstrating that an increase in the frequency of these substitutions correlated with an enhanced risk in MM patients. Univariate analysis revealed that C>T substitution had the most significant impact (*p*-value <0.05) on the overall survival (OS) owing to the highest hazard ratio followed by T>C and C>A while T>G had the most significant impact (*p*-value <0.05) on PFS followed by C>T and C>A. However, only C>A was significant in multivariate analysis with *p*-values less than 0.05 (0.04 for PFS and 0.03 for OS).

Comparison of mutational signature profiles between MGUS and MM

A total of 29 and 61 SBS signatures were extracted from the mutation data of MGUS and NDMM patients, respectively. Union of 29 and 61 signatures resulted in 66 unique signatures. Signatures SBS37, SBS49, and SBS55 were found only in MGUS. However, their frequency is low as they were found in a single sample in MGUS (1/61=1.6%). SBS49 and SBS55 signatures are possible sequencing artifacts, and the proposed etiology of signature 37 is unknown according to the COSMIC v3.2 mutational signature database. Further, 37 signatures were discovered only in MM. However, 7 out of 37 were mutated in more than 1% MM samples. They include SBS6, SBS7d, SBS9, SBS17b, SBS19, SBS40, and SBS42. The rest of the 30 signatures were found in less than 1% MM samples and include SBS7c, SBS8, SBS10d, SBS14, SBS20, SBS21, SBS22, SBS23, SBS25, SBS26, SBS27, SBS28, SBS30, SBS32, SBS33, SBS34, SBS35, SBS36, SBS39, SBS41, SBS43, SBS46, SBS47, SBS50, SBS52, SBS53, SBS57, SBS86, SBS88, and SBS89. SBS27, SBS43, SBS46, SBS47, SBS50, SBS52, SBS53, and SBS57 are possible sequencing artifacts, as described previously. Clock-like signatures SBS1 and SBS5 were present in both MGUS and MM. Defective DNA mismatch repair signatures SBS15 and SBS44 were present in both MGUS and MM while SBS6, SBS14, SSB20, SBS21, SBS26 were present only in MM. SBS2 and SBS13 are associated with the activity of the AID/APOBEC family of cytidine deaminases and were found in both MGUS and MM. MM patients with APOBEC signatures were investigated further using survival data. APOBEC signature was present in 27 out of 177 MM patients with poor OS outcome and 52 out of 655 MM patients with superior OS outcome. Fisher's exact test revealed a statistically significant association between the APOBEC activity and poor overall survival in MM (*p*-value =0.0056). However, there was no significant association between APOBEC activity and progression-free survival (*p*-value =0.9). KM curves showed a significant difference (*p*-value =1.8e-4) in the overall survival pattern of MM patients with and without APOBEC activity (Supplementary Figure 2). SBS84 and SBS85 are related to indirect effects of activation-induced cytidine deaminase (AID) induced somatic mutagenesis in

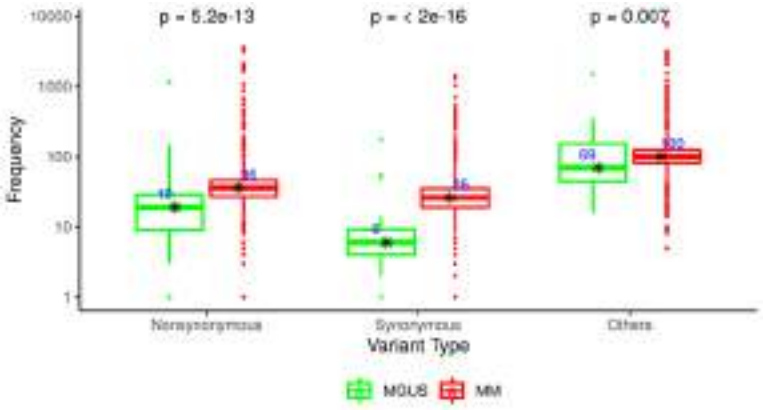


Figure 3. Boxplot showing the variation in the frequency of the three different categories of variants-Nonsynonymous, Synonymous, and Others between MGUS and MM. Wilcoxon rank-sum test was applied to determine if the change is statistically significant or not. There was a statistically significant variation in all the categories of variants with *p*-values less than 0.05.

lymphoid cells and were found in both MGUS and MM.

Frequency of the variants increases significantly from MGUS to MM

According to the Wilcoxon rank-sum test, there was a statistically significant increase in all the three categories of variants from MGUS to MM (**Figure 3**). The median value of nonsynonymous variants increased from 19 to 36 (*p*-value =5.2e-13) as the disease progressed from MGUS to MM. Median value of synonymous variants increased from 6 to 26 (*p*-value <2e-16) while that of other variants increased from 69 to 100 (*p*-value =0.007). Within the nonsynonymous category, there was a statistically significant increase in the nonsynonymous snv (*p*-value =2.9e-13) from 14 to 30 and stop-gain (*p*-value =0.016) variants from 0 to 2 as the disease progressed from MGUS to MM (**Figure 4A**). Within the synonymous category, there was a statistically significant increase in the UTR3 (*p*-value <2e-16) and UTR5 variants (*p*-value =2.7e-7) (**Figure 4B**). Within the other variant category, there was a statistically significant increase in the intronic, intergenic, and downstream variants (Supplementary Figure 1). The median value of UTR3 variants increased from 4 to 21, while that of UTR5 increased from 1 to 4.

Comparison of TMB values between MGUS and MM

Tumor mutational burden (TMB) was calculated using the three different categories of variants-nonsynonymous (NS), synonymous (SYN), and others (OTH). A statistically significant increase was observed for TMB_NS and TMB_SYN with *p*-values less than 0.05 (**Figure 5**). For TMB_OTH, the difference in the KM survival curve was not significant (**Figure 5**).

Calculation of TMB cut-offs and comparison between high and low TMB MM groups

Survival data were available for 832 MM patients. Hence, threshold values of TMB were calculated using the K-adaptive partitioning (KAP) algorithm and Cutoff Finder. Both the tools inferred almost the same cut-offs (**Table 2**, Supplementary Tables 1 and 2; Supplementary Figures 3 and 4). **Table 2** and Supplementary Table 1 reveal the different cut-offs obtained via KAP for progression-free survival (PFS) and overall survival (OS). For TMB_NS, 0.63 and 0.62 are the threshold values obtained via PFS and OS.

Similarly, for TMB_SYN, 0.55 and 0.52 are the threshold values obtained for PFS and OS. The patients were then organized into two groups, one with TMB values less than the selected cut-offs and the other one with TMB values greater than the chosen cut-offs. There was a significant difference (*p*-value <0.05) on the KM survival curves of the patients below 0.63/0.62 and above 0.63/0.62. There is a significant difference (*p*-value <0.05) on the KM survival curves of the patients below 0.55/0.52 and above 0.55/0.52. Univariate and multivariate hazard analysis was also done using the cut-offs via KAP, as shown in the Supplementary Table 4. Hazard ratios for TMB_NS, TMB_SYN and TMB_OTH were greater than 1 in both the univariate and multivariate analysis and indicate the enhanced risk associated with an increase in the mutation burden. Multivariate analysis showed the combined effect of the TMB values on the survival patterns where TMB_NS had the highest impact, followed by TMB_OTH and TMB_SYN, respectively.

MM patients with very high TMB_NS load and very low TMB_NS load were analyzed separately. Cut-off of 35 and 0.1 was deduced using the

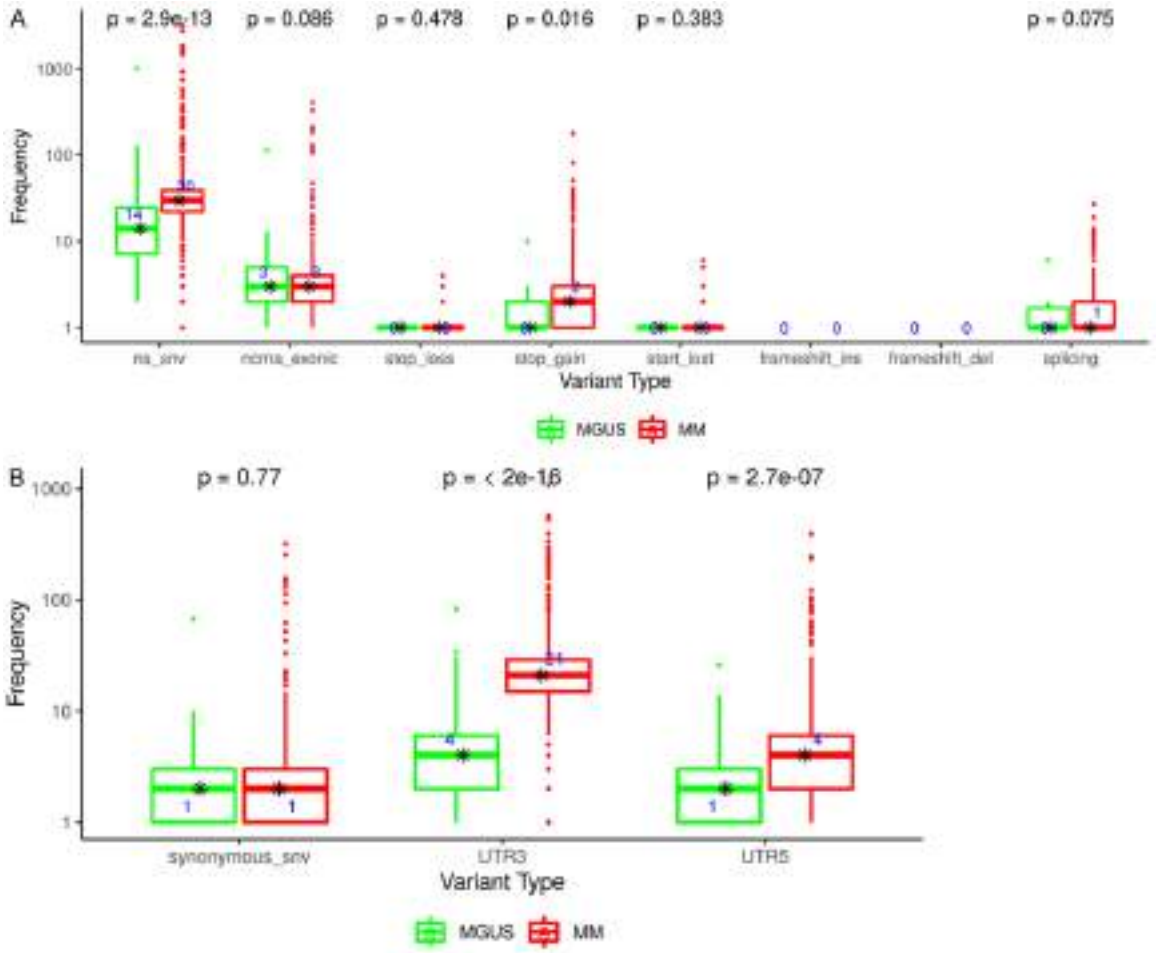


Figure 4. A. Boxplot showing the variation in the frequency of the variants under the nonsynonymous category. There was a statistically significant variation in the frequency of nonsynonymous_snv and stop_gain variants with *p*-values less than 0.05. B. Boxplot showing the variation in the frequency of the variants under the synonymous category. There was a statistically significant variation in the frequency of UTR3 and UTR5 variants with *p*-values less than 0.05. Wilcoxon rank-sum test was applied to determine if the change is statistically significant or not.

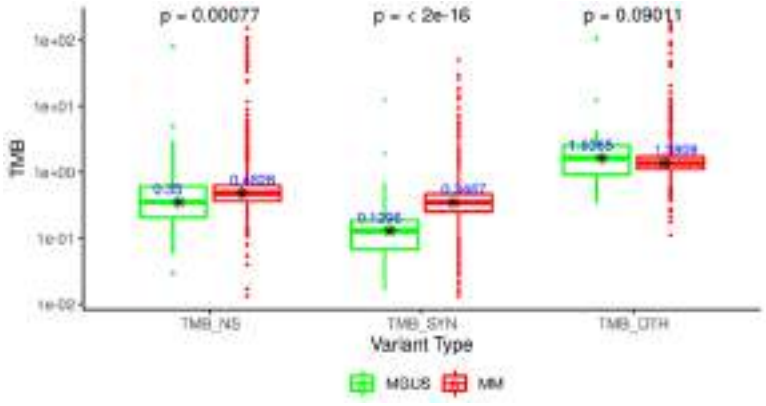


Figure 5. Boxplot reveals that the difference in the low TMB and high TMB groups is statistically significant with *p*-values less than 0.05 for TMB_NS and TMB_SYN. Wilcoxon rank-sum test was applied to determine if the change is statistically significant or not.

maximum separability on the KM survival curves. There were 822 patients with TMB_NS less than 35 and only 10 with TMB_NS greater than 35. There were 6 patients with TMB_NS less than 0.1 and 826 patients with TMB_NS greater than 0.1. A significant difference in the survival patterns of patients with TMB_NS less than 35 and greater than 35 were observed. For PFS, the observed *p*-value was 0.045, and for OS, the observed *p*-value was 0.022 (**Figure 6**). The

Table 2. The table shows the cut-offs obtained for TMB_NS and TMB_SYN via KAP. Two cut-offs were obtained, one using PFS and the other using OS. The two cut-offs are close to each other and KM analysis was done using both the cut-offs

	Min	Median	Max	KAP on PFS			KAP on OS		
				Cut-off (≤, >)	PFS	OS	Cut-off (≤, >)	PFS	OS
TMB_NS	0	0.496	154.2	0.63 (612, 220)	3.19E-07	3.52E-08	0.62 (611, 221)	3.90E-07	2.09E-08
TMB_SYN	0	0.3487	50.84	0.55 (703, 129)	4.12E-05	2.05E-08	0.52 (668, 164)	5.60E-04	3.50E-08

There was a significant difference (*p*-value <0.05) on the KM survival curves of the patients below and above the selected cut-offs.

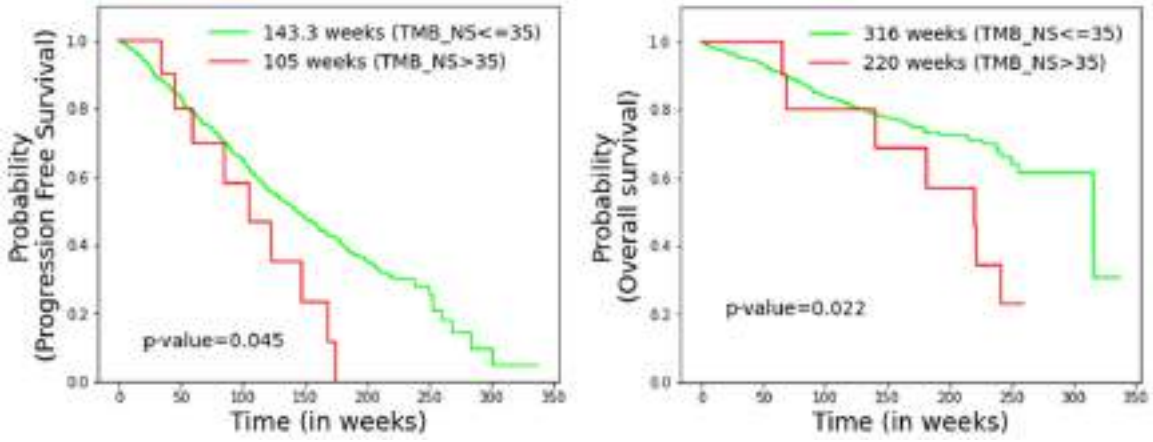


Figure 6. High TMB is associated with poor overall survival in NDMM patients. The difference in the overall survival probability between low and high TMB_NS is statistically significant with *p*-values 0.045 and 0.022 for PFS and OS respectively.

Table 3. The table shows the median values of TMB and SBS for the two groups of MM patients, one where the death event was observed and the other where the death event was not observed

		Median (OS event =0)	Median (OS event =1)	*p-value*
SBS	C>A	17	18	0.018
	C>G	20	21	0.1205
	C>T	59	64	0.038
	T>A	17	16	0.07
	T>C	36	33	0.08
	T>G	19	17	0.02
TMB	TMB_NS	0.4828	0.5766	4.26E-07
	TMB_SYN	0.3487	0.4023	0.002
	TMB_OTH	1.341	1.5288	3.08E-04

Wilcoxon rank-sum test was applied to determine if the change is statistically significant or not. For substitutions, C>A, C>T, and T>G, the frequency was statistically different between the two groups.

patients with TMB_NS greater than 35 are hypermutators, and the characteristics specific to these high-risk patients were examined thoroughly.

Comparison of TMB and SBS based on the overall survival event

Out of 832 MM patients for which survival data were available, 177 observed poor OS outcome while the rest of the 655 MM patients observed superior OS outcome. SBS and TMB values of the two groups were examined, and Wilcoxon rank-sum test was used to deduce if the change in the TMB and SBS values is statistically significant or not. The median SBS and TMB values for the two groups are shown in **Table 3**. There was a significant change (*p*-value <0.05) for SBS T>G, C>A, and C>T. An increase was observed in the C>A and C>T substitution values, while a decrease was observed in T>G substitutions. Further, there is a statistically significant difference in the TMB values of TMB_NS, TMB_SYN, and TMB_OTH, i.e. there was a considerable increase in the tumor mutational burden of the patients with poor outcome as compared to patients with a superior outcome.

Discussion

The fundamental goal of the study was to investigate the entire spectrum of the mutations altered in MGUS and MM, thereby identifying the critical factors responsible for the progression of the disease from MGUS to MM. In this study, we have explored the nonsynonymous and synonymous variants due to their impact on protein expression and function. First of all, variants were identified using four different variant callers to reduce the false positives from the study. Our approach ensured that the variants discovered in our research are the closest possible estimation of the true variants present in the MM and MGUS patients. Variants were then categorized into three main categories-nonsynonymous (NS), synonymous (SYN), and other (OTH) variants. TMB was calculated for each of the three categories of variants. This study reveals changes in the mutational spectrum from MGUS to MM. There was a statistically significant rise in the single base substitutions as the disease progressed from MGUS to MM (**Figure 2**). The frequency pattern of the substitutions in MM is similar to what was observed in a previous study [36]. The highest rise in the frequency was observed in C>T transitions, where the median almost doubled from 30 to 59. An increase in the C>T transitions in MM can be attributed to the overexpression of A3B, an APOBEC cytidine deaminase, that has an essential part in immunity against diseases [37]. Aberrant expression of A3B is known to be correlated with drug resistance, metastasis, and poor prognosis in breast cancer [38], lung cancer [39], and ovarian cancer [40]. Yamazaki et al. [37] proposed that A3B may promote disease progression and drug resistance in MM, which validates our observation of the hike in C>T transitions from MGUS to MM. The association of the frequency of substitutions in the MM patients and their survival outcome was further explored. Frequency of C>T, C>A, and T>G substitutions were significantly higher in MM patients with poor overall survival outcome as compared to MM patients with superior overall survival outcome (**Table 3**). However, in multivariate Cox Hazard analysis (Supplementary Table 5), only C>A transitions have a statistically significant impact on the survival outcome of MM patients.

In addition, SBS2 and SBS13 mutational signatures are linked to APOBEC activity reported in MM in multiple studies [41, 42]. APOBEC signatures were found in nearly 9.63% (98/1018) of the total MM patients, while they were primarily absent in MGUS patients (present in only 1 out of 61 MGUS patients). This finding suggests that ABOPEC activity may be responsible for the molecular mechanisms driving tumor progression from MGUS to MM. The association of ABOPEC activity with overall and progression-free survival in MM was also explored. There was a statistically significant association between the ABOPEC activity and poor overall survival in MM (p-value =0.0056). The KM survival analysis validated this, which yielded significant separation (p-value = 1.8e-4) in the OS curves of MM patients with and without APOBEC activity. Contrary to these findings, no significant association was found between PFS and APOBEC activity. Further, signatures SBS6, SBS14, SSB20, SBS21, SBS26 were found only in MM and are associated with defective DNA mismatch repair and microsatellite instability (MSI) as described previously. MSI has been reported in Multiple Myeloma [43]. However, its frequency is low (~10%) [44]. MSI has been observed to be an effective indicator of response to immunotherapy in solid tumors [45], like colorectal carcinoma [46]. Therefore, it is vital to look for these signatures in MM to help identify the high-risk MM patients in need of immunotherapy.

In the present study, synonymous mutations have been examined along with nonsynonymous mutations. Though synonymous mutations do not change the amino acid sequence of the resulting protein, they have a profound influence on RNA stability, RNA folding [19] or splicing [20], translation [21], or co-translational protein folding. Hence, their role in cancer progression cannot be ignored. There are three different variants categorized under synonymous-synonymous snvs, 3'UTRs and 5'UTRs. A statistically significant rise in the 3'UTR (p-value =2e-16) and 5'UTR (p-value =2.7e-7) mutations were observed from MGUS and MM. 3' untranslated region (UTR) are a part of mRNA containing regulatory binding sites that post-transcriptionally influence gene expression and may lead to disruption in critical pathways associated with different types of cancers. Multiple studies have demonstrated that 3'UTR variants are linked to the risk of developing tumor or tumor progression. Zhang

et al. [47] discovered that a polymorphism detected in the IL-1α 3'UTR of the miRNA-122-binding site was associated with the risk of epithelial ovarian cancer. A unique variant located in the 3'UTR was identified in the gene PCM1, which was significantly associated with ovarian cancer [48]. Recently, Melaiu et al. [49] evaluated the significance of germline genetic variants located within the 3'-untranslated region (polymorphic 3'UTR, i.e., p3UTR) of candidate genes involved in multiple myeloma. Their findings suggested that *IL10*-rs3024496 was associated with an increased risk of developing MM and worse overall survival in MM patients. They also observed that *IL10*-rs3024496 SNP might regulate the *IL10* mRNA expression and hence, could help in the stratification of MM patients in terms of risk progression and prognosis.

5'UTR regions are a part of mRNA, which regulates the protein expression by controlling the translation initiation. Hence, single nucleotide polymorphisms (SNPs) located at 5'UTR regions may alter the protein levels by regulating the mRNA translation efficiency, thereby disturbing consequential biological pathways. The role of 5'UTR variants in multiple cancers has been explored in previous studies. A 5'UTR variant was the driving factor leading to familial breast and ovarian cancer in two independent families [50]. 5'UTR SNP in the PLA2G2A gene was associated with PC metastasis [51]. Thus, it can be concluded that 3' and 5'UTR mutations are more frequent in MM and drive MGUS to MM via regulatory binding sites.

TMB has become a prominent biomarker of enhanced responsiveness to immunotherapy and better outcomes. High TMB is often associated with longer survival after treatment with immune checkpoint inhibitors (ICIs) [16]. However, in non-ICI-treated patients, high TMB was associated with poor prognosis and overall survival in many cancer types [17]. Correlation of high TMB with response to targeted immunotherapies has been established in solid tumors [52, 53]. High somatic mutation and neoantigen loads have been correlated with reduced PFS in MM [54]. However, the association of TMB with overall survival is still unknown in newly diagnosed multiple myeloma (NDMM) patients. Patients with very high TMB_NS values were further analyzed to examine the relation of TMB with OS. These are known as hypermutators and are high-risk patients. Hypermutators demonstrated a significant poor overall survival (p-value =0.022) and poor progression-free survival (p-value =0.045) as compared to non-hypermutators (TMB_NS≤ 35) (**Figure 6**). The median overall survival of hypermutators was 220 weeks compared to 316 weeks of non-hypermutators, while the median progression-free survival of hypermutators was 105 weeks compared to 143.3 weeks non-hypermutators. Mutational signatures SBS1, SBS5, and SBS54 were observed in hypermutators and death events in 7 out of 10 hypermutators. DBS4, DBS5, DBS9, DBS10, and DBS11 are the mutational signatures reflective of double base substitutions (DBS) and were found to be present in hypermutators. On the contrary, no DBS signatures were found in low TMB patients (TMB_NS< 0.1; n=6). SBS1 and SBS5 were present in low TMB patients, including SBS7a, SBS17b, SBS27, SBS51, and SBS86. Our study establishes that the frequency of hypermutators is low in the MM population, and hypermutators are associated with poor OS and poor PFS outcome. Since TMB is a predictor of enhanced responsiveness to immunotherapy, hypermutators may be treated with immunotherapy drugs such as Daratumumab/Elotuzumab [55], Isatuximab [56], and Belantamab Mafodotin [57] to improve their overall survival.

In conclusion, the present study reveals the factors responsible for disease progression from MGUS to MM and poor survival outcome in MM via a detailed investigation of the mutations present in MGUS and MM. The entire landscape of the mutational spectrum involving both synonymous and nonsynonymous mutations was examined. This study finds a change in the mutational spectrum with a statistically significant increase from MGUS to MM. There was a statistically significant increase in the frequency of all the three categories of variants-non-synonymous, synonymous, and others from MGUS to MM (P<0.05). However, there was a statistically significant rise in the TMB values for TMB_NS and TMB_SYN only. We also observed that 3' and 5'UTR mutations were more frequent in MM and might be responsible for driving MGUS to MM via regulatory binding sites. In addition, NDMM patients were also examined separately along with their survival outcome. 10 out of 832 NDMM patients had TMB_NS values greater

than 35 and were designated as hypermutators. It could be concluded that the frequency of hypermutators was low in MM with poor OS and PFS outcome. We also observed a statistically significant rise in the frequency of C>A and C>T substitutions and a statistically significant decline in T>G substitutions. There was a statistically significant increase in the tumor mutational burden of the patients with poor outcome as compared to patients with a superior outcome. Further, a statistically significant association between the APOBEC activity and poor overall survival in MM was discovered. A limitation of the current study is that the number of MGUS patients is significantly less than the number of MM patients. Comparison with a larger dataset of MGUS patients can substantiate the study findings of the significant increase in the mutational frequencies from MGUS to MM. A coherent analysis of evolving mutational landscapes and cancer signatures could assist in designing therapies to impede the transformation of benign MGUS to malignant MM. Additionally, a systematized comparison of high-risk MM patients with low-risk MM patients can aid in identifying the risk factors responsible for disease progression and ultimately guide towards a personalized cure, thereby improving the overall survival of MM patients. A significant rise in 3' and 5'UTR mutations from MGUS to MM was observed in our study. A detailed investigation of these mutations might help understand the mechanism of the progression of MGUS to internecine MM and may be explored in future studies.

Availability of data and materials

Mutation data of 936 MM patients was obtained via dbGaP (MMRF CoMMpass study; phs000748; phs000348), while the remaining 82 patients' exome data was obtained from AIIMS. Exome data of 33 MGUS patients out of 61 patients was obtained from EGA (EGAD-00001001901), and data of the remaining 28 patients was obtained from AIIMS. Variant files of MM patients from the MMRF CoMMpass study were downloaded from the GDC portal via dbGaP authorized access.

Acknowledgements

This work was supported by a grant from the Department of Biotechnology, Govt. of India [Grant: BT/MED/30/SP11006/2015] and Department of Science and Technology, Govt. of India [Grant: DST/ICPS/CPS-Individual/2018/279(G)]. Authors acknowledge dbGaP (Project #18964) for providing authorized access to the MM datasets (phs000748 and phs000-348). Funding support for the phs000348 study was provided by the Multiple Myeloma Research Foundation in collaboration with the Multiple Myeloma Research Consortium. Assistance with data generation, processing, and analysis was provided by the Broad Institute Genome Sequencing, Genetic Analysis, and Biological Samples Platforms. The datasets used for the analyses described in this work were obtained from dbGaP through dbGaP accession number phs000348.v1.p1. Data of study phs000748 were generated as part of the Multiple Myeloma Research Foundation CoMMpass [SM] (Relating Clinical Outcomes in MM to Personal Assessment of Genetic Profile) study (www.themmrf.org). We also acknowledge EGA (EGAD00001001901) for providing authorized access to the MGUS data. The authors would also like to thank the Centre of Excellence in Healthcare, IIIT-Delhi for support in their research. This work was supported by a grant from the Department of Biotechnology, Govt. of India [Grant: BT/MED/30/SP11006/2015] and the Department of Science and Technology, Govt. of India [Grant: DST/ICPS/CPS-Individual/2018/279(G)]. The funding bodies had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding authors had full access to all the data used in the study and had final responsibility for the decision to submit for publication.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Ritu Gupta, Laboratory Oncology Unit, Dr. B.R.A. IRCH, AIIMS, New Delhi 110029, India. E-mail: drritugupta@gmail.com; drritu.laboncology@aiims.edu; Dr. Anubha Gupta, SBILab, Department of ECE, IIIT-Delhi, Core Member, Centre of Excellence in Healthcare, IIIT-Delhi Member, Infosys Centre for AI, IIIT-Delhi, New Delhi 110020, India. E-mail: anubha@iiitd.ac.in

References

[1] Kyle RA, Therneau TM, Rajkumar SV, Larson DR, Plevak MF, Offord JR, Dispenzieri A,

Katzmann JA and Melton LJ 3rd. Prevalence of monoclonal gammopathy of undetermined significance. N Engl J Med 2006; 354: 1362-1369.

[2] Kyle RA, Therneau TM, Rajkumar SV, Offord JR, Larson DR, Plevak MF and Melton LJ 3rd. A long-term study of prognosis in monoclonal gammopathy of undetermined significance. N Engl J Med 2002; 346: 564-569.

[3] Manier S, Salem K, Glavey SV, Roccaro AM and Ghobrial IM. Genomic Aberrations in Multiple Myeloma. Cancer Treat Res 2016; 169: 23-34.

[4] Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Schinzel AC, Harview CL, Brunet JP, Ahmann GJ, Adli M, Anderson KC, Ardlie KG, Auclair D, Baker A, Bergsagel PL, Bernstein BE, Drier Y, Fonseca R, Gabriel SB, Hofmeister CC, Jagannath S, Jakubowiak AJ, Krishnan A, Levy J, Liefeld T, Lonial S, Mahan S, Mfuko B, Monti S, Perkins LM, Onofrio R, Pugh TJ, Rajkumar SV, Ramos AH, Siegel DS, Sivachenko A, Stewart AK, Trudel S, Vij R, Voet D, Winckler W, Zimmerman T, Carpten J, Trent J, Hahn WC, Garraway LA, Meyerson M, Lander ES, Getz G and Golub TR. Initial genome sequencing and analysis of multiple myeloma. Nature 2011; 471: 467-472.

[5] Fonseca R, Blood EA, Oken MM, Kyle RA, Dewald GW, Bailey RJ, Van Wier SA, Henderson KJ, Hoyer JD, Harrington D, Kay NE, Van Ness B and Greipp PR. Myeloma and the t(11;14)(q13;q32); evidence for a biologically defined unique subset of patients. Blood 2002; 99: 3735-3741.

[6] Walker BA, Wardell CP, Melchor L, Brioli A, Johnson DC, Kaiser MF, Mirabella F, Lopez-Corral L, Humphray S, Murray L, Ross M, Bentley D, Gutiérrez NC, Garcia-Sanz R, San Miguel J, Davies FE, Gonzalez D and Morgan GJ. Intraclonal heterogeneity is a critical early event in the development of myeloma and precedes the development of clinical symptoms. Leukemia 2014; 28: 384-390.

[7] Dhodapkar MV. MGUS to myeloma: a mysterious gammopathy of underexplored significance. Blood 2016; 128: 2599-2606.

[8] Dutta AK, Fink JL, Grady JP, Morgan GJ, Mullighan CG, To LB, Hewett DR and Zannettino ACW. Subclonal evolution in disease progression from MGUS/SMM to multiple myeloma is characterised by clonal stability. Leukemia 2019; 33: 457-468.

[9] Pleasance ED, Stephens PJ, O'Meara S, McBride DJ, Meynert A, Jones D, Lin ML, Beare D, Lau KW, Greenman C, Varela I, Nik-Zainal S, Davies HR, Ordoñez GR, Mudie LJ, Latimer C, Edkins S, Stebbings L, Chen L, Jia M, Leroy C, Marshall J, Menzies A, Butler A, Teague JW, Mangion J, Sun YA, McLaughlin SF, Peckham HE, Tsung EF, Costa GL, Lee CC, Minna JD, Gazdar A, Birney E, Rhodes MD, McKernan KJ, Stratton MR, Futreal PA and Campbell PJ. A small-cell lung cancer genome with complex signatures of tobacco exposure. Nature. 2010; 463: 184-190.

[10] Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, Bignell G, Davies H, Teague J, Butler A, Stevens C, Edkins S, O'Meara S, Vastrik I, Schmidt EE, Avis T, Barthorpe S, Bhamra G, Buck G, Choudhury B, Clements J, Cole J, Dicks E, Forbes S, Gray K, Halliday K, Harrison R, Hills K, Hinton J, Jenkinson A, Jones D, Menzies A, Mironenko T, Perry J, Raine K, Richardson D, Shepherd R, Small A, Tofts C, Varian J, Webb T, West S, Widaa S, Yates A, Cahill DP, Louis DN, Goldstraw P, Nicholson AG, Brasseur F, Looijenga L, Weber BL, Chiew YE, DeFazio A, Greaves MF, Green AR, Campbell P, Birney E, Easton DF, Chenevix-Trench G, Tan MH, Khoo SK, Teh BT, Yuen ST, Leung SY, Wooster R, Futreal PA and Stratton MR. Patterns of somatic mutation in human cancer genomes. Nature 2007; 446: 153-158.

[11] Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, Greenman CD, Varela I, Lin ML, Ordóñez GR, Bignell GR, Ye K, Alipaz J, Bauer MJ, Beare D, Butler A, Carter RJ, Chen L, Cox AJ, Edkins S, Kokko-Gonzales PI, Gormley NA, Grocock RJ, Haudenschild CD, Hims MM, James T, Jia M, Kingsbury Z, Leroy C, Marshall J, Menzies A, Mudie LJ, Ning Z, Royce T, Schulz-Trieglaff OB, Spiridou A, Stebbings LA, Szajkowski L, Teague J, Williamson D, Chin L, Ross MT, Campbell PJ, Bentley DR, Futreal PA and Stratton MR. A comprehensive catalogue of somatic mutations from a human cancer genome. Nature 2010; 463: 191-196.

[12] Kasar S, Kim J, Improgo R, Tiao G, Polak P, Haradhvala N, Lawrence MS, Kiezun A, Fernandes SM, Bahl S, Sougnez C, Gabriel S, Lander ES, Kim HT, Getz G and Brown JR. Whole-genome sequencing reveals activation-induced cytidine deaminase signatures during indolent chronic lymphocytic leukaemia evolution. Nat Commun 2015; 6: 1-12.

[13] Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K, Jones D, Hinton J, Marshall J, Stebbings LA, Menzies A, Martin S, Leung K, Chen L, Leroy C, Ramakrishna M, Rance R, Lau KW, Mudie LJ, Varela I, McBride DJ, Bignell GR, Cooke SL, Shlien A, Gamble J, Whitmore I, Maddison M, Tarpey PS, Davies HR, Papaemmanuil E, Stephens PJ, McLaren S, Butler AP, Teague JW, Jönsson G, Garber JE, Silver D, Miron P, Fatima A, Boyault S, Langerød A, Tutt A, Martens JW, Aparicio SA, Borg Å, Salomon AV, Thomas G, Børresen-Dale AL, Richardson AL, Neuberger MS, Futreal PA, Camp-

bell PJ and Stratton MR; Breast Cancer Working Group of the International Cancer Genome Consortium. Mutational processes molding the genomes of 21 breast cancers. Cell 2012; 149: 979-993.
[14] Davies H, Morganella S, Purdie CA, Jang SJ, Borgen E, Russnes H, Glodzik D, Zou X, Viari A, Richardson AL, Børresen-Dale AL, Thompson A, Eyfjord JE, Kong G, Stratton MR and Nik-Zainal S. Whole-genome sequencing reveals breast cancers with mismatch repair deficiency. Cancer Res 2017; 77: 4755-4762.
[15] Alexandrov LB, Ju YS, Haase K, Van Loo P, Martincorena I, Nik-Zainal S, Totoki Y, Fujimoto A, Nakagawa H, Shibata T, Campbell PJ, Vineis P, Phillips DH and Stratton MR. Mutational signatures associated with tobacco smoking in human cancer. Science 2016; 354: 618-622.
[16] Fusco MJ, West HJ and Walko CM. Tumor mutation burden and cancer treatment. JAMA Oncol 2021; 7: 316.
[17] Valero C, Lee M, Hoen D, Wang J, Nadeem Z, Patel N, Postow MA, Shoushtari AN, Plitas G, Balachandran VP, Smith JJ, Crago AM, Long Roche KC, Kelly DW, Samstein RM, Rana S, Ganly I, Wong RJ, Hakimi AA, Berger MF, Zehir A, Solit DB, Ladanyi M, Riaz N, Chan TA, Seshan VE and Morris LGT. The association between tumor mutational burden and prognosis is dependent on treatment context. Nat Genet 2021; 53: 11-15.
[18] Sharma Y, Miladi M, Dukare S, Boulay K, Caudron-Herger M, Groß M, Backofen R and Diederichs S. A pan-cancer analysis of synonymous mutations. Nat Commun 2019; 10: 1-14.
[19] Goodman DB, Church GM and Kosuri S. Causes and effects of N-terminal codon bias in bacterial genes. Science 2013; 342: 475-479.
[20] Parmley JL, Chamary JV and Hurst LD. Evidence for purifying selection against synonymous mutations in mammalian exonic splicing enhancers. Mol Biol Evol 2006; 23: 301-309.
[21] Drummond DA and Wilke CO. Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. Cell 2008; 134: 341-352.
[22] Supek F, Skunca N, Repar J, Vlahovicek K and Smuc T. Translational selection is ubiquitous in prokaryotes. PLoS Genet 2010; 6: e1001004.
[23] Savisaar R and Hurst LD. Exonic splice regulation imposes strong selection at synonymous sites. Genome Res 2018; 28: 1442-1454.
[24] Kimura M. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. Nature 1977; 267: 275-276.
[25] Fan Y, Xi L, Hughes DS, Zhang J, Zhang J, Futreal PA, Wheeler DA and Wang W. MuSE: accounting for tumor heterogeneity using a sample-specific error model improves sensitivity and specificity in mutation calling from sequencing data. Genome Biol 2016; 17: 1-11.
[26] Benjamin D, Sato T, Cibulskis K, Getz G, Stewart C and Lichtenstein L. Calling somatic SNVs and indels with Mutect2. BioRxiv 2019; 861054.
[27] Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L and Wilson RK. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 2012; 22: 568-576.
[28] Larson DE, Harris CC, Chen K, Koboldt DC, Abbott TE, Dooling DJ, Ley TJ, Mardis ER, Wilson RK and Ding L. SomaticSniper: identification of somatic point mutations in whole genome sequencing data. Bioinformatics 2012; 28: 311-317.
[29] Wang K and Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010; 38: e164.
[30] Rogers MF, Shihab HA, Mort M, Cooper DN, Gaunt TR and Campbell C. FATHMM-XF: accurate prediction of pathogenic point mutations via extended features. Bioinformatics 2018; 34: 511-513.
[31] Islam SA, Wu Y, Díaz-Gay M, Bergstrom EN, He Y, Barnes M, Vella M, Wang J, Teague JW, Clapham P and Moody S. Uncovering novel mutational signatures by de novo extraction with SigProfilerExtractor. BioRxiv 2021;2020-12.
[32] Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, Boot A, Covington KR, Gordenin DA, Bergstrom EN, Islam SMA, Lopez-Bigas N, Klimczak LJ, McPherson JR, Morganella S, Sabarinathan R, Wheeler DA and Mustonen V; PCAWG Mutational Signatures Working Group, Getz G, Rozen SG and Stratton MR; PCAWG Consortium. The repertoire of mutational signatures in human cancer. Nature 2020; 578: 94-101.
[33] Farswan A, Jena L, Kaur G, Gupta A, Gupta R, Rani L, Sharma A and Kumar L. Branching clonal evolution patterns predominate mutational landscape in multiple myeloma. Am J Cancer Res 2021; 11: 5659-5679.
[34] Eo SH, Kang HJ, Hong SM and Cho H. K-adaptive partitioning for survival data, with an application to cancer staging. arXiv preprint arXiv:1306.4615. 2013.
[35] Budczies J, Klauschen F, Sinn BV, Győrffy B, Schmitt WD, Darb-Esfahani S and Denkert C. Cutoff Finder: a comprehensive and straightforward Web application enabling rapid biomarker cutoff optimization. PLoS One 2012; 7: e51862.

[36] Bolli N, Avet-Loiseau H, Wedge DC, Van Loo P, Alexandrov LB, Martincorena I, Dawson KJ, Iorio F, Nik-Zainal S, Bignell GR, Hinton JW, Li Y, Tubio JM, McLaren S, O'Meara S, Butler AP, Teague JW, Mudie L, Anderson E, Rashid N, Tai YT, Shammas MA, Sperling AS, Fulciniti M, Richardson PG, Parmigiani G, Magrangeas F, Minvielle S, Moreau P, Attal M, Facon T, Futreal PA, Anderson KC, Campbell PJ and Munshi NC. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nat Commun 2014; 5: 1-13.

[37] Yamazaki H, Shirakawa K, Matsumoto T, Hirabayashi S, Murakawa Y, Kobayashi M, Sarca AD, Kazuma Y, Matsui H, Maruyama W, Fukuda H, Shirakawa R, Shindo K, Ri M, Iida S and Takaori-Kondo A. Endogenous APOBEC3B overexpression constitutively generates DNA substitutions and deletions in myeloma cells. Sci Rep 2019; 9: 1-14.

[38] Law EK, Sieuwerts AM, LaPara K, Leonard B, Starrett GJ, Molan AM, Temiz NA, Vogel RI, Meijer-van Gelder ME, Sweep FC, Span PN, Foekens JA, Martens JW, Yee D and Harris RS. The DNA cytosine deaminase APOBEC3B promotes tamoxifen resistance in ER-positive breast cancer. Sci Adv 2016; 2: e1601737.

[39] Yan S, He F, Gao B, Wu H, Li M, Huang L, Liang J, Wu Q and Li Y. Increased APOBEC3B predicts worse outcomes in lung cancer: a comprehensive retrospective study. J Cancer 2016; 7:618-625.

[40] Du Y, Tao X, Wu J, Yu H, Yu Y and Zhao H. APOBEC3B up-regulation independently predicts ovarian cancer prognosis: a cohort study. Cancer Cell Int 2018; 18: 1-10.

[41] Walker BA, Wardell CP, Murison A, Boyle EM, Begum DB, Dahir NM, Proszek PZ, Melchor L, Pawlyn C, Kaiser MF, Johnson DC, Qiang YW, Jones JR, Cairns DA, Gregory WM, Owen RG, Cook G, Drayson MT, Jackson GH, Davies FE and Morgan GJ. APOBEC family mutational signatures are associated with poor prognosis translocations in multiple myeloma. Nat Commun 2015; 6: 1-11.

[42] Hoang PH, Cornish AJ, Dobbins SE, Kaiser M and Houlston RS. Mutational processes contributing to the development of multiple myeloma. Blood Cancer J 2019; 9: 1-11.

[43] Timurağaoğlu A, Demircin S, Dizlek S, Alanoğlu G and Kiriş E. Microsatellite instability is a common finding in multiple myeloma. Clin Lymphoma Myeloma 2009; 9: 371-374.

[44] Miyashita K, Fujii K, Suehiro Y, Taguchi K, Uike N, Yoshida MA and Oda S. Heterochronous occurrence of microsatellite instability in multiple myeloma - an implication for a role of defective DNA mismatch repair in myelomagenesis. Leuk Lymphoma 2018; 59: 2454-2459.

[45] Chang L, Chang M, Chang HM, Chang F. Microsatellite Instability: A Predictive Biomarker for Cancer Immunotherapy. Appl Immunohistochem Mol Morphol 2018; 26: e15-e21.

[46] Oh DY, Venook AP and Fong L. On the Verge: Immunotherapy for Colorectal Carcinoma. J Natl Compr Canc Netw 2015; 13: 970-978.

[47] Zhang Z, Zhou B, Gao Q, Wu Y, Zhang K, Pu Y, Song Y, Zhang L and Xi M. A polymorphism at miRNA-122-binding site in the IL-1α 3'UTR is associated with risk of epithelial ovarian cancer. Fam Cancer 2014; 13: 595-601.

[48] Chen X, Paranjape T, Stahlhut C, McVeigh T, Keane F, Nallur S, Miller N, Kerin M, Deng Y, Yao X, Zhao H, Weidhaas JB and Slack FJ. Targeted resequencing of the microRNAome and 3'UTRome reveals functional germline DNA variants with altered prevalence in epithelial ovarian cancer. Oncogene 2015; 34: 2125-2137.

[49] Melaiu O, Macauda A, Sainz J, Calvetti D, Facioni MS, Maccari G, Ter Horst R, Netea MG, Li Y, Grząśko N, Moreno V, Jurczyszyn A, Jerez A, Watek M, Varkonyi J, Garcia-Sanz R, Kruszewski M, Dudziński M, Kadar K, Jacobsen SEH, Mazur G, Andersen V, Rybicka M, Zawirska D, Raźny M, Zaucha JM, Ostrovsky O, Iskierka-Jazdzewska E, Reis RM, Stępień A, Beider K, Nagler A, Druzd-Sitek A, Marques H, Martìnez-Lopez J, Lesueur F, Avet-Loiseau H, Vangsted AJ, Krawczyk-Kulis M, Butrym A, Jamroziak K, Dumontet C, Vogel U, Rymko M, Pelosini M, Subocz E, Szombath G, Sarasquete ME, Silvestri R, Morani F, Landi S, Campa D, Canzian F and Gemignani F. Common gene variants within 3'-untranslated regions as modulators of multiple myeloma risk and survival. Int J Cancer 2021; 148: 1887-1894.

[50] Evans DGR, van Veen EM, Byers HJ, Wallace AJ, Ellingford JM, Beaman G, Santoyo-Lopez J, Aitman TJ, Eccles DM, Lalloo FI, Smith MJ and Newman WG. A dominantly inherited 5'UTR variant causing methylation-associated silencing of BRCA1 as a cause of breast and ovarian cancer. Am J Hum Genet 2018; 103: 213-220.

[51] Ozturk K, Onal MS, Efiloglu O, Nikerel E, Yildirim A and Telci D. Association of 5'UTR polymorphism of secretory phospholipase A2 group IIA (PLA2G2A) gene with prostate cancer metastasis. Gene 2020; 742: 144589.

[52] Halbert B and Einstein DJ. Hot or Not: Tumor mutational Burden (TMB) as a biomarker of immunotherapy response in genitourinary cancers. Urology 2021; 147: 119-126.

[53] Boumber Y. Tumor mutational burden (TMB) as a biomarker of response to immunotherapy in small cell lung cancer. J Thorac Dis 2018; 10: 4689-4693

[54] Miller A, Cattaneo L, Asmann YW, Braggio E, Keats JJ, Auclair D, Lonial S, Network TM, Russell SJ and Stewart AK. Correlation between somatic mutation burden, neoantigen load and progression free Survival in multiple myeloma: analysis of MMRF CoMMpass study. Blood 2016; 128: 193.

[55] Laubach JP, Paba Prada CE, Richardson PG and Longo DL. Daratumumab, elotuzumab, and the development of therapeutic monoclonal antibodies in multiple myeloma. Clin Pharmacol Ther 2017; 101: 81-88.

[56] Moreno L, Perez C, Zabaleta A, Manrique I, Alignani D, Ajona D, Blanco L, Lasa M, Maiso P, Rodriguez I, Garate S, Jelinek T, Segura V, Moreno C, Merino J, Rodriguez-Otero P, Panizo C, Prosper F, San-Miguel JF and Paiva B. The mechanism of action of the anti-CD38 monoclonal antibody isatuximab in multiple myeloma. Clin Cancer Res 2019; 25: 3176-3187.

[57] Offidani M, Corvatta L, Morè S and Olivieri A. Belantamab mafodotin for the treatment of multiple myeloma: an overview of the clinical efficacy and safety. Drug Des Devel Ther 2021; 15: 2401-2415.

Supplementary Table 1. The table shows the cut-offs obtained for TMB_OTH via KAP

	Min	Median	Max	Cut-off (≤, >)	PFS	OS
TMB_OTH	0.1114	1.3742	193.673	1.84 (666, 166)	4.90E-06	9.16E-09

The same cut-off was obtained using PFS and OS. KM analysis was done using the proposed cut-offs. There was a significant difference (*p*-value <0.05) on the KM survival curves of the patients below and above the selected cut-offs of TMB_OTH.

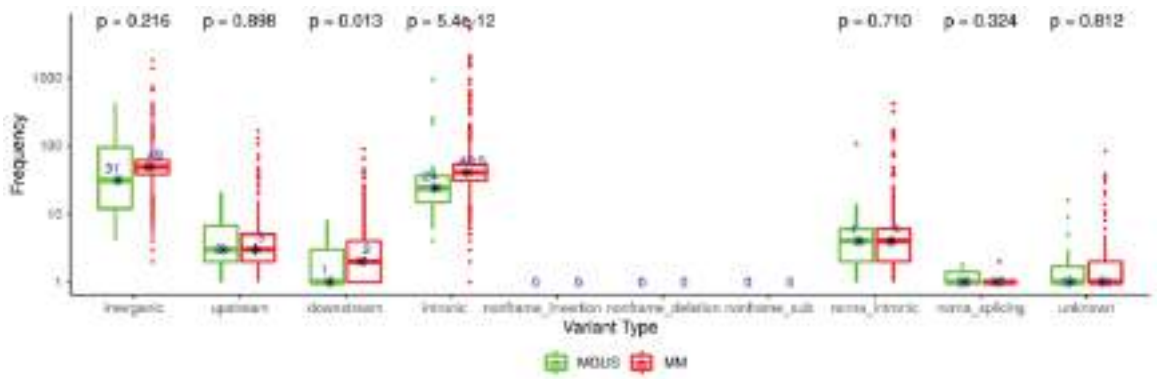
Supplementary Table 2. The table shows the cut-offs obtained for TMB_NS, TMB_SYN and TMB_OTH via Cutoff Finder

	Cut-off via PFS	*p*-value	Cut-off via OS	*p*-value
TMB_NS	0.6282	3.2E-07	0.6216	2.1E-08
TMB_SYN	0.5565	4.1E-05	0.5265	2.3E-09
TMB_OTH	1.84	4.9E-06	1.84	9.2E-09

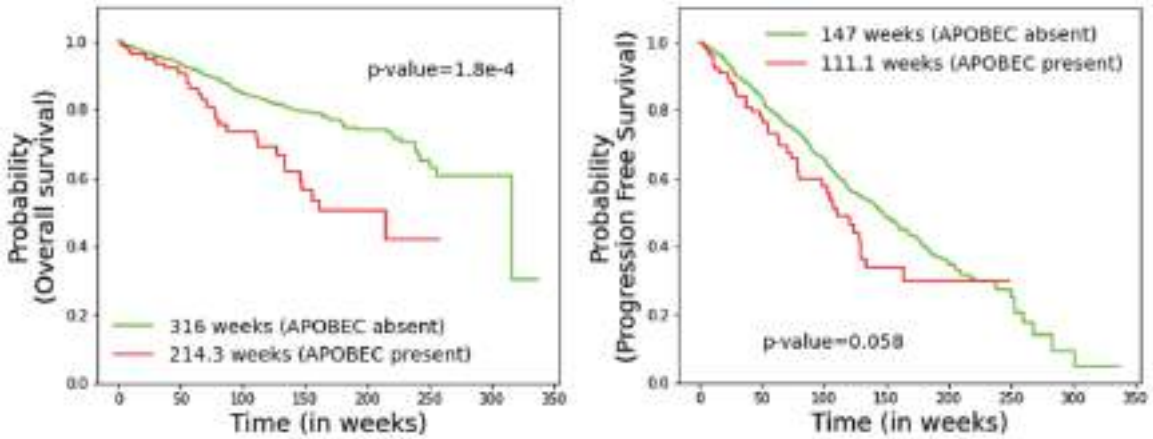
Similar cut-offs were obtained using PFS and OS. KM analysis revealed a significant difference (*p*-value <0.05) on the survival curves of the patients below and above the selected cut-offs.

Supplementary Table 3. The table shows the cut-offs obtained for the six different types of substitutions via Cutoff Finder

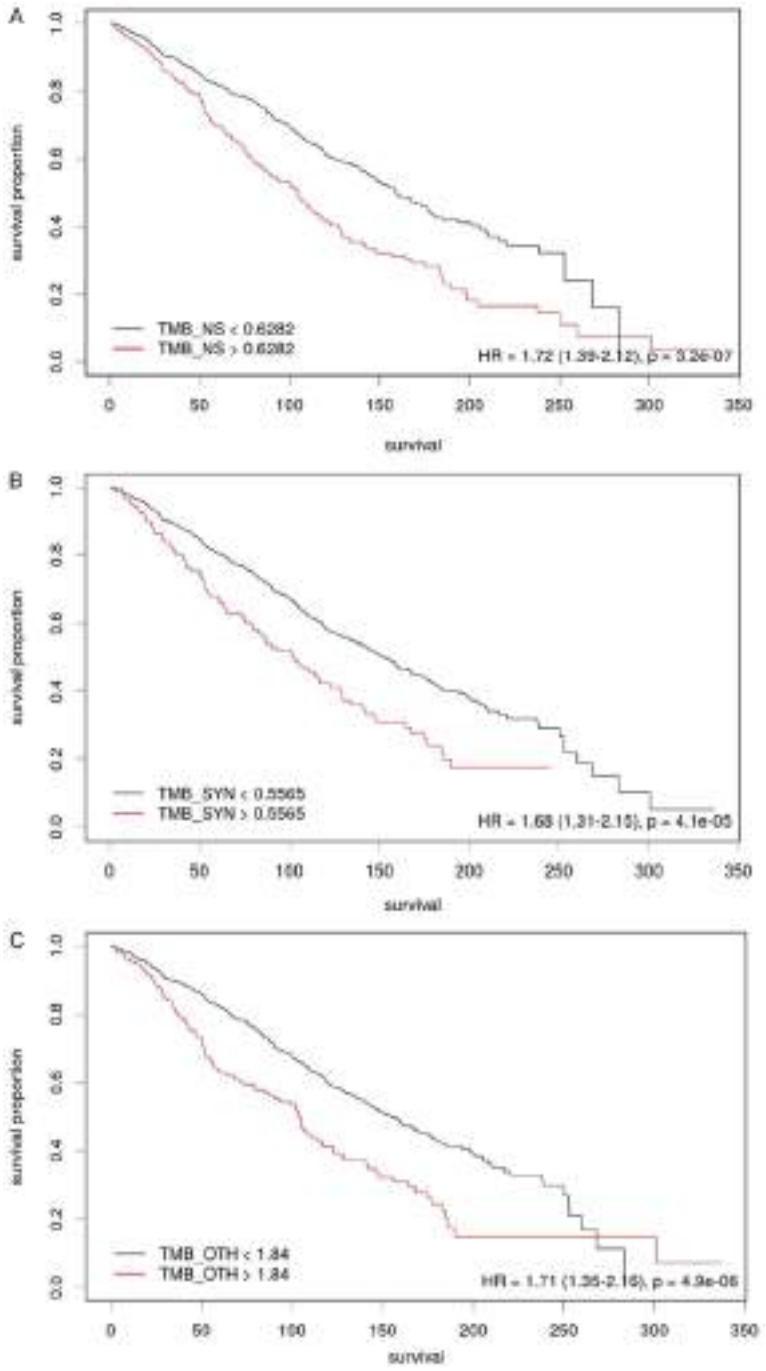
SBS	Min	Median	Max	PFS cutoff	OS cutoff	PFS *p*-value	OS *p*-value
C>A	0	17	1251	26.5	28.5	1.9E-5	5.1E-06
C>G	0	21	1575	3.5	34.5	0.027	8.6E-05
C>T	1	59	7315	79.5	110	0.00055	5.8E-07
T>A	0	17	684	5.5	2.5	0.0046	0.0024
T>C	0	35	4498	12.5	11.5	0.00019	0.0027
T>G	0	19	915	6.5	6.5	0.0013	0.0029



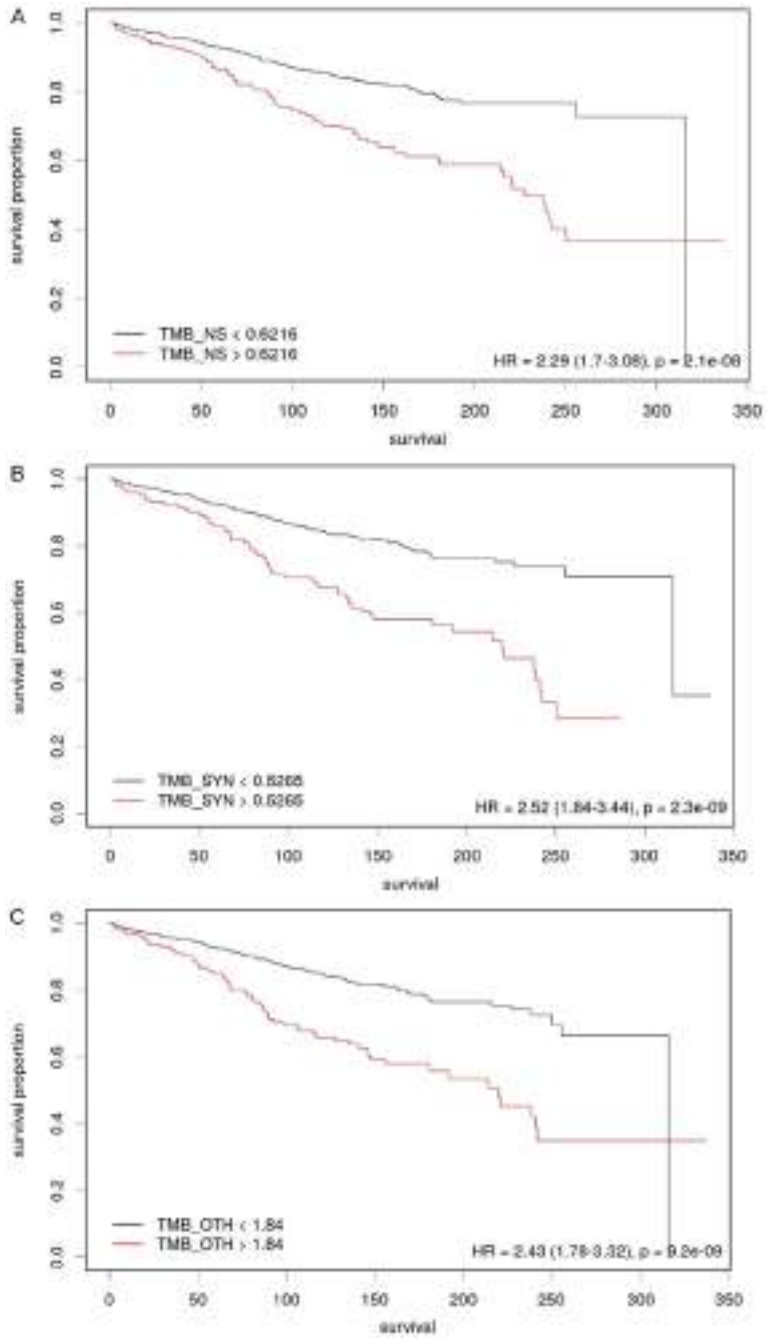
Supplementary Figure 1. Boxplot showing the variation in the frequency of the variants under the other variants category. There was a statistically significant rise in the frequency of intronic, intergenic, and downstream variants with *p*-values less than 0.05.



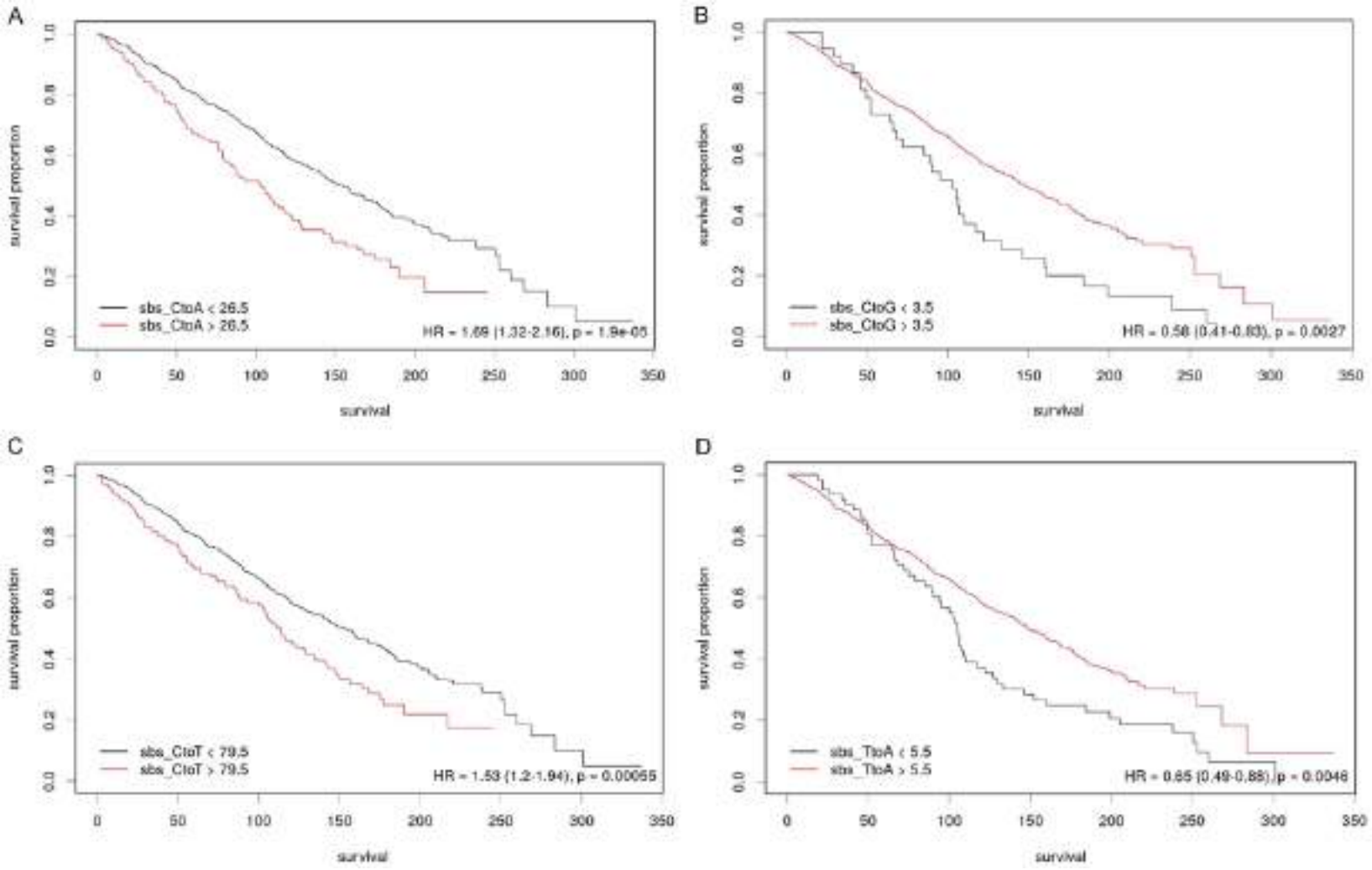
Supplementary Figure 2. KM curves reveal that APOBEC activity is associated with poor overall survival in NDMM patients. The difference in the overall survival probability between low and high TMB_NS is statistically significant with *p*-values 1.8e-4. However, there is no statistically significant difference between progression-free survival and APOBEC activity.

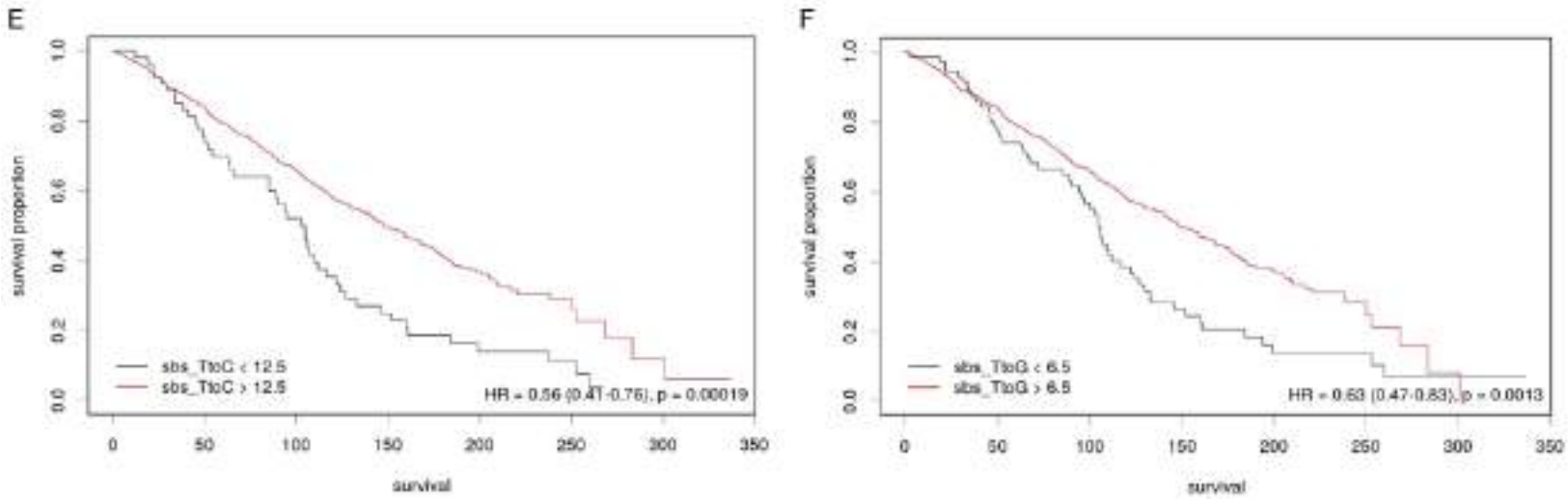


Supplementary Figure 3. KM curves reveal significant differences in the PFS survival patterns of (A) TMB_NS, (B) TMB_SYN and (C) TMB_OTH at the thresholds obtained via Cutoff Finder.

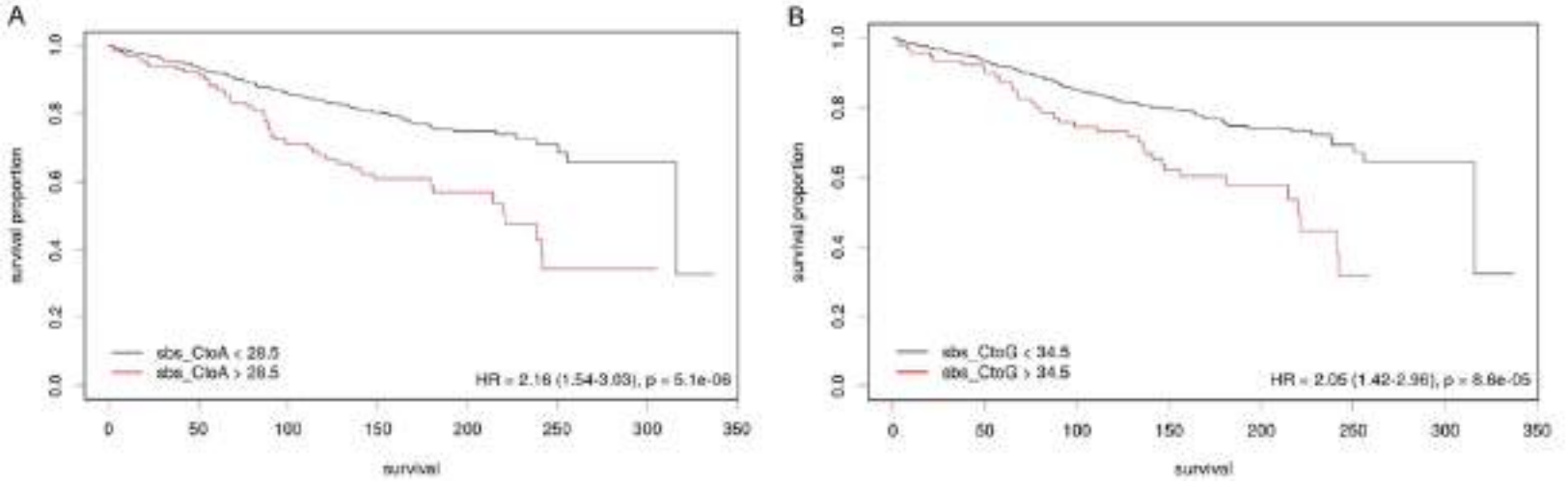


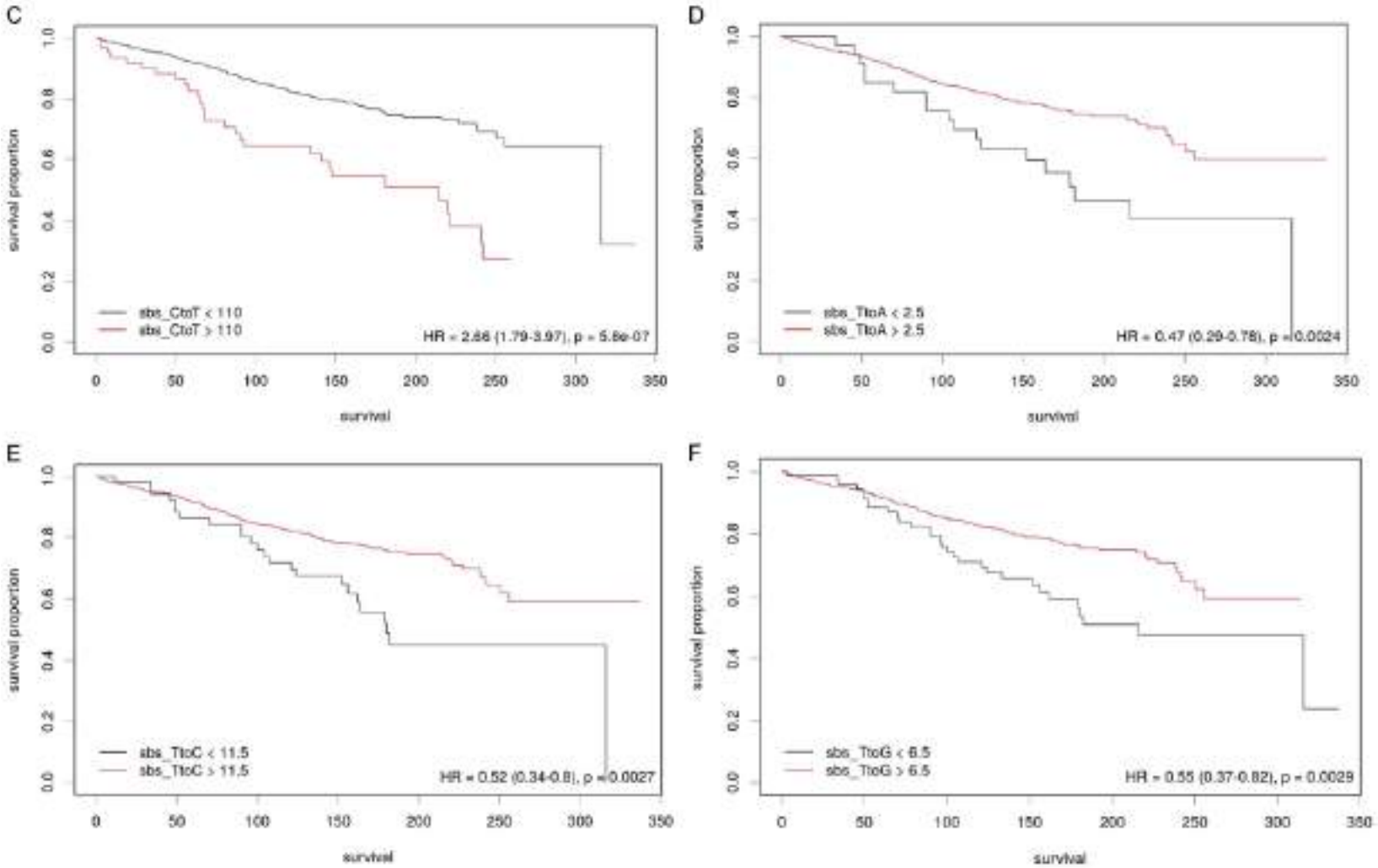
Supplementary Figure 4. KM curves reveal significant differences in the OS survival patterns of (A) TMB_NS, (B) TMB_SYN and (C) TMB_OTH at the thresholds obtained via Cutoff Finder.


A
survival proportion
survival
sbs_CtoA < 26.5
sbs_CtoA > 26.5
HR = 1.69 (1.32-2.16), p = 1.9e-05
B
survival proportion
survival
sbs_CtoG < 3.5
sbs_CtoG > 3.5
HR = 0.58 (0.41-0.83), p = 0.0027
C
survival proportion
survival
sbs_CtoT < 79.5
sbs_CtoT > 79.5
HR = 1.53 (1.2-1.94), p = 0.00055
D
survival proportion
survival
sbs_TtoA < 5.5
sbs_TtoA > 5.5
HR = 0.65 (0.49-0.88), p = 0.0046



Supplementary Figure 5. KM curves reveal differences in the PFS survival patterns of substitutions (A) C>A, (B) C>G, (C) C>T, (D) T>A, (E) T>C and (F) T>G at the thresholds obtained via Cutoff Finder. Separation in the survival curves is significant if *p*-values <0.05.





Supplementary Figure 6. KM curves reveal differences in the OS survival patterns of substitutions (A) C>A, (B) C>G, (C) C>T, (D) T>A, (E) T>C and (F) T>G at the thresholds obtained via Cutoff Finder. Separation in the survival curves is significant if *p*-values <0.05.

Supplementary Table 4. The table shows the univariate hazard analysis and multivariate hazard analysis obtained on TMB_NS, TMB_SYN and TMB_OTH

	pfs				os			
	HR	CI	*p*-value	C-index	HR	CI	*p*-value	C-index
Univariate								
TMB_NS	1.71	1.39-2.12	<0.005	0.56	2.26	1.68-3.05	<0.005	0.58
TMB_SYN	1.68	1.31-2.15	<0.005	0.54	2.46	1.78-3.40	<0.005	0.56
TMB_OTH	1.71	1.35-2.16	<0.005	0.55	2.43	1.78-3.32	<0.005	0.58
Multivariate								
TMB_NS	1.45	1.11-1.90	0.01	0.57	1.55	1.04-2.31	0.03	0.6
TMB_SYN	1.13	0.81-1.58	0.48		1.41	0.89-2.24	0.14	
TMB_OTH	1.26	0.92-1.74	0.16		1.48	0.94-2.34	0.09	

Supplementary Table 5. The table shows the univariate hazard analysis and multivariate hazard analysis on the six different substitutions

	PFS				OS			
	HR	CI	*p*-value	C-index	HR	CI	*p*-value	C-index
Univariate								
C>A	1.63	1.26-2.11	<0.005	0.54	2.16	1.54-3.03	<0.005	0.55
C>G	1.46	1.04-2.04	0.03	0.52	2.11	1.41-3.16	<0.005	0.53
C>T	1.65	1.22-2.24	<0.005	0.53	2.36	1.61-3.45	<0.005	0.55
T>A	1.61	1.11-2.32	0.01	0.51	1.93	1.20-3.11	0.01	0.52
T>C	1.47	0.83-2.61	0.19	0.50	2.27	1.19-4.32	0.01	0.51
T>G	1.73	1.09-2.75	0.02	0.51	2.14	1.21-3.77	0.01	0.51
Multivariate								
C>A	1.43	1.02-1.99	0.04	0.55	1.67	1.06-2.63	0.03	0.58
C>G	0.84	0.49-1.43	0.52		0.97	0.49-1.93	0.93	
C>T	1.38	0.86-2.22	0.18		1.71	0.91-3.22	0.10	
T>A	1.19	0.71-1.97	0.65		1.22	0.61-2.44	0.58	
T>G	1.03	0.52-2.05	0.94		0.82	0.34-1.99	0.66	

T>C was removed from multivariate analysis as it was not significant for PFS in univariate analysis.

Translational Oncology 23 (2022) 101472

Contents lists available at ScienceDirect

Translational Oncology

journal homepage: www.elsevier.com/locate/tranon

Original Research

Correlation of changes in subclonal architecture with progression in the MMRF CoMMpass study

Gurvinder Kaur [a,1], Lingaraja Jena [a,1], Ritu Gupta [a,*], Akanksha Farswan [b], Anubha Gupta [b,*], K Sriram [c]

[a] Laboratory Oncology Unit, Dr. B. R.A. IRCH, AIIMS, New Delhi
[b] SBILab, Department of Electronics and Communication Engineering, IIIT, Delhi
[c] Department of Computational Biology & Centre for Computational Biology, IIIT, Delhi

ARTICLE INFO

Keywords:
Multiple myeloma
Clonal evolution
Actionable genes
Oncogenic correlations
Gene enrichment pathways
Molecular signatures

ABSTRACT

Multiple myeloma (MM) is a heterogeneous plasma cell proliferative disorder that arises from its premalignant precursor stages through a complex cascade of interactions between clonal mutations and co-evolving microenvironment. The temporo-spatial evolutionary trajectories of MM are established early during myelomatogenesis in precursor stages and retained in MM. Such molecular events impact subsequent disease progression and clinical outcomes. Identification of clonal sweeps of actionable gene targets in MM could reveal potential vulnerabilities that may exist in early stages and thus potentiate prognostication and customization of early therapeutic interventions. We have evaluated clonal evolution at multiple time points in 76 MM patients enrolled in the MMRF CoMMpass **study**. The major findings of this study are (a) MM progresses predominantly through branching evolution, (b) there is a heterogeneous spectrum of mutational landscapes that include unique actionable gene targets at diagnosis compared to progression, (c) unique clonal gains/ losses of mutant driver genes can be identified in patients with different cytogenetic aberrations, (d) there is a significant correlation between co-occurring oncogenic mutations/ co-occurring subclones e.g., with mutated TP53+SYNE1, NRAS+MAGI3, and anticorrelative dependencies between FAT3+FCGBP gene pairs. Such co-trajectories may synchronize molecular events of drug response, myelomatogenesis and warrant future studies to explore their potential for early prognostication and development of risk stratified personalized therapies in MM.

Introduction

Multiple myeloma (MM) is a plasma cell malignancy characterized by a complex genomic landscape, heterogeneity in response to therapy and clinical outcomes [1]. Myelomatogenesis is a multistep process that is initiated and driven by accumulation of composite genomic aberrations [2–7]. The primary cytogenetic aberrations include translocations involving IgH locus (found in about 55% patients) and hyperdiploid trisomies of odd numbered chromosomes (found in about 40% patients) [8,9]. While IgH translocations t(4;14) (partner *FGFR3*), t(14;16) (*MAF*) and t(14;20) (*MAFB*) have been associated with high risk disease and poor prognosis; other translocations t(11;14) (*CCND1*), t(6;14) (*CCND3*) and trisomies are considered to be associated with standard risk [10]. Deletion 13q is also considered to be an early event and is observed in about 35% of the patients. There are additional multiple secondary events that evolve gradually and include mutations in RAS/NFkB pathway genes, haploinsufficiency of p73, high expression of *MYC*, 1q gain, 1p loss and others. Recently, 8 copy number subtypes [11] and 12 RNA subgroups [12] of MM, have also been characterized from the Multiple Myeloma Research Foundation (MMRF) CoMMpass study [13].

Molecular analysis of myeloma genome through conventional cytogenetics and fluorescence in situ hybridization (FISH), gene expression profiling, Whole Genome Sequencing (WGS), Whole Exome Sequencing (WES) and microarrays has provided novel insights into the genomics of MM [14–17], role of driver genes [18–22] and chromothripsis [23,24], in risk prediction strategies and identified different oncogenic events involved in the immunity-malignancy equilibrium in MM [25,26]. Multiple mechanisms involving deregulated signaling cascades can derail oncogenic pathways [27,28] and their evaluation can lead to development of pathway directed therapies [29] for MM.

* Corresponding authors.
E-mail addresses: drritu.laboncology@aiims.edu (R. Gupta), anubha@iiitd.ac.in (A. Gupta).
[1] Both Gurvinder Kaur and Lingaraja Jena contributed equally.

https://doi.org/10.1016/j.tranon.2022.101472
Received 11 September 2021; Received in revised form 3 June 2022; Accepted 20 June 2022
Available online 28 June 2022

A series of studies have established clonal heterogeneity in MM [30–33]. There is evidence that as disease progresses from its premalignant stages of MGUS (Monoclonal Gammopathy of Undetermined Significance) and SMM (Smouldering MM) to MM [34], the (sub)clones tend to evolve new mutations and CNAs (copy number alterations), imposed by positive selection pressure, and sweep to dominance [35–39]. It has been reported that the genomic landscapes of SMM are very similar to those at MM with two major patterns of progression [35]. While the static progression model retains the subclonal architecture as disease progresses to MM, a spontaneous evolution model supports changes in subclonal composition. A recent study [37] evaluated both Darwinian positive selection and neutrality in evolution in WGS CoMMpass data and surmised that neutral mutations might occur in a few tumors but Darwinian evolution appears to be the overall dominant trajectory of MM.

The temporal and spatial clonal trajectories established by the time of SMM are thus preserved and can predict the course of subsequent tumor evolution and contribute to drug resistance [35–37]. This also provides a unique opportunity to revisit and deploy prognostication at premalignant stages and adopt early intervention. There is limited information available on subclonal evolution associated with progression in MM and, in particular, on clinically relevant actionable targets. Moreover, there is no report on associations between co-occurring clonal gains/ losses in MM. This enthused us to evaluate clonal landscapes of the MMRF CoMMpass WES data at baseline time point (TP1) and compare to subsequent time points of progression (TP2, TP3, TP4 and TP5). We have analyzed 76 MM patients for whom multiple time point WES data were available in the MMRF IA12 dataset, and compared clonal mutational landscapes and pathways, at baseline with those at progression.

This study has shown that clonal expansion to progression in MM follows predominantly branching pattern of evolution, with differential waxing and waning of mutated (sub)clones. A series of subclonal mutations in driver and actionable genes could be identified, some of which showed patterns of mutational/ subclonal co-occurrence or exclusions with significant correlations. Identification of serial correlative clonal trajectories of drivers and actionable targets in MM may advance prognostication and assist in personalization of targeted therapies particularly at early stages of myelomatogenesis.

Materials and methods

A complete workflow of data analysis starting from selection of MMRF CoMMpass IA12 dataset to generation of FISH plots showing patterns of clonal evolution and interpretations is summarized in Fig. 1.

Selection of dataset

In this study, we have used the WES data obtained from the MMRF CoMMpass database (IA12) (https://themmrf.org/finding-a-cure/our-work/the-mmrf-commpass-study/), which includes whole genome/ exome sequencing data of over 1000 newly diagnosed MM patients with

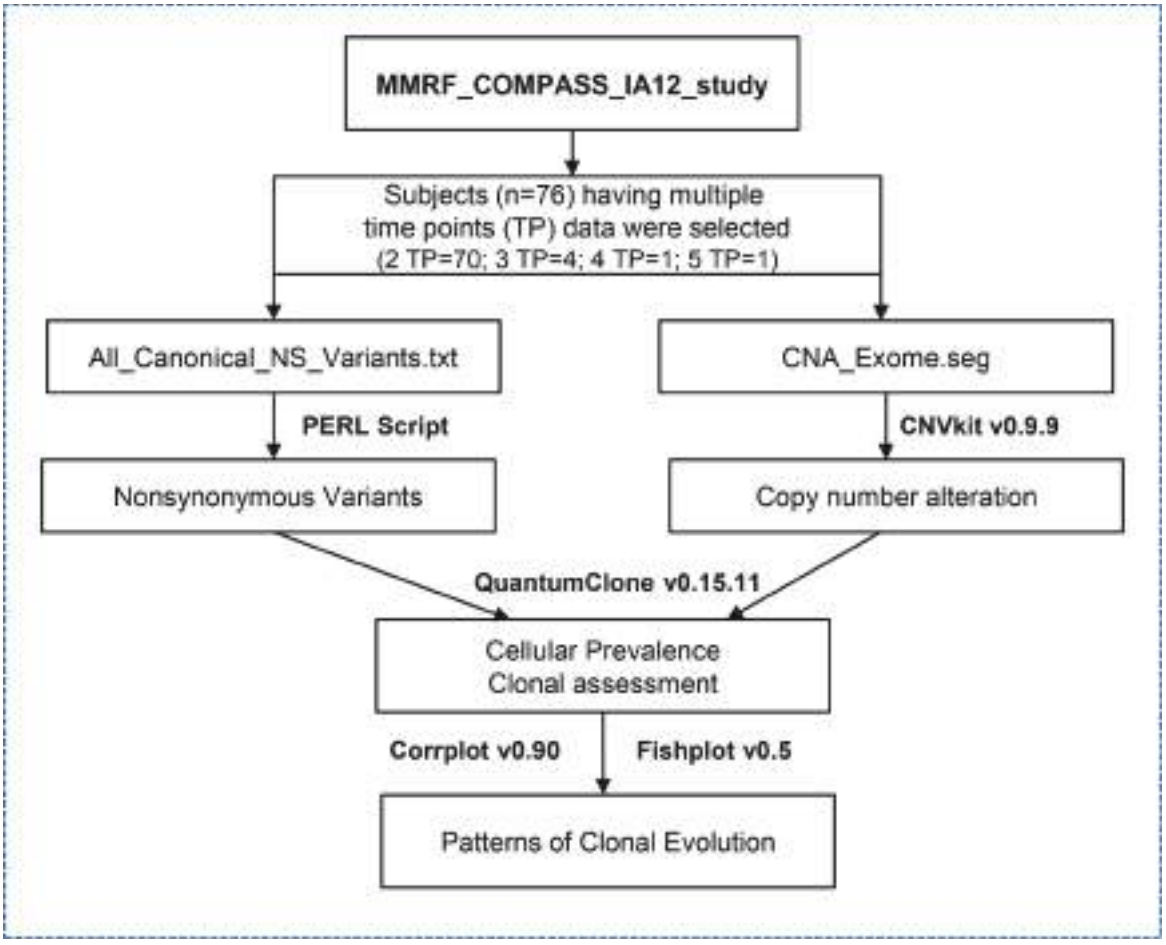


Fig. 1. Overall study design and workflow. Data on WES was mined for 76 MM patients sampled at multiple time points TP (ranging between 2 and 5) in the MMRF CoMMpass study IA12. Nonsynonymous mutations and copy number alterations were collated with PERL and CNVkit respectively. Data was converted into cellular prevalence with QuantumClone and analyzed further for mutation clustering and clonal assignments. Fishplot was used to deduce patterns of clonal evolution while corrplot was used to compare pairwise correlations among clonal shifts.

enriched tumor and matched constitutional samples. The variant file (MMRF_CoMMpass_IA12a_All_Canonical_NS_Variants.txt) and the copy number file (MMRF_CoMMpass_IA12a_CNA_Exome.seg) were used for this clonal evolution study. Clinical and demographic details were obtained from the Clinical_Flat_Files and are summarized in Table 1. The MM subjects ($n = 76$) having multiple time points data on Non Synonymous (NS) variants and CNAs (at 2 time points, $n = 70$; at 3 time points, $n = 4$; at 4 time points, $n = 1$; at 5 time points, $n = 1$) were selected and included in this study.

Evaluation of multiple time point whole exome data for cellular prevalence

We developed in house PERL (practical extraction and report language) scripts to parse single nucleotide variant / copy number variant (CNV) files for selected subjects at all timepoints. The CNVkit program (version 0.9.9; https://github.com/etal/cnvkit [40]) was used to extract absolute copy number from the copy number segment files (MMRF_CoMMpass_IA12a_CNA_Exome.seg). The output obtained from the CNVkit program was used to assign genotype information for each variant. The data obtained was further processed with QuantumClone [41] (https://www.rdocumentation.org/packages/QuantumClone/versions/0.15.11). Somatic variants along with their genotype information were used to infer the clonal progression in MM. The cellular prevalence values $\theta^{'}$ of each cluster were calculated as follows:

$$\widehat{\theta} = VAF \times \frac{N_{Ch} + N_{Ch(Normal)} \times \frac{1-p}{p}}{NC}$$

Table 1
Clinical, laboratory parameters and cytogenetic aberrations found in 76 MM patients.

Parameter	No. of patients	%
Median Age (Range) In Years	67 (35 to 93)	
Gender		
Male	43	56.6
Female	33	43.4
Hemoglobin (g/dL)		
≤10	76	100
>10	0	0
Platelet Count (/dL)		
<100	3	3.9
≥100	73	96.0
Serum creatinine (mg/dL)		
≤2	66	86.8
>2	10	13.2
Serum albumin (g/dL)		
<3.5	31	40.8
≥3.5	44	57.9
NA	1	1.3
ISS 1/2/3/NA	22/26/26/2	
RISS I/II/III/NA	8/38/6/24	
Serum calcium, mg/dL		
0–11	76	100
>11	0	0
IgG Isotype		
IgA	9	11.8
IgG	18	23.6
NA	49	64.5
BM plasma cells,%		
≤40	48	63.2
>40	6	7.9
NA	22	28.9
Serum LDH (IU/L)		
≤420	47	61.8
>420	13	17.1
NA	16	21.1
β2-microglobulin, mg/L		
<3.5	29	38.2
≥3.5	45	59.2
NA	2	2.6

NA = Not available.

where N_{Ch} is the number of copies of the corresponding locus in cancer cells, $N_{Ch(Normal)}$ is the number of copies of the corresponding locus in the normal cells ($N_{Ch(Normal)} = 2$ for autosomes), VAF is the Variant Allele Frequency and *NC* is the number of chromosomal copies bearing the variant and *p* is the tumor purity [41].

Assessment of patterns of clonal evolution

The values of cellular prevalence obtained from QuantumClone were subjected to Fish plot R package [42] in order to visualize the patterns of clonal evolution (https://github.com/chrisamiller/fishplot). Cellular prevalence values higher than 1 were set to 1 as suggested [41]. Clonal lineage was inferred by following sum rules and cross rules as described in earlier studies [43,44]. Patterns of evolution [45] were classified as branching if characterized both by the gain and loss of mutational clusters at relapse or as linear if characterized by the gain of mutations at relapse but no evidence of clonal loss. Stable evolution pattern was characterized by a preserved clonal structure at both the time points. Stable with loss patterns had a predominantly preserved clonal structure at relapse but with some loss of cluster of mutations.

Estimation of clonality and subclonality

Based on the value of estimated cellular prevalence percentage, gene variants were classified into four categories: Clonal (C, cellular prevalence > = 85%), High Sub Clonal (HSC, cellular prevalence > = 25% and < 85%), Low Sub Clonal (LSC cellular prevalence > = 0.6% and <25%), Very Low Sub clonal (VLSC, cellular prevalence <0.6%)] [46].

Identification of mutated genes and gene functions

Genes found to be mutated were classified as drivers/ oncogenes/ tumor suppressor/actionable genes as defined by intOgen [47] (https://www.intogen.org/search); OncoKB [48] (https://www.oncokb.org/); cBioPortal [49,50] (https://www.cbioportal.org/) and COSMIC actionability data v93 (https://cancer.sanger.ac.uk/cosmic).

Computation of pairwise correlations between subclones

Pearson correlation between pairwise subclonal mutations was evaluated and visualized with corrplot function of the corrplot package v0.90 (https://cran.r-project.org/web/packages/corrplot/index.html).

Analysis of biological pathways

The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways relating to mutations at TP1 and TP2 (time point 2) were deduced by gene enrichment approach using Enrichr (https://maayanlab.cloud/Enrichr/) as described [51] while MSigDB [52] was used to compare molecular signatures at both timepoints respectively.

Results

NS mutational spectrum at baseline compared to progression

A median of 60 NS mutations (range 2 to 414) and 65 NS variations (range 2 to 312) were identified per patient at baseline and on progression respectively. Supplementary Table ST1 lists 915, 1173 and 2444 genes that were found to have NS mutations in whole exomes of samples evaluated at TP1 or TP at progression (TP2) or in common at both the time points respectively.

Among these, 16 genes (ALK, ATR, AXL, BIRC3, BRIP1, CSF3R, DDR1, MERTK, MET, PTPN11, RAD52, RICTOR, RPTOR, TET2, U2AF1, XRCC3) and 14 genes (BLM, ETV6, FANCM, FLT4, JAK2, KMT2A, PIK3R1, POLE, RAD51C, RFC1, RFC2, SLX4, SMO, TSC1) were

actionable genes and were identified only at baseline or on progression respectively. A group of 32 actionable genes (APC, ASXL1, ATM, ATRX, BRAF, CBFB, CCND1, DAXX, DDR2, ERBB3, ERBB4, ERCC3, FGFR3, FGFR4, GNAQ, KIT, KRAS, MLH1, MLH3, MTOR, MYC, NRAS, NTRK2, NTRK3, PML, POLD1, PRKDC, RB1, SF3B1, SMARCA4, TEK, TP53) were found to be mutated at both the time points (Supplementary Table ST1).

Gains and losses of mutated genes in subclones with progression

A comparison of clonal gains or losses of mutated genes from baseline to progression in patients is shown in Supplementary Table ST2. The most commonly mutated actionable genes included NRAS followed by *KRAS, TP53, BRAF, GNAQ, NTRK3, APC, RB1* and *FGFR3* in descending order while mutations were also common in other genes such as *TTN, RRBP1, FAT3, DNAH5, MUC16, DIS3, USH2A* and *IRF4* (Supplementary Table ST2).

The genes that showed topmost number of clonal gains or losses with progression in MM patients are shown in Fig. 2a. The topmost frequent clonal gains were observed in *NRAS, RRBP1, TP53, FCGBP, SYNE1, KCNMA1, BRAF* and *MAGI3* while the most frequent clonal losses on progression were found among *ZFHX3, LAMA1, MASP1* and *PTPRF* (Fig. 2a). All these mutated genes were observed at comparable high subclonal levels except *RRBP1* that was found at low subclonal levels at both the time points (Fig. 2b). Heatmaps in Fig. 3 and Supplementary Fig. 1 depict frequencies of clonal losses or gains of mutated actionable and non actionable genes in each of the 76 MM patients respectively as well as their cytogenetic profiles.

Comparison of cellular prevalence of mutated genes at baseline with progression

Average%CP of mutated genes at baseline (TP1) and second TP of progression (TP2) were calculated with QuantumClone and are plotted in Fig. 4. Some of the genes that were clonal (average%CP > = 85%) at diagnosis dropped down to low subclonal levels (*LOXHD1*%CP = 24.871), or high subclonal levels (*ICOSLG*%CP = 28.972, *VCAN*%CP =

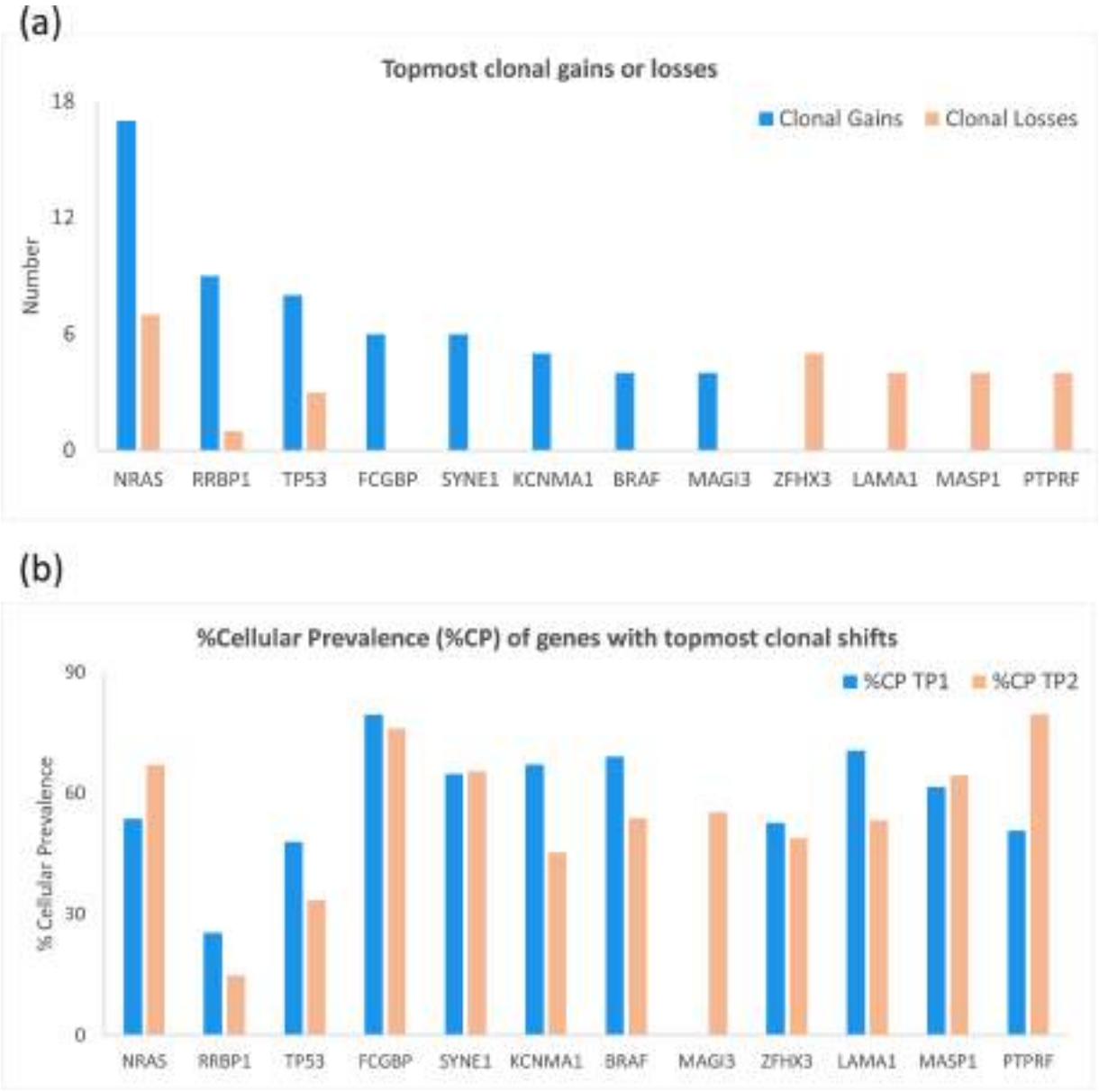


Fig. 2. Comparison of topmost clonal gains and losses associated with progression of MM. (a) Genes showing largest numbers of clonal gains and losses, and (b) parallel comparison of cellular prevalence of genes shown in (a).

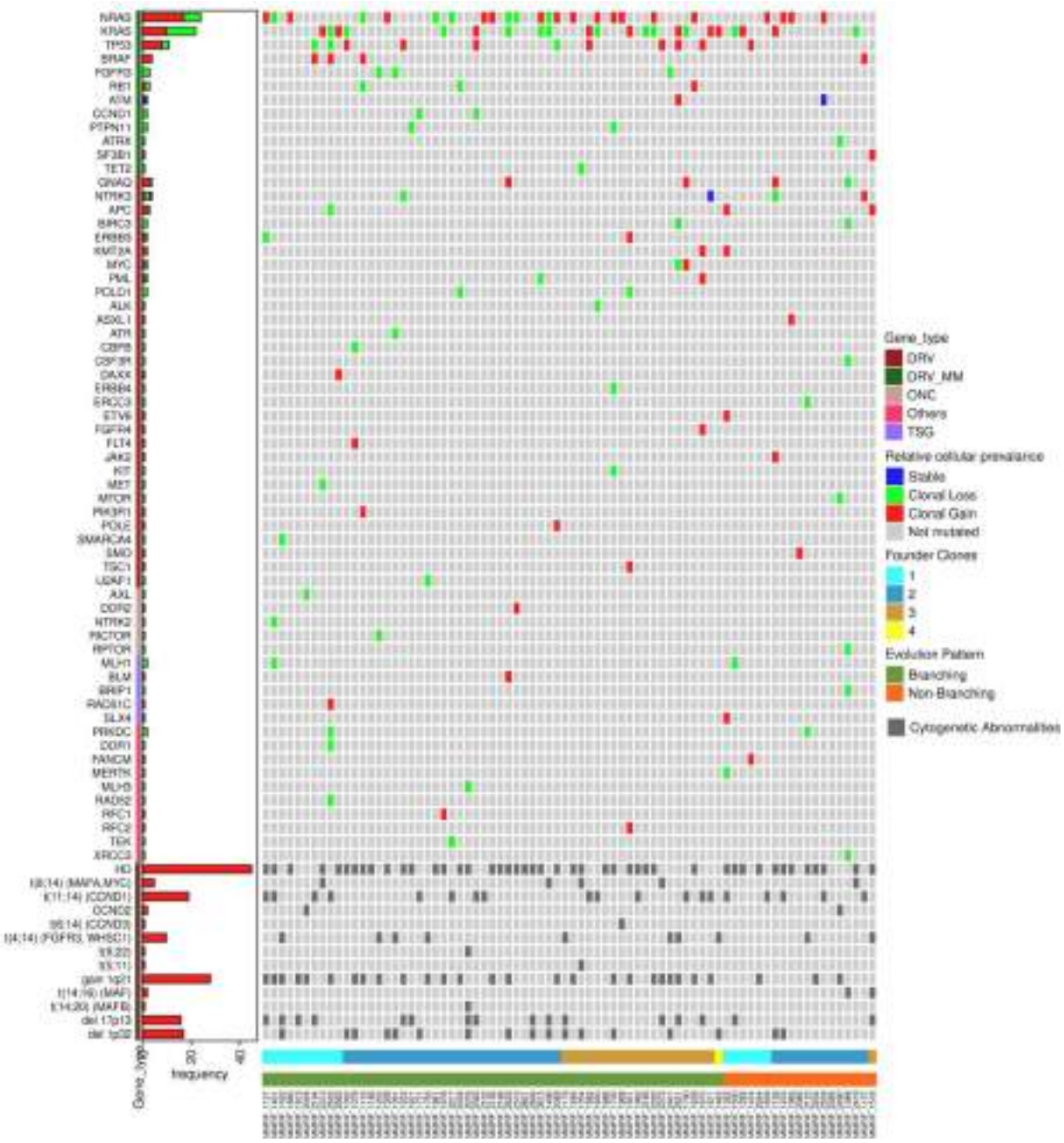


Fig. 3. Heatmap depicting distribution of clonal shifts in actionable genes and cytogenetic aberrations in 76 MM patients categorised on the basis of branching/ non branching evolution and number of founder clones. DRV = Driver, DRV_MM = Driver known for MM, ONC = Oncogene, TSG = Tumor suppressor gene.

47.3278, *DST*%CP = 41.394, *RYR2*%CP = 66.980, *ZNF462*%CP = 69.204, *GNAQ*%CP = 65.122, *EGR1*%CP = 66.450 and *TRRAP*%CP = 68.803) on progression (Supplementary Table ST2 and Fig. 4). Conversely, a rise in cellular prevalence with progression was also observed for other set of genes. The average%CP of *MAGI3, PCDHA3, COL6A5* and *TUSC3* at baseline were 0, 55.157, 61.372,79.871 and increased to 55.356, 92.269, 91.531 and 86.732% on progression respectively (Supplementary Table ST2 and Fig. 4).

Patterns of clonal evolution in MM

QuantumClone analysis revealed most of the patients had 2 founder clones (51.3%) followed by 3 (26.3%) and 1 (21.1%) while 4 founder clones were rarely observed only in 1.3% of the patients (Fig. 5a). Four different patterns of clonal evolution were identified (Fig. 5). Among them, branching pattern was the most predominant found in 75% patients, followed by Linear in 15.8%, Stable in 5.3% and stable with loss of clone in 3.9% patients.

Further casewise details on clonal heterogeneity, individual patterns of clonal evolution, trajectories of founder clones and subclones identified for each of the 76 patients are shown in Supplementary Fig. 2. This figure shows fish plots, evolution plots as well as the clinical, therapy and demographic details for each patient included in this study. Each subclone had unique cluster of mutations that evolved from baseline to progression, their detailed molecular profiles, cellular prevalence, functional implications and consequences are summarized for each sample in Supplementary Table ST3.

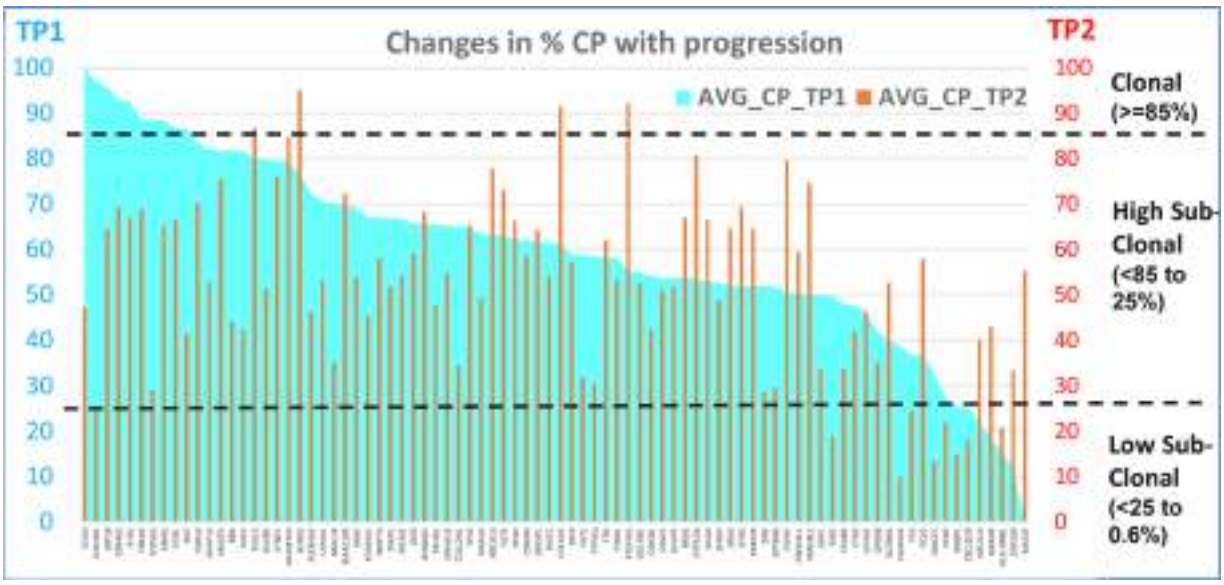


Fig. 4. A comparison of changes in average% Cellular prevalence (%CP) of mutated genes at two time points. Gene mutations were considered as clonal if the%CP was ≥ 85%, high subclonal if between <85 to 25% and low subclonal if between <25 to 0.6%.

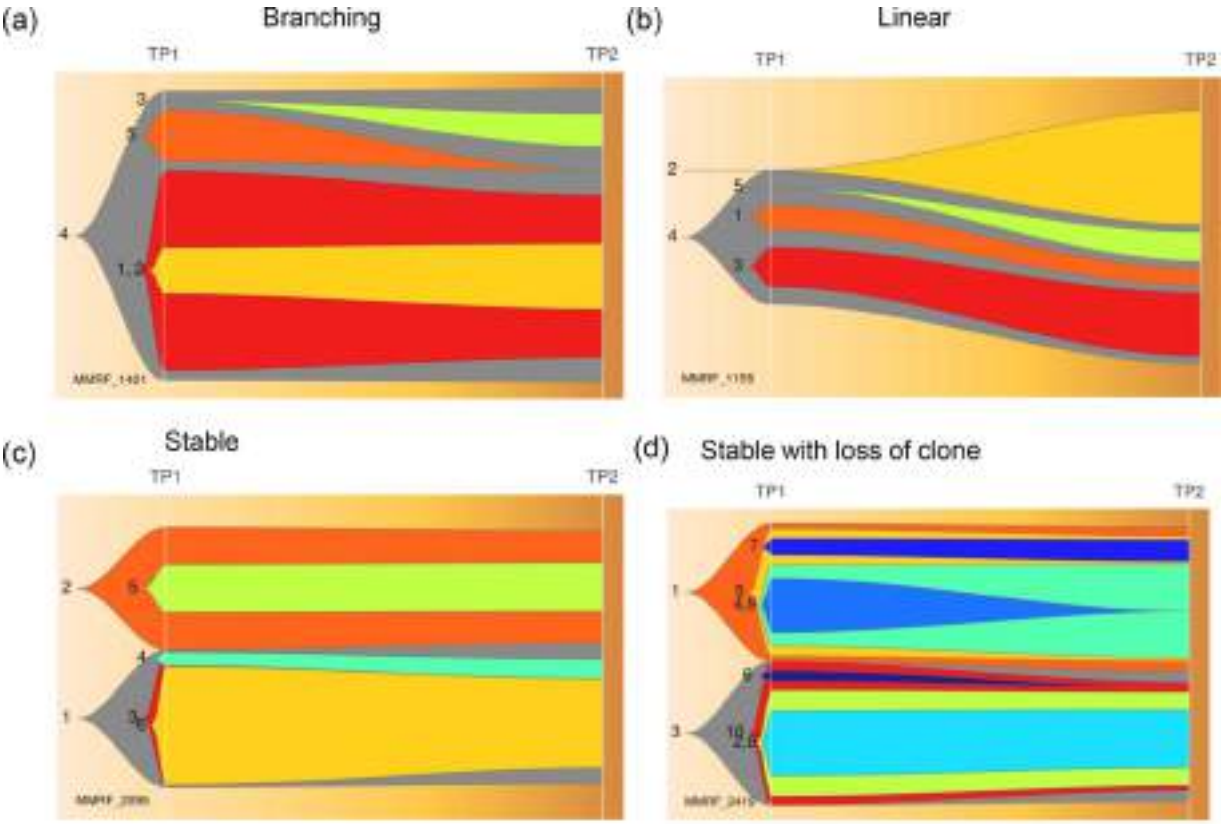


Fig. 5. Four types of clonal evolution patterns observed in MM patients. A representative Fishplot corresponding to each of the four patterns of clonal evolution is shown in (a) Branching, (b) Linear, (c) Stable, and (d) Stable with loss of clone.

Clonal sweeps and correlations of their co-occurrences

Fig. 6 shows a corrgram of individual and concomitant clonal shifts of pairwise mutated genes with disease progression in MM. Strength of Pearsons correlation coefficient r is depicted in lower triangle in Fig. 6 and in Supplementary Table ST4 while the significant p values are shown in upper triangle in Fig. 6 and in Supplementary Table ST5. Trajectories of clonal co-gains were identified in at least 3 patients for *NRAS*+*MAGI3* genes ($p = 0.023$, $r = 0.283$), *KRAS*+*TP53* ($p = 0.070$, $r = 0.228$), *TP53*+ *SYNE1* ($p = 0.003$, $r = 0.365$), while gain of *FCGBP* was observed along with reciprocal loss of *FAT3* ($p = 2.969\text{E-}10$, $r = 0.689$). Similar corrplots were also generated for patient subsets grouped on the basis of their cytogenetic aberrations (Fig. 6b–f) including hyperdiploid, del(1)p, 1q21+, del(17)p and t(11;14) respectively. Fig. 7a shows

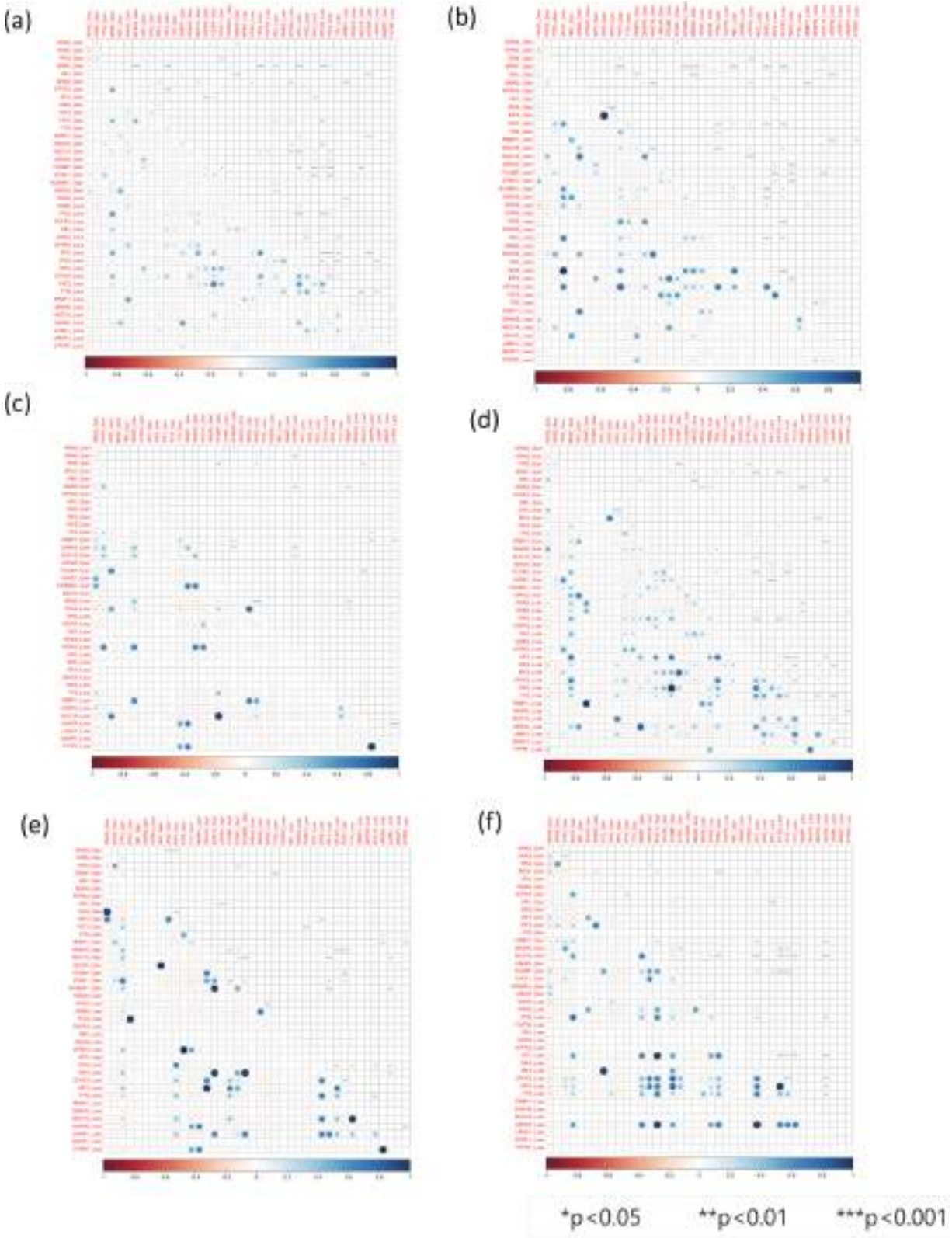


Fig. 6. Corrplot showing pairwise Pearson correlation coefficient r values for co-occurrences of clonal shifts in lower triangle and p values of significance in upper triangle. *$p<0.05$ **$p<0.01$ ***$p<0.001$. Corrplot (a) refers to total patient cohort; (b) patients with Hyperdiploidy (HD); (c) 1p del; (d) 1q gain; (e) 17p del; and (f) samples with t(11;14).

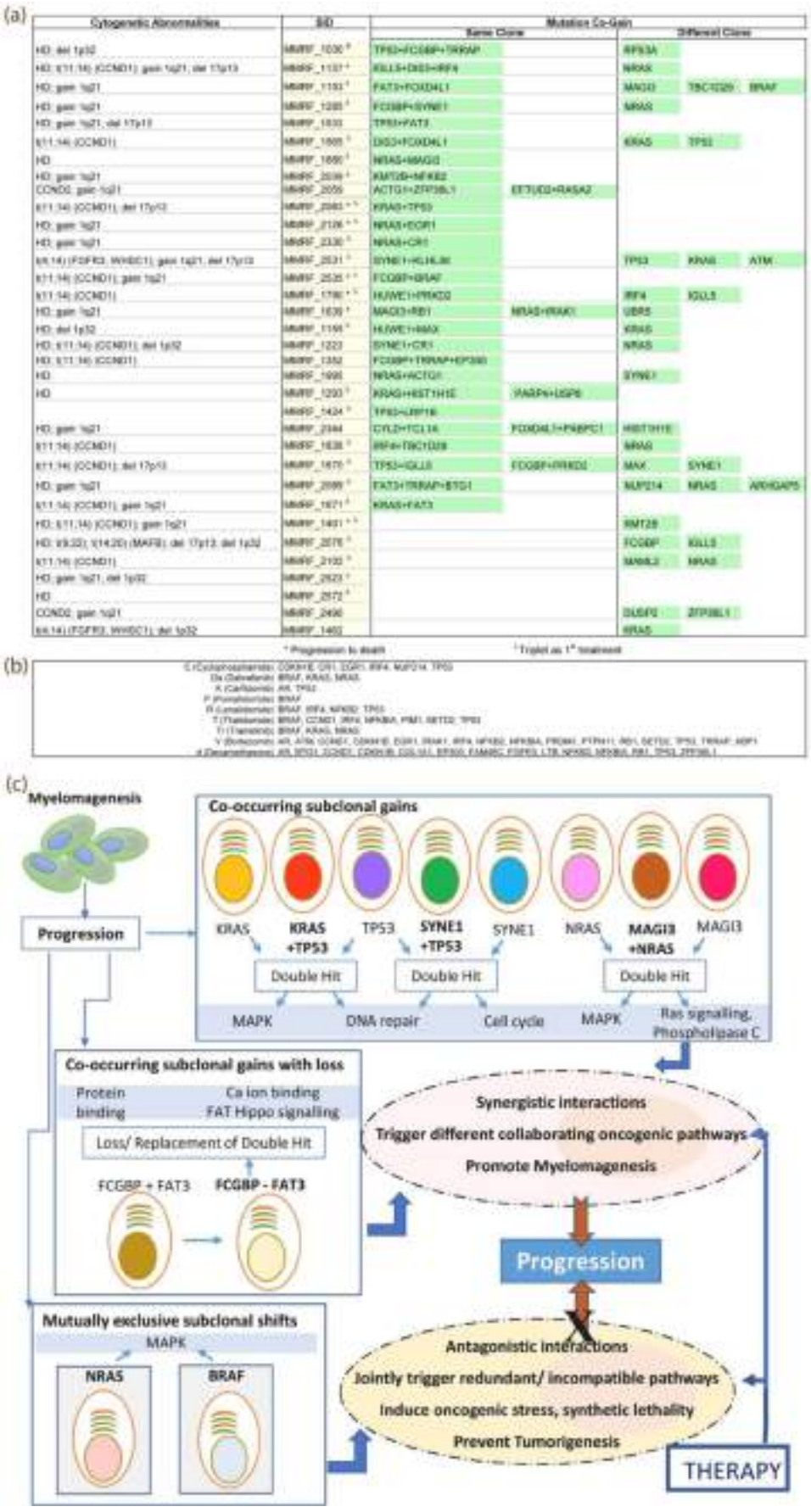


Fig. 7. Molecular profiles of mutational landscapes and 'Double hits' of subclonal shifts observed during progression of myelomatogenesis. (a) Depicts list of co-occurrences of mutations in same clones or in different clones in patients and the cytogenetic abnormalities, (b) Summary of drug pathway genes that may act as target genes or are involved in regulating drug sensitivities, (c) Cartoon showing three types of subclonal shifts (Co-occurring subclonal gains, Co-occurring subclonal gains with loss and mutually exclusive) that can be observed either within same clone or separate clones.. The co-occurring clonal shifts (Co-occurring subclonal gains or Co-occurring subclonal gains with loss) act like 'Double hits' and when present in same clone may promote progression by altering multiple pathways synergistically whereas mutually exclusive clonal shifts tend to result in antagonist synthetic lethal interactions that prevent tumorigenesis. The pathways triggered by such clonal shifts could be explored for development of novel and risk appropriate combinatorial therapies.

subclonal mutational landscapes of MM patients and co-occurring mutations in same or different subclones, their probable impact on drug sensitivities (Fig. 7b) and how the subclonal double hits/ loss of double hits and mutual exclusivities could impact myelomatogenesis (Fig. 7c).

Pathway enrichment analysis at two time points

The KEGG biological pathways predicted by Enrichr to be altered significantly in MM at baseline and on progression are shown in Supplementary Table ST6. The top 10 KEGG pathways significantly enriched at TP1 and TP2 are shown in Supplementary Fig. 3a and c respectively. These included ECM receptor interaction, Focal adhesion, PI3K.Akt signaling, calcium signaling and others.

The KEGG pathways that were uniquely perturbed at baseline included Adherens junction, Hippo signaling pathway, Lysine degradation and ATP-binding cassette (ABC) transporters while those altered at progression exclusively included MAPK signaling, autophagy, cellular senescence, apoptosis, protein digestion and others. Similarly, KEGG pathways that were predicted to be altered throughout baseline to progression included ECM receptor interaction, focal adhesion, Calcium signaling, PI3K-Akt signaling, Gap junction and others (Supplementary Table ST6).

Similarly, the top ten hallmark molecular signatures related to various cellular processes at baseline and progression are shown in Supplementary Table ST6 and Supplementary Fig. 3b and d respectively. The topmost significant MSigs included Epithelial mesenchymal transition, UV response downregulation, Mitotic spindle, Myogenesis and G2-M checkpoint at TP1 while Mitotic Spindle, Myogenesis, Epithelial mesenchymal transition, UV response downregulation, Apical junction, TNF alpha signaling and apoptosis at the time of progression.

Discussion

One of the major reasons for why MM remains undetectable at pre-malignant precursor stages and why it remains incurable is the extent of huge interpatient heterogeneity in genomic composition of malignant plasma cells and their differential abilities to respond to treatment. This highlights the need to identify and better understand clonal evolution of mutational landscapes during myelomatogenesis over time. This would help dissect out drivers and clinically relevant targets of action spanning earliest stages of MGUS/SMM to malignant MM or its relapse.

We have evaluated evolution of clonal mutational signatures with progression in MM. One strength of our study is that we have datamined WES of 76 MM patients that have been sampled at multiple time points in the largest ever genomic study on MM (MMRF CoMMpass). We have used a novel approach of comparing subclonal gains and losses across all the patients (Supplementary Tables ST2, ST3), and for the first time interpreted pairwise correlations of subclonal co-occurrences and exclusions with progression (Fig. 7). For example, in patient MMRF1137, a simultaneous co-gain of mutations in *DIS3* and *IRF4* was observed within the same clone which was also significant in the correlation plot in Fig. 6. This patient ultimately died from disease progression. This study has, therefore, provided a novel insight on how paired driver mutations may preferentially coexist and interact in the same subclones in patients with different cytogenetic aberrations. Their subclonal oncogenic dependencies may have a double hit impact, drive drug resistance and worsen outcomes. Analysis of differential cytogenetic subgroups revealed a few unique significant pairwise correlations (Fig. 6). These correlation profiles included *BRAF*+*FAT3, TP53*+*FAT3, DNAH5*+*GNAQ, BRAF*+*KCNMA1, MUC16*+*RRBP1* among patients with hyperdiploidy; *NRAS*+*DNAH5, RB1*+*FCGBP, USH2A*+*FCGBP* with 1q gain; *TP53*+*FCGBP, NRAS*+*KCNMA1, RRBP1*+*KCNMA1, DNAH5*+*KCNMA1* with 1pdel; *NRAS*+*DIS3, MUC16*+*KCNMA1* with 17pdel; and *BRAF*+*MUC16, RRBP1*+*MUC16, APC*+*FCGBP, FCGBP*+*MUC16* with t (11;14). Further studies will be needed in future to understand the biological and clinical impact of such subclonal correlations in MM.

Our study has revealed three categories of clonal shifts (Fig. 7c). Firstly, co-occurring gains of mutations in drivers such as *KRAS*+*TP53, SYNE1*+*TP53, MAGI3*+*NRAS* that may occur in different clones or within the same clone and result in collaborative oncogenic 'Double Hits'. The second category is represented by co-occurring subclonal gains with gradual loss among *FCGBP*+*FAT3* that may result in loss of 'Double hits' yet oncogenic. The third category includes mutual exclusion of driver genes e.g., *NRAS* and *BRAF* (Fig. 7c). In line with earlier studies [21], subclonal gains of 'double hits' of driver mutations may lead to synergistic interactions across different oncogenic pathways and promote tumorigenesis. Gradual loss of 'Double hits' may also trigger different collaborating oncogenic cascade of events and favor progression. On the contrary, mutual exclusion of other drivers such as *NRAS* and *BRAF* may jointly activate either redundant or incompatible divergent pathways resulting in oncogenic stress, synthetic lethality and prevent tumorigenesis [21]. Mutations in *BRAF, NRAS* and *KRAS* genes are frequently encountered in MM at frequencies of about 20%, 25% and 8% respectively. These are involved in the MAPK pathway [53], regulate proteasome assembly [54], and are potential targets for checkpoint therapy and IMiDs [55]. The mutual exclusion propensities of *NRAS* and *BRAF* are also well established in case of melanoma [56] and our study like others suggests it is operational in MM as well. However, these mutually exclusive genes may tend to co-emerge on treatment with RAF inhibitors. Co-operativity between *KRAS* and *TP53* has been reported in pancreatic cancer [57] and substantiates our findings in the context of MM.

The genes identified to be involved in synergistic or antagonistic interactive clonal 'double hits' leading to progression in MM in this study are key regulatory genes involved in oncogenesis and foster further investigations. Gene *SYNE1* (alias Nesprin1) Spectrin Repeat containing nuclear envelope protein1 is located on chromosome 6q25.2. It links actin cytoskeleton to organelles, maintains nuclear polarity, spindle orientation and subcellular cellular organization. It has a role in nucleotide binding, cell cycle, mitosis, meiosis and DNA damage response. It is known to be somatically mutated or silenced by methylation in breast, prostate, colon, lung cancers, head and neck squamous cell carcinoma suggesting its loss of function may promote tumorigenesis. It has been reported in an isolated case of MM with drug resistance to Bortezomib [58] and needs to be explored further. *MAGI3* gene (alias KIAA1634) (1p13.2) is a membrane associated guanylate kinase that cooperates with *PTEN* to modulate kinase activity of AKT1. It regulates JNK signaling, Ras signaling and phospholipase C pathways. Aberrations in *MAGI3* have been associated with ovarian cancer and *MAGI3-AKT3* fusions occur in breast cancer. It is a known level 3 presumptive evidence based gene implicated in MMSET group of myeloma [59]. *FAT3* (alias KIAA1989) (11q14.3) is a FAT atypical cadherin3 involved in FAT-Hippo signaling and calcium ion binding. It has been associated with TMB and poor prognosis in oesophageal carcinoma. It is found in 4–7% MM patients [60]. The *FCGBP* (Fc fragment of IgG binding protein) (19q13.2) gene has been implicated in melanoma and it forms an integral component of transcriptional networks in MM [61].

In addition to above, we have also analyzed genewise cellular prevalence during clonal shifts and shown presence of trajectories ranging from clonal to subclonal levels for genes such as *LOXHD1* and *ICOSLG* (Fig. 4). Another novel finding of our study is that we have been able to compare the actionable mutational blueprints of subclones at multiple time points in each patient (summarised in Supplementary Table ST3, Supplementary Fig. SF1) that provides a direction to eventually infer personalized genome prescribed combinatorial therapies in future. This study has analyzed association of subclonal mutational gains with probable drug resistance (Fig. 7b). For example, subclonal mutations in *BRAF, IRF4, NFKB2* and *TP53* may affect response to Lenalidomide. Similarly, mutations in *ATM, CCND1, RB1* etc. are involved in Bortezomib response pathway. Hence, regular monitoring of subclonal sweeps of driver mutations at diagnosis and at every relapse may help predict responsiveness to different drugs.

Parallel studies carried out on clonal evolution of MM by others also lend support to predominance of branching evolution [30,35,62] as observed in our analysis. A similar analysis of sequential WES of 10 patients from MGUS to MM highlighted that the transformed subclonal PC populations identified at MM are already present in asymptomatic MGUS/SMM stages and are thus reflective of clonal stability [36]. Another study [38] on WES of 19 patients enrolled in MMRF CoMMpass genomic project study showed that stable/resistant clones were characterized by concurrent 17pdel and 13qdel, and/or mutations in *NRAS, DIS3, FAM46C, ROBO1* and *CCND1*. A longitudinal follow-up of clonal trajectories from early initiation or time of diagnosis to malignant transformation / relapse may therefore allow identification of subclones that are stable/resistant or sensitive to line of treatment [36].

A recent study [37] has further proposed that alternate punctuated episodes of emergence of new subclones driven to dominance followed by their static expansion under positive selection continue across multiple time windows of myelomatogenesis. This process during early initiation prior to SMM leads to preservation of a clonal architecture with malignant potential that might expand and evolve further until clinical significance. It has been suggested that retention of a parallel architecture in MM as of SMM could be probed for better prognostication and early intervention [37].

Conclusions

The differences in oncogenic genes and associated pathway networks at two time points of evaluation reflect a continuum of clonal evolution that may be explored as an additional independent mechanism of drug resistance observed in clinical practice. Besides, these findings are important conceptually as they highlight the need for (1) molecular testing immediately before initiating a new therapy especially if aiming at a targeted or immune therapy, (2) estimation of cellular prevalence to predict the likelihood of the depth of response achievable with targeted therapy and rationalized approach on combining multiple therapeutics and (3) single cell mutation spectrum profiling at least for patients under evaluation in clinical trials with novel therapeutics and to explore the relevance of subclonal double hits.

Data availability

The IA12 datasets used for the analyses described in this work were downloaded from MMRF CoMMpass (Multiple Myeloma Research Foundation CoMMpass [SM] (Relating Clinical Outcomes in MM to Personal Assessment of Genetic Profile) study (www.themmrf.org)) researcher gateway.

Author contributions

RG conceived and designed the study, contributed in clinical analysis, data interpretation and wrote the manuscript. GK, LJ, AF and SK performed analysis, data interpretations and wrote the manuscript. AG coordinated computational analysis and wrote the manuscript.

Ethical information statement

This is a computational study that has analyzed publicly available online datasets generated by the Multiple Myeloma Research Foundation (MMRF) Personalized Medicine Initiatives (https://research.themmrf.org and www.themmrf.org). These data were downloaded directly from MMRF researcher gateway with due permission granted by the MMRF. Since no samples were collected from human/animal subjects in this study, ethical approval and informed consents are not applicable.

Conflicts of interest

The authors declare that there is no conflict of interests of any kind regarding the publication of this paper.

Acknowledgements

This work was supported by grants from Department of Biotechnology, Govt. of India [BT/ PR7136 /MED /30/901/2012; BT/MED/30/SP11006/2015] and Department of Science and Technology, Govt. of India [DST/ICPS/CPS-Individual/2018/279(G)]. Akanksha Farswan would like to thank University Grants Commission, Govt. of India for UGC-Senior Research Fellowship. Authors acknowledge Multiple Myeloma Research Foundation (MMRF) for providing the CoMMpass IA12 dataset. These data were generated as part of the Multiple Myeloma Research Foundation Personalized Medicine Initiatives (https://research.themmrf.org and www.themmrf.org). Dr. Anubha Gupta would also like to thank Infosys Center for Artificial Intelligence, IIIT-Delhi for support in their research.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tranon.2022.101472.

References

[1] NWCJ van de Donk, C. Pawlyn, K.L Yong, Multiple myeloma. The Lancet, Elsevier 397 (2021) 410–427, https://doi.org/10.1016/S0140-6736(21)00135-5.
[2] B.A. Walker, C.P. Wardell, L. Melchor, A. Brioli, D.C. Johnson, M.F. Kaiser, Intraclonal heterogeneity is a critical early event in the development of myeloma and precedes the development of clinical symptoms, Leukemia 28 (2014) 384–390, https://doi.org/10.1038/leu.2013.199.
[3] S.V. Rajkumar, Updated diagnostic criteria and staging system for multiple myeloma, Am. Soc. Clin. Oncol. Educ. Book Am. Soc. Clin. Oncol. Meet. 35 (2016) e418–e423, https://doi.org/10.14694/EDBK_159009.
[4] S. Manier, K.Z. Salem, J. Park, D.A. Landau, G. Getz, I.M. Ghobrial, Genomic complexity of multiple myeloma and its clinical implications, Nat. Rev. Clin. Oncol. 14 (2016) 100–113, https://doi.org/10.1038/nrclinonc.2016.122.
[5] S.K. Kumar, S.V. Rajkumar, The multiple myelomas — Current concepts in cytogenetic classification and therapy, in: Nat Rev Clin Oncol, 15, Nature Publishing Group, 2018, pp. 409–421, https://doi.org/10.1038/s41571-018-0018-y.
[6] O. Landgren, G.J. Morgan, Biologic frontiers in multiple myeloma: from biomarker identification to clinical practice, Clin Cancer Res. Am. Assoc. Cancer Res. 20 (2014) 804–813, https://doi.org/10.1158/1078-0432.CCR-13-2159.
[7] H. Kaufmann, J. Ackermann, C. Baldia, T. Nösslinger, R. Wieser, S. Seidl, V. Sagaster, H. Gisslinger, U. Jäger, M. Pfeilstöcker, C. Zielinski, J. Drach, Both IGH translocations and chromosome 13q deletions are early events in monoclonal gammopathy of undetermined significance and do not evolve during transition to multiple myeloma, Leukemia 18 (2004) 1879–1882, https://doi.org/10.1038/sj.leu.2403518.
[8] H. Avet-Loiseau, C. Li, F. Magrangeas, W. Gouraud, C. Charbonnel, J.-. L. Harousseau, M. Attal, G. Marit, C. Mathiot, T. Facon, P. Moreau, K.C. Anderson, L. Campion, et al., Prognostic significance of copy-number alterations in multiple myeloma, J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol. 27 (2009) 4585–4590, https://doi.org/10.1200/JCO.2008.20.6136.
[9] D.R. Carrasco, G. Tonon, Y. Huang, Y. Zhang, R. Sinha, B. Feng, J.P. Stewart, F. Zhan, D. Khatry, M. Protopopova, A. Protopopov, K. Sukhdeo, I. Hanamura, et al., High-resolution genomic profiles define distinct clinico-pathogenetic subgroups of multiple myeloma patients, Cancer Cell 9 (2006) 313–325, https://doi.org/10.1016/j.ccr.2006.03.019.
[10] N. Abdallah, S.V. Rajkumar, P. Greipp, P. Kapoor, M.A. Gertz, A. Dispenzieri, L. B. Baughn, M.Q. Lacy, S.R. Hayman, F.K. Buadi, D. Dingli, R.S. Go, Y.L. Hwa, et al., Cytogenetic abnormalities in multiple myeloma: association with disease characteristics and treatment response, Blood Cancer J. 10 (2020) 82, https://doi.org/10.1038/s41408-020-00348-5.
[11] CoMMpass Network TM A. Christofferson, S. Skerget, J. Aldrich, C. Legendre, S. Nasser, J. Yesil, D. Auclair, S. Lonial, J.J Keats, Unsupervised Clustering of DNA Copy Number Profiles Identifies a High-Risk Subtype of Hyperdiploid Multiple Myeloma: an Mmrf Commpass Analysis, Blood. Am. Soc. Hematol. 134 (2019), https://doi.org/10.1182/blood-2019-132152, 1805–1805.
[12] S. Skerget, A. Christofferson, S. Nasser, C. Legendre, J. Yesil, D. Auclair, S. Lonial, J. Keats, Identification and Molecular Characterization of High-Risk Multiple Myeloma Patients from the MMRF CoMMpass Study at Diagnosis and Progression, in: Clin Lymphoma Myeloma Leuk, 19, Elsevier, 2019, p. e7, https://doi.org/10.1016/j.clml.2019.09.009.

[13] Skerget S., Penaherrera D., Chari A., Jagannath S., Siegel D.S., Vij R., Orloff G., Jakubowiak A., Niesvizky R., Liles D., Berdeja J., Levy M., Wolf J., et al. Genomic Basis of Multiple Myeloma Subtypes from the MMRF CoMMpass Study [Internet]. 2021 Aug p. 2021.08.02.21261211. Available from https://www.medrxiv.org/content/10.1101/2021.08.02.21261211v1.

[14] O. Castaneda, R. Baz, Multiple Myeloma Genomics - A Concise Review, Acta Medica Acad. 48 (2019) 57–67, https://doi.org/10.5644/ama2006-124.242.

[15] J.R. Sawyer, The prognostic significance of cytogenetics and molecular profiling in multiple myeloma, Cancer Genet. 204 (2011) 3–12, https://doi.org/10.1016/j.cancergencyto.2010.11.002.

[16] N.C. Munshi, H. Avet-Loiseau, Genomics in multiple myeloma, Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res. 17 (2011) 1234–1242, https://doi.org/10.1158/1078-0432.CCR-10-1843.

[17] S. Robiou du Pont, A. Cleynen, C. Fontan, M. Attal, N. Munshi, J. Corre, H. Avet-Loiseau, Genomics of Multiple Myeloma, J. Clin. Oncol. 35 (2017) 963–967, https://doi.org/10.1200/jco.2016.70.6705.

[18] N. Bolli, G. Biancon, M. Moarii, S. Gimondi, Y. Li, C. de Philippis, F. Maura, V. Sathiaseelan, Y.-.T. Tai, L. Mudie, S. O'Meara, K. Raine, J.W. Teague, et al., Analysis of the genomic landscape of multiple myeloma highlights novel prognostic markers and disease subgroups, Leukemia (2017), https://doi.org/10.1038/leu.2017.344.

[19] M. Hultcrantz, Genomic Landscape and Mechanisms of Disease Evolution and Progression in Multiple Myeloma, Am. J. Hematol. Oncol. (2017) [Internet][cited 2021 Aug 30]; 13. Available from, https://www.gotoper.com/publications/ajho/2017/2017december/genomic-landscape-and-mechanisms-of-disease-evolution-and-progression-in-multiple-myeloma.

[20] A. Laganà, D. Perumal, D. Melnekoff, B. Readhead, B.A. Kidd, V. Leshchenko, P.-.Y. Kuo, J. Keats, M. DeRome, J. Yesil, D. Auclair, S. Lonial, A. Chari, et al., Integrative network analysis identifies novel drivers of pathogenesis and progression in newly diagnosed multiple myeloma, Leukemia 32 (2018) 120–130, https://doi.org/10.1038/leu.2017.197.

[21] G. El Tekle, T. Bernasocchi, A.M. Unni, F. Bertoni, D. Rossi, M.A. Rubin, J.-.P. Theurillat, Co-occurrence and mutual exclusivity: what cross-cancer mutation patterns can tell us, Trends Cancer 7 (2021) 823–836, https://doi.org/10.1016/j.trecan.2021.04.009.

[22] B.A. Walker, K. Mavrommatis, C.P. Wardell, T.C. Ashby, M. Bauer, F.E. Davies, A. Rosenthal, H. Wang, P. Qu, A. Hoering, M. Samur, F. Towfic, M. Ortiz, et al., Identification of novel mutational drivers reveals oncogene dependencies in multiple myeloma, Blood 132 (2018) 587–597, https://doi.org/10.1182/blood-2018-03-840132.

[23] F. Maura, E.M. Boyle, E.H. Rustad, C. Ashby, D. Kaminetzky, B. Bruno, M. Braunstein, M. Bauer, P. Blaney, Y. Wang, H. Ghamlouch, L. Williams, J. Stoeckle, et al., Chromothripsis as a pathogenic driver of multiple myeloma, Semin. Cell Dev. Biol. (2021), https://doi.org/10.1016/j.semcdb.2021.04.014 [Internet][cited 2021 Aug 30];.

[24] G. Kaur, R. Gupta, N. Mathur, L. Rani, L. Kumar, A. Sharma, V. Singh, A. Gupta, O. D Sharma, Clinical impact of chromothriptic complex chromosomal rearrangements in newly diagnosed multiple myeloma, Leuk. Res. 76 (2019) 58–64, https://doi.org/10.1016/j.leukres.2018.12.005.

[25] T. Perini, M. Materozzi, E. Milan, The Immunity-malignancy equilibrium in multiple myeloma: lessons from oncogenic events in plasma cells, FEBS J. [Internet] (2021), https://doi.org/10.1111/febs.16068 [citedAug 30]; n/a.

[26] R. Liu, Q. Gao, S.M. Foltz, J.S. Fowles, L. Yao, J.T. Wang, S. Cao, H. Sun, M. C. Wendl, S. Sethuraman, A. Weerasinghe, M.P. Rettig, E.P. Storrs, et al., Co-evolution of tumor and immune cells during progression of multiple myeloma, Nat. Commun. 12 (2021) 2559, https://doi.org/10.1038/s41467-021-22804-x.

[27] P.H. Hoang, R.S. Houlston, Multiple mechanisms can disrupt oncogenic pathways in multiple myeloma, Oncotarget 9 (2018) 35801–35802, https://doi.org/10.18632/oncotarget.26301.

[28] E.M. Boyle, S. Deshpande, R. Tytarenko, C. Ashby, Y. Wang, M.A. Bauer, S. K. Johnson, C.P. Wardell, S. Thanendrarajan, M. Zangari, T. Facon, C. Dumontet, B. Barlogie, et al., The molecular make up of smoldering myeloma highlights the evolutionary pathways leading to multiple myeloma, Nat. Commun. 12 (2021), https://doi.org/10.1038/s41467-020-20524-2 [Internet]. 2021 [citedJun 10].

[29] L. John, M.T. Krauth, K. Podar, M.-.S. Raab, Pathway-Directed Therapy in Multiple Myeloma, Cancers. Multidisciplinary Digital Publishing Institute; 13 (2021) 1668, https://doi.org/10.3390/cancers13071668.

[30] N. Bolli, H. Avet-Loiseau, D.C. Wedge, P. Van Loo, L.B. Alexandrov, I. Martincorena, K.J. Dawson, F. Iorio, S. Nik-Zainal, G.R. Bignell, J.W. Hinton, Y. Li, J.M.C. Tubio, et al., Heterogeneity of genomic evolution and mutational profiles in multiple myeloma, Nat. Commun. 5 (2014) 2997, https://doi.org/10.1038/ncomms3997.

[31] J.G. Lohr, P. Stojanov, S.L. Carter, P. Cruz-Gordillo, M.S. Lawrence, D. Auclair, C. Sougnez, B. Knoechel, J. Gould, G. Saksena, K. Cibulskis, A. McKenna, M. A. Chapman, et al., Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy, Cancer Cell 25 (2014) 91–101, https://doi.org/10.1016/j.ccr.2013.12.015.

[32] B.A. Walker, E.M. Boyle, C.P. Wardell, A. Murison, D.B. Begum, N.M. Dahir, P. Z. Proszek, D.C. Johnson, M.F. Kaiser, L. Melchor, L.I. Aronson, M. Scales, C. Pawlyn, et al., Mutational spectrum, copy number changes, and outcome: results of a sequencing study of patients with newly diagnosed myeloma, J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol. 33 (2015) 3911–3920, https://doi.org/10.1200/JCO.2014.59.1503.

[33] M.C. Da Vià, B. Ziccheddu, A. Maeda, F. Bagnoli, G. Perrone, N. Bolli, A Journey Through Myeloma Evolution: from the Normal Plasma Cell to Disease Complexity, HemaSphere 4 (2020) e502, https://doi.org/10.1097/HS9.0000000000000502.

[34] M. Ho, A. Patel, C.Y. Goh, M. Moscvin, L. Zhang, G. Bianchi, Changing paradigms in diagnosis and treatment of monoclonal gammopathy of undetermined significance (MGUS) and smoldering multiple myeloma (SMM), Leukemia 34 (2020) 3111–3125, https://doi.org/10.1038/s41375-020-01051-x.

[35] N. Bolli, F. Maura, S. Minvielle, D. Gloznik, R. Szalat, A. Fullam, I. Martincorena, K. J. Dawson, M.K. Samur, J. Zamora, P. Tarpey, H. Davies, M. Fulciniti, et al., Genomic patterns of progression in smoldering multiple myeloma, Nat. Commun. 9 (2018) 3363, https://doi.org/10.1038/s41467-018-05058-y.

[36] A.K. Dutta, J.L. Fink, J.P. Grady, G.J. Morgan, C.G. Mullighan, L.B. To, D. R. Hewett, A.C.W. Zannettino, Subclonal evolution in disease progression from MGUS/SMM to multiple myeloma is characterised by clonal stability, Leukemia 33 (2019) 457–468, https://doi.org/10.1038/s41375-018-0206-x.

[37] B. Diamond, V. Yellapantula, E.H. Rustad, K.H. Maclachlan, M. Mayerhoefer, M. Kaiser, G. Morgan, O. Landgren, F. Maura, Positive selection as the unifying force for clonal evolution in multiple myeloma, Leukemia 35 (2021) 1511–1515, https://doi.org/10.1038/s41375-021-01130-7.

[38] A. Lagana, D. Melnekoff, I. Beno, V. Leshchenko, D. Perumal, J.J. Keats, M. DeRome, J. Yesil, D. Auclair, D. Madduri, A. Chari, H.J. Cho, B. Barlogie, et al., Clonal Evolution in Newly Diagnosed Multiple Myeloma Patients: a Follow-up Study from the Mmrf Commpass Genomics Project, Blood 130 (2017) 325, https://doi.org/10.1182/blood.V130.Suppl_1.325.325.

[39] F. Maura, E.H. Rustad, E.M. Boyle, G.J. Morgan, Reconstructing the evolutionary history of multiple myeloma, Best Pract. Res. Clin. Haematol. 33 (2020), 101145, https://doi.org/10.1016/j.beha.2020.101145.

[40] E. Talevich, A.H. Shain, T. Botton, Bastian BC. CNVkit, Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing, PLoS Comput. Biol. 12 (2016), e1004873, https://doi.org/10.1371/journal.pcbi.1004873.

[41] P. Deveau, L. Colmet Daage, D. Oldridge, V. Bernard, A. Bellini, M. Chicard, N. Clement, E. Lapouble, V. Combaret, A. Boland, V. Meyer, J.-.F. Deleuze, I. Janoueix-Lerosey, et al., QuantumClone: clonal assessment of functional mutations in cancer based on a genotype-aware method for clonal reconstruction, Bioinformatics. Oxford Acad. 34 (2018) 1808–1816, https://doi.org/10.1093/bioinformatics/bty016.

[42] M. Ca, M. J, D. Hx, M. Ca, D. L, L. Tj, M. Er, W Rk, Visualizing tumor evolution with the fishplot package for R, BMC Genomics [Internet]. BMC Genom. 17 (2016), https://doi.org/10.1186/s12864-016-3195-z [cited 2021 Aug 25].

[43] W. Jiao, S. Vembu, A.G. Deshwar, L. Stein, Q. Morris, Inferring clonal evolution of tumors from single nucleotide somatic mutations, BMC Bioinformatics 15 (2014) 35, https://doi.org/10.1186/1471-2105-15-35.

[44] X. Ma, M. Edmonson, D. Yergeau, D.M. Muzny, O.A. Hampton, M. Rusch, G. Song, J. Easton, R.C. Harvey, D.A. Wheeler, J. Ma, H. Doddapaneni, B. Vadodaria, et al., Rise and fall of subclones from diagnosis to relapse in pediatric B-acute lymphoblastic leukaemia, Nat. Commun. 6 (2015) 6604, https://doi.org/10.1038/ncomms7604.

[45] J.R. Jones, N. Weinhold, C. Ashby, B.A. Walker, C. Wardell, C. Pawlyn, L. Rasche, L. Melchor, D.A. Cairns, W.M. Gregory, D. Johnson, D.B. Begum, S. Ellis, et al., Clonal evolution in myeloma: the impact of maintenance lenalidomide and depth of response on the genetics and sub-clonal structure of relapsed disease in uniformly treated newly diagnosed patients, Haematologica 104 (2019) 1440–1450, https://doi.org/10.3324/haematol.2018.202200.

[46] F. Nadeu, G. Clot, J. Delgado, D. Martín-García, T. Baumann, I. Salaverria, S. Beà, M. Pinyol, P. Jares, A. Navarro, H. Suárez-Cisneros, M. Aymerich, M. Rozman, et al., Clinical impact of the subclonal architecture and mutational complexity in chronic lymphocytic leukemia, Leukemia 32 (2018) 645–653, https://doi.org/10.1038/leu.2017.291.

[47] A. Gonzalez-Perez, C. Perez-Llamas, J. Deu-Pons, D. Tamborero, M.P. Schroeder, A. Jene-Sanz, A. Santos, N Lopez-Bigas, IntOGen-mutations identifies cancer drivers across tumor types, Nat. Methods 10 (2013) 1081–1082, https://doi.org/10.1038/nmeth.2642.

[48] D. Chakravarty, J. Gao, S.M. Phillips, R. Kundra, H. Zhang, J. Wang, J.E. Rudolph, R. Yaeger, T. Soumerai, M.H. Nissan, M.T. Chang, S. Chandarlapaty, T.A. Traina, et al., OncoKB: a precision oncology knowledge base, JCO Precis. Oncol. (2017), https://doi.org/10.1200/PO.17.00011, 2017PO.17.00011.

[49] J. Gao, B.A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S.O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, E. Cerami, C. Sander, N. Schultz, Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal, Sci. Signal. 6 (2013) pl1, https://doi.org/10.1126/scisignal.2004088.

[50] E. Cerami, J. Gao, U. Dogrusoz, B.E. Gross, S.O. Sumer, B.A. Aksoy, A. Jacobsen, C. J. Byrne, M.L. Heuer, E. Larsson, Y. Antipin, B. Reva, A.P. Goldberg, et al., The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data, Cancer Discov. 2 (2012) 401–404, https://doi.org/10.1158/2159-8290.CD-12-0095.

[51] M.V. Kuleshov, M.R. Jones, A.D. Rouillard, N.F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S.L. Jenkins, K.M. Jagodnik, A. Lachmann, M.G. McDermott, C. D. Monteiro, G.W. Gundersen, et al., Enrichr: a comprehensive gene set enrichment analysis web server 2016 update, Nucleic Acids Res. 44 (2016) W90–W97, https://doi.org/10.1093/nar/gkw377.

[52] A. Liberzon, C. Birger, H. Thorvaldsdóttir, M. Ghandi, J.P. Mesirov, P. Tamayo, The Molecular Signatures Database (MSigDB) hallmark gene set collection, Cell Syst. 1 (2015) 417–425, https://doi.org/10.1016/j.cels.2015.12.004.

[53] M. Lionetti, M. Barbieri, K. Todoerti, L. Agnelli, S. Marzorati, S. Fabris, G. Ciceri, S. Galletti, G. Milesi, M. Manzoni, M. Mazzoni, A. Greco, G. Tonon, et al., Molecular spectrum of BRAF, NRAS and KRAS gene mutations in plasma cell dyscrasias: implication for MEK-ERK pathway activation, Oncotarget. Impact J. 6 (2015) 24205–24217, https://doi.org/10.18632/oncotarget.4434.

[54] F. Shirazi, R.J. Jones, R.K. Singh, J. Zou, I. Kuiatse, Z. Berkova, H. Wang, H.C. Lee, S. Hong, L. Dick, N. Chattopadhyay, R.Z. Orlowski, Activating KRAS, NRAS, and BRAF mutants enhance proteasome capacity and reduce endoplasmic reticulum stress in multiple myeloma, Proc. Natl. Acad. Sci. Proc. Natl. Acad. Sci. 117 (2020) 20004–20014, https://doi.org/10.1073/pnas.2005052117.

[55] S. Pasca, C. Tomuleasa, P. Teodorescu, G. Ghiaur, D. Dima, V. Moisoiu, C. Berce, C. Stefan, A. Ciechanover, H Einsele, KRAS/NRAS/BRAF mutations as potential targets in multiple myeloma, Front. Oncol. 9 (2019) 1137, https://doi.org/10.3389/fonc.2019.01137.

[56] C. Petti, A. Molla, C. Vegetti, S. Ferrone, A. Anichini, M. Sensi, Coexpression of NRASQ61R and BRAFV600E in human melanoma cells activates senescence and increases susceptibility to cell-mediated cytotoxicity, Cancer Res. 66 (2006) 6503–6511, https://doi.org/10.1158/0008-5472.CAN-05-4671.

[57] M.P. Kim, X. Li, J. Deng, Y. Zhang, B. Dai, K.L. Allton, T.G. Hughes, C. Siangco, J. J. Augustine, Y. Kang, J.M. McDaniel, S. Xiong, E.J. Koay, et al., Oncogenic KRAS recruits an expansive transcriptional network through mutant p53 to drive pancreatic cancer metastasis, Cancer Discov. Am. Assoc. Cancer Res. 11 (2021) 2094–2111, https://doi.org/10.1158/2159-8290.CD-20-1228.

[58] H. Ikeda, K. Ishiguro, Y. Aoki, T. Igarashi, T. Hayashi, T. Ishida, Y. Sasaki, T. Tokino, Y. Shinomura, Molecular diagnostics of a single drug-resistant multiple myeloma case using targeted next-generation sequencing, OncoTargets Ther. (2015) 2805, https://doi.org/10.2147/ott.s86515.

[59] Alberge J.-.B., Magrangeas F., Wagner M., Denié S., Guérin-Charbonnel C., Campion L., Attal M., Avet-Loiseau H., Carell T., Moreau P., Minvielle S., Sérandour A.A. DNA hydroxymethylation reveals transcription regulation networks and prognostic signatures in multiple myeloma [Internet]. 2019 Oct p. 806133. Available from https://www.biorxiv.org/content/10.1101/806133v1.

[60] I.J. Cardona-Benavides, C. de Ramón, N.C. Gutiérrez, Genetic abnormalities in multiple myeloma: prognostic and therapeutic implications, Cells 10 (2021) 336, https://doi.org/10.3390/cells10020336.

[61] L. Agnelli, M. Forcato, F. Ferrari, G. Tuana, K. Todoerti, B.A. Walker, G.J. Morgan, L. Lombardi, S. Bicciato, A. Neri, The reconstruction of transcriptional networks reveals critical genes with implications for clinical outcome of multiple myeloma, Clin Cancer Res. Am. Assoc. Cancer Res. 17 (2011) 7402–7412, https://doi.org/10.1158/1078-0432.CCR-11-0596.

[62] Y. Furukawa, J. Kikuchi, Molecular pathogenesis of multiple myeloma, Int. J. Clin. Oncol. 20 (2015) 413–422, https://doi.org/10.1007/s10147-015-0837-0.

Am J Cancer Res 2023;13(4):1155-1187
www.ajcr.us /ISSN:2156-6976/ajcr0148190

Original Article

BDL-SP: A Bio-inspired DL model for the identification of altered Signaling Pathways in Multiple Myeloma using WES data

Vivek Ruhela[1,2], Lingaraja Jena[3], Gurvinder Kaur[3], Ritu Gupta[3], Anubha Gupta[2]

[1]Department of Computational Biology, Indraprastha Institute of Information Technology-Delhi (IIIT-D), New Delhi, India; [2]SBI Lab, Department of Electronics and Communication Engineering & Centre of Excellence in Healthcare, Indraprastha Institute of Information Technology-Delhi (IIIT-D), New Delhi, India; [3]Laboratory Oncology Unit, Dr. B.R.A. IRCH, All India Institute of Medical Sciences (AIIMS), New Delhi, India

Received December 2, 2022; Accepted March 23, 2023; Epub April 15, 2023; Published April 30, 2023

Abstract: Identification of the genomic features responsible for the progression of Multiple Myeloma (MM) cancer from its precancerous stage MGUS can improve the understanding of the disease pathogenesis and, in devising suitable preventive and treatment measures. We have designed an innovative AI-based model, namely, the Bio-inspired Deep Learning architecture for the identification of altered Signaling Pathways (BDL-SP) to discover pivotal genomic biomarkers that can potentially distinguish MM from MGUS. The proposed BDL-SP model comprehends gene-gene interactions using the PPI network and analyzes genomic features using a deep learning (DL) architecture to identify significantly altered genes and signaling pathways in MM and MGUS. For this, whole exome sequencing data of 1174 MM and 61 MGUS patients were analyzed. In the quantitative benchmarking with the other popular machine learning models, BDL-SP performed almost similar to the two other best performing predictive ML models of Random Forest and CatBoost. However, an extensive post-hoc explainability analysis, capturing the application specific nuances, clearly established the significance of the BDL-SP model. This analysis revealed that BDL-SP identified a maximum number of previously reported oncogenes, tumor-suppressor genes, and actionable genes of high relevance in MM as the top significantly altered genes. Further, the post-hoc analysis revealed a significant contribution of the total number of single nucleotide variants (SNVs) and genomic features associated with synonymous SNVs in disease stage classification. Finally, the pathway enrichment analysis of the top significantly altered genes showed that many cancer pathways are selectively and significantly dysregulated in MM compared to its precursor stage of MGUS, while a few that lost their significance with disease progression from MGUS to MM were actually related to the other disease types. These observations may pave the way for appropriate therapeutic interventions to halt the progression to overt MM in the future.

Keywords: AI in cancer, haematological malignancy, multiple myeloma, MGUS, genomic aberrations, ShAP

Introduction

Multiple Myeloma (MM) is a neoplasm of malignant plasma cells in the bone marrow, preceded by the precancerous stage of Monoclonal Gammopathy of Undetermined Significance (MGUS). Presently, the distinction between MM and its precursor states (MGUS and smoldering multiple myeloma (SMM)) is based on the clinical symptoms and disease load including the percentage of aberrant plasma cells in the bone marrow, levels of monoclonal protein secreted by the aberrant plasma cells, and the extent of dysregulation of normal homeostasis. However, in clinical practice, distinction between different stages is at times ambiguous. The role of an early treatment and the type of such treatment to prevent progression to MM or to reduce the associated morbidity is also not clear. Although survival in MM has improved notably over the last few years, myeloma remains an incurable disease with an overall median survival of 2 to 10 years, depending on the response to the treatment. Thus, it would be interesting to decipher genes, genomic biomarkers and crucial pathogenic prognostic fac-

tors that are representative of MGUS and MM in order to develop appropriate therapeutic interventions to halt the progression to overt MM.

Multiple studies involving exome data have been performed to understand the genomic abnormalities driving tumor progression in MM. Exome data analysis of MM patients has revealed that the primary events in MM are either hyperdiploidy, i.e., trisomy of chromosomes 3, 5, 7, 9, 11, 15, 17 and/or 21, or non-hyperdiploidy involving translocations affecting the genes encoding immunoglobulin (Ig) heavy chains (IGH)-mainly t(4;14), t(6;14), t(11;14), t(14;16), and t(14;20) [1]. Primary events are then followed by multiple secondary events that are secondary translocations: t(8;14) linked with *MYC* overexpression, loss of heterozygosity, copy number variations (CNV), acquired mutations, and epigenetic modifications [1], contributing to tumorigenesis. Initial deep sequencing studies on 38 whole-genome sequencing (WGS) and 23 whole-exome sequencing (WES) MM patients revealed frequent mutations in NF-kB signaling pathway and activating mutations in the oncogene *BRAF* [2]. In another study based on the WES data of 84 MM patients, *SP140, LTB, ROBO1*, and *EGR1* genes were identified as the novel drivers of MM [3]. Similarly, the analysis of 463 WES data of MM patients revealed 15 recurrently mutated genes: *IRF4, KRAS, NRAS, MAX, HIST1H1E, RB1, EGR1, TP53, TRAF3, FAM46C, DIS3, BRAF, LTB, CYLD, and FGFR3* [4]. Further, the analysis of same 463 MM samples reported *RAS* and *NF-Kappa-B* pathways as most altered signaling pathways. Furthermore, the same study reported that the mutations in *CCND1* and DNA repair pathway genes-*TP53, ATM*, and *ATR*, adversely impacted the overall survival, while the alterations in *IRF4* and *EGR1* were associated with a favorable overall survival.

Another study on the exome data analysis of 203 MM patients demonstrated tumor heterogeneity with subclonal pattern of mutations and multiple mutations within the same pathway in the same patient [5]. A recent study on 62 newly diagnosed MM (NDMM) patients reported the association of changes in the cellular prevalence of mutations with disease progression [6]. Another study explored oncogenic dependencies between mutations in driver genes, hyperdiploidy events, primary translocations, and copy number alterations in MM patients [7]. Associations were established between t(4;14) and mutations in *FGFR3*, *DIS3*, and *PRKD2*; t(11;14) and mutations in *CCND1* and *IRF4*; t(14;16) and mutations in *MAF, BRAF, DIS3*, and *ATM*; and hyperdiploidy with gain 11q and mutations in *FAM46C*, and *MYC* rearrangements [7]. A recent study demonstrated the co-occurrence of mutations within the same or a different clone and the clonal shifts in the co-occurring and mutually exclusive mutations with progression in MM [8]. Similar phenomena may be occurring from the stage of MGUS to overt MM and require to be evaluated. Analysis of WES data of unpaired samples of MGUS and MM has been carried out by several groups [9-12]. These studies have demonstrated a less complex genomic architecture in MGUS compared to MM with fewer mutations and lower TMB in MGUS. In a landmark study, the analysis of paired samples of MGUS and MM reaffirmed the clonal heterogeneity and presence of majority of genomic changes at MGUS stage [13]. The existence of the majority of genomic abnormalities seen in MM at the MGUS stage poses a challenge in distinguishing MM from MGUS based on the genomic signatures and in defining critical genomic events responsible for the progression of MGUS to MM [9-13].

The early diagnosis of MM and the identification of relevant differentiating genomic biomarkers between MGUS and MM present several challenges at the genomic-level and the subject-level. The unavailability of paired sequencing data (that is, sequencing data of MGUS and MM from the same sample), because all the MGUS subjects do not progress to MM, and the unavailability of reliable workflows for analyzing a pool of a large mutational information to decipher accurate and reliable genomic information, biomarkers, and significantly altered pathways pose key challenges at the genomic-level. Moreover, at the subject-level, limited information in the studies about the time intervals of a subject's treatment and death times pose key challenges in pursuing disease progression and a reliable identification of critical genes, genomic features, and signaling pathways for targeted therapeutics.

With advancements in bioinformatics and increasing inclination toward machine learning (ML) or deep learning (DL), newer methods are

being developed for deducing salient information from the genomics data. For example, ML models have been developed to predict the survival outcome and treatment sensitivity in multiple myeloma [14, 15]. Similarly, AI-assisted risk stratification models for the prediction of survival and deciding the treatment regimen have been developed for the newly diagnosed multiple myeloma patients [16, 17]. Pathway enrichment analysis and classification has been shown to improve with the imputation of missing values in the microarray data of blood cancers via ML methods [18, 19]. ML/DL methods have also been proposed to detect somatic mutations from whole exome sequencing data [20, 21], prediction of copy number variants from whole exome data [22-24], driver genes in cancer [25-29] and, prediction of the survival-outcome and treatment-sensitivity in MM [14, 15].

However, the multi-dimensional analysis of exonic mutational profiles from exome sequencing data with gene-gene interaction has not yet been explored. This can be a promising direction for detecting key biomarkers in any cancer type. In recent years, geometric deep learning (GDL) has emerged to incorporate graph structures into a deep learning framework. Graph Convolutional Networks (GCNs) [30, 31], a type of GDL, can learn gene regulatory networks and do disease classification based on the network topology and disease-associated features, enabling an integration of graph-based data with genomic profiles [32]. The protein-protein interaction (PPI) network captures the physical interactions between proteins in an organism. Since the level of proteins and their interplay govern the molecular, cellular, and signaling controls which are the key to gene-level functionality and can help in capturing disease specific information, PPI networks can be immensely helpful if utilized alongside genomic information. A study on the exploration of the PPI network reported that the disease-related components in the PPI network are likely to be found in the network-based vicinity of disease components [33]. Similarly, another study on the PPI network revealed that the genes that contribute to a common disorder show an increased tendency of their protein-protein interactions [34]. These observations indicate that, due to the interconnected nature of a PPI network, genes belonging to similar diseases have a high predilection for interacting with other genes, forming a disease module. Therefore, identifying such genes or disease modules with the help of the PPI network can divulge the disease-related signaling pathways or other disease genes. These observations motivated us to incorporate the biological interactions in between genes as a key attribute of the bio-inspired BDL-SP model. Thus, we have incorporated the PPI information from the STRING database [35], which is the most comprehensive and global PPI network.

Motivated by the above discussion, this study addresses the problem of identifying significant biomarkers that differentiate MGUS from MM by incorporating a multidimensional analysis of exome profiles and their PPI network in a bio-inspired deep learning-based architecture from signaling pathways (BDL-SP) model. One of the challenges with this task is the ability to analyze a large amount of mutational information, a significant amount of which overlap in MGUS and MM samples. Since this mutational information is not easy to decipher for extracting differentiable patterns among MGUS and MM, the current literature shows this gap. To address the above gap, we have designed and implemented a GCN-based model, a *bio-inspired deep learning-based architecture from signaling pathways (BDL-SP)*, for extracting important genomic information to discern MGUS and MM. BDL-SP model uses single nucleotide variation (SNV) profiles of the significantly altered genes from the exome sequencing data along with the topological features of the PPI network, with an aim to identify pivotal biomarkers that can distinguish MGUS from MM. An in-depth analysis has been carried out for the identification of significantly altered genes and pathways that are specifically associated with MM and may be beneficial for the early identification of MGUS patients who are at a high risk of progression to the malignant MM stage. This work can further lead to the identification of novel therapeutic targets, thereby, preventing or delaying the malignant transformation of MGUS to MM.

For post-hoc model explainability, ShAP (SHapley Additive exPlanations) [36] algorithm is considered as one of the emerging and preferred approaches for decoding a DL model as well as for estimating feature importance based on

their contribution to the model's predictions. The ShAP algorithm incorporates model-agnostic approximations and uniformly characterizes an approach for model explainability [37, 38]. We aimed to use ShAP for post-hoc explainability in order to extract the underlying cause of the model's predictions by analyzing the ShAP score of each individual gene and genomic feature. We ranked the significantly altered genes based on their contribution to disease classification using the ShAP score. Among all the ML models trained in this study, BDL-SP model reported the highest numbers of previously reported driver genes, oncogenes, TSGs, and actionable genes in the top-ranked significantly altered genes compared to the other models. BDL-SP model also shows novel genes in the top-ranking genes that are not reported in MM but found significantly altered and contributing significantly to the disease prediction. We performed pathway enrichment analysis for top-500 significantly mutated genes. We analyzed whether an altered signaling pathway becomes more or less significant with disease progression from MGUS to MM. We observed that several signaling pathways either become significant (from being insignificant at MGUS) or become more significant with disease progression from MGUS to MM.

We benchmarked the BDL-SP with several baseline ML models both quantitatively and qualitatively, and observed that BDL-SP outperformed the other models in both aspects. With the help of the BDL-SP model, we identified the genes and their corresponding enriched signaling pathways that significantly contributed to MM disease development. The BDL-SP model's findings helped us to improve the understanding of cell transformation from pre-malignant to malignant state and strategic diagnosis to support the early detection of transformation to MM.

Material and methods

Whole-exome sequencing datasets of MM and MGUS patients

In this work, we utilized two external whole-exome sequencing (WES) datasets available with controlled access and one in-house WES dataset of MM and MGUS patients. These datasets are: 1) Multiple Myeloma Research Foundation (MMRF) CoMMpass data (of American population), 2) EGA dataset (of European population), and 3) AIIMS WES dataset (of Indian population). The MMRF CoMMpass (https://research.themmrf.org) is an open-source, extensive clinical and molecular database of multiple myeloma. The majority of MM samples in MMRF CoMMpass dataset (>75%) were collected from the people of American ethnicity. The MMRF CoMMpass dataset is aimed to provide molecular characterization and to correlate clinical datasets of MM patients for finding new, actionable targets to facilitate future clinical trial designs [39]. In our study, we have included 1092 bone marrow (BM) samples of MM collected from the GDC portal via dbGaP authorized access (phs-000748; phs000348). This is to note that the MMRF dataset also contained 20 peripheral blood (PB) samples that were not included in this study for the uniformity of the data. Similarly, the European Genome-phenome Archive (EGA) contains more than 700 studies of multiple diseases (including cancer and non-cancer) worldwide. EGA (http://www.ebi.ac.uk/ega/) was launched in 2008 by the European Molecular Biology Laboratory's European Bioinformatics Institute (EMBL-EBI) to provide secure storage of biological data and distribution only to authorized users [40]. The whole exome sequencing data of 33 MGUS European patients were obtained from the EGA repository EGAD00001001901. Besides the above two external datasets, we also included the WES data collected in-house from patients of Indian origin registered at All India Institute of Medical Sciences (AIIMS), New Delhi, India. This dataset included 82 MM and 28 MGUS samples. We have used the tumor-normal matched pairs of all BM samples obtained from MMRF, EGA and AIIMS WES data repository. Thus, we have included MGUS and MM WES datasets from three different databases.

Methods

Data pre-processing: Four variant callers, namely, MuSE [41], Mutect2 [42], VarScan2 [43], and Somatic-Sniper [44], were used for finding the variants in MM patients from the MMRF CoMMpass study. Therefore, for each patient, four variant call format (VCF) files corresponding to each variant caller were downloaded from the GDC portal via dbGaP authorized access (phs000748; phs000348). Exome

data obtained from EGA and AIIMS were processed with an exome sequencing pipeline [45] using BWA [46] and GATK [47], which is also considered a standard pipeline and mostly adopted to process the exome sequencing data. Similar to the MMRF data, the single nucleotide variants (SNVs) in EGA and AIIMS exome sequencing data were extracted using MuSE, Mutect2, VarScan2, and SomaticSniper variant callers. SNVs were annotated using ANNOVAR tool [48] that provides information about mutated genes, mutation type, the property of being deleterious or not, and clinical validation. In our study, we considered 23 types of functionally significant SNVs clustered into three groups based on their functional impact as follows: 1) *Non-Synonymous (NS) SNV Group*: This group consists of non-synonymous SNVs, exonic, ncRNA_exonic, stop gain, stop loss, start loss, exonic; splicing, splicing, frameshift insertion, and frameshift deletion type SNVs; 2) *Synonymous SNV Group*: This group consists of synonymous SNVs, UTR3 and UTR5 SNVs; and 3) *Other SNV Group*: This group consists of non-frameshift insertion/deletion/substitution, intronic, intergenic, ncRNA_intronic, upstream, downstream, unknown, and ncRNA_splicing SNVs. The benign SNVs were filtered out using the FATHMM-XF method [49]. Genomic annotations of SNVs (i.e., SNV type, mutated gene name, etc.) obtained from ANNOVAR were pooled and analyzed to identify the top significantly mutated genes using the 'dndscv' tool [50] based on the *q*-value (≤0.05) in both MM and MGUS individually. Union of significantly mutated genes from all four variant callers for MM (1174 patients) and MGUS (61 patients) groups led to 617 and 362 genes, respectively, and further union of the genes mentioned above yielded a total of 824 genes (Table S1 of supplementary material). For each gene, a total of 28 genomic features were created that includes total variant count and the distributive statistics (maximum, mean, median, and standard deviation) of variant allele frequency (VAF) and allele depth (AD) of each of the three groups of SNVs (NS SNV group, synonymous SNV group, and Other SNV group). A detailed description of the 28 genomic features is presented in **Figure 1**. The complete AI workflow is presented in **Figure 2**. For gene-gene interaction network information, we used the STRING database to get protein-protein interaction (PPI) of 824 significantly altered genes. The STRING database contains all the known and predicted associations of protein-protein interactions, including physical and functional associations for more than 14000 organisms.

Proposed shallow bio-inspired deep learning architecture from signaling pathways (BDL-SP): The conventional convolutional neural network (CNN) often fails to learn data of non-Euclidean space because non-Euclidean data cannot be modeled into *n*-dimensional linear space. The protein-protein interaction (PPI) network used in our model has a similar underlying non-Euclidean structure. Thus, a Graph Convolutional Network (GCN) could help us learn PPI data of non-Euclidean space. The proposed BDL-SP model carries out disease classification using a graph convolutional network that learns significant features from the exomic mutational profiles of genes interacting among each other according to the PPI network interactions. The mathematical description of GCN model is as follows:

For a given undirected graph, $g = (v, \varepsilon)$ where v is a collection of a finite set of nodes and ε is a collection of the finite set of edges, a graph convolution network learn the node representation by applying the graph laplacian with the input feature matrix ($X \epsilon R^{N \times p}$, where N denotes the number of nodes and p the number of features) and follows the propagation rule for each layer shown below:

$$H^{(l+1)} = \sigma(LH^{(l)}W^{(l)})$$

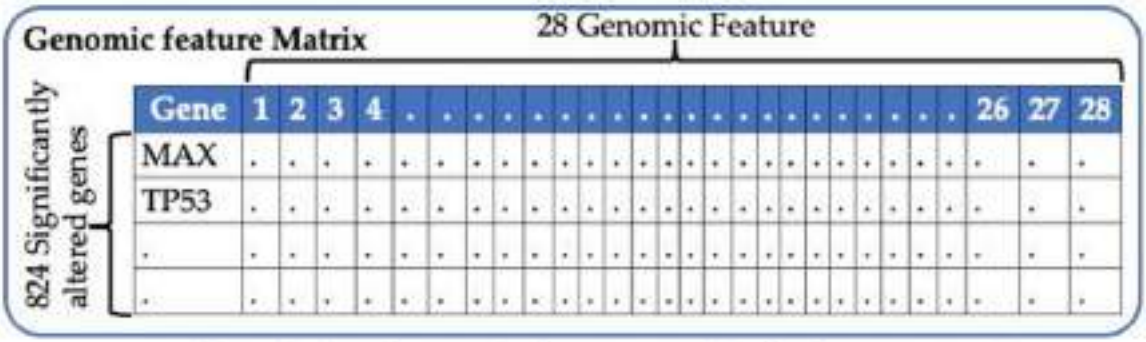
Where L denoted the normalized graph laplacian defined below.

$$L = I - D^{-\frac{1}{2}}\tilde{A}D^{-\frac{1}{2}} = U \wedge U^{T}$$

Where $D_{i,j} = \sum_{j=1}^{n} A(i,j)$, degree matrix of the graph and $\tilde{A} = A+I$ where A is the adjacency matrix, U is the matrix of eigenvectors of graph, $\wedge$ denote the respective eigenvectors, and $W \epsilon R^{p \times m}$ (where m corresponds to the number of filters in the graph convolution) denotes a learnable weight matrix. A GCN model transform a graph into the spectral domain by graph Fourier transformation [30] defined as below:

$$x * g = UgU^{T}x$$

The above Fourier transformation can be computed by approximating Chebyshev polynomi-



Description of genomic features used in the feature matrix

Feature Number	Feature Name
1	Total number of the SNVs
2	Total number of the SNVs in synonymous group
3	Maximum VAF of the SNVs in synonymous group
4	Median VAF of the SNVs in synonymous group
5	Mean VAF of the SNVs in synonymous group
6	VAF's standard deviation of the SNVs in synonymous group
7	Maximum AD of the SNVs in synonymous group
8	Median AD of SNVs in synonymous group
9	Mean AD of SNVs in synonymous group
10	AD's standard deviation of the SNVs in synonymous group
11	Total number of SNVs in non-synonymous group
12	Maximum VAF of SNVs in non-synonymous group
13	Median VAF of SNVs in non-synonymous group
14	Mean VAF of SNVs in non-synonymous group
15	VAF's standard deviation of the SNVs in non-synonymous group
16	Maximum AD of SNVs in non-synonymous group
17	Median AD of SNVs in non-synonymous group
18	Mean AD of SNVs in non-synonymous group
19	AD's standard deviation of the SNVs in non-synonymous group
20	Total number of SNVs in other group
21	Maximum VAF of SNVs in other group
22	Median VAF of SNVs in other group
23	Mean VAF of SNVs in other group
24	VAF's standard deviation of the SNVs in other group
25	Maximum AD of SNVs in other group
26	Median AD of SNVs in other group
27	Mean AD of SNVs in other group
28	AD's standard deviation of the SNVs in other group

Figure 1. Schematic layout of genomic feature matrix used for the training of proposed BDL-SP model. The dimension of the genomic feature matrix is 824×28 with 824 significantly altered genes (See Table S1 of supplementary material) and 28 genomic features obtained from MMRF, EGA and AIIMS WES datasets using the AI-based workflow shown in **Figure 2**. The genomic features were extracted from three groups of SNVs, namely, 1. Non-synonymous SNV group, 2. Synonymous SNV group, and 3. Other SNV group. A total of nine features were extracted for each SNV group to learn the distributive statistics (maximum, mean, median, and standard deviation). The full form of abbreviations used in this figure are as follows: SNVs = Single Nucleotide Variations, VAF = Variant Allele Frequency, and AD = Allele Depth.

als and the renormalization trick mentioned in [30] as:

$$Z = \tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}X\,W$$

The infographic representation of the architecture of BDL-SP with end-to-end pipeline is shown in **Figure 3** and is explained here. The BAM files from EGA and AIIMS datasets, and VCF files from the MMRF dataset are processed to extract 824 significantly altered genes using the dndscv tool (as shown in the WES Data pre-processing block) in **Figure 3**. The interaction among these 824 genes is extracted using the protein-protein interaction (PPI) network (from the STRING database). A network of nodes and edges is set-up using this information, where each node denotes one of these 824 genes and each link implies that the two nodes/genes of that link were connected as per the PPI network. Each node is set up with its 28-length feature-vector extracted earlier (as shown in **Figure 1**). Hence, the 28-length genomic feature-vectors of all 824 genes are added to the network established using the PPI network. This input layer is followed by two hidden layers of GCN, that are further followed by one fully connected layer of 824 neurons to 2 neurons giving output through log-softmax activation function. Since there were 95% samples of MM class and 5% samples of MGUS class, which made the data highly imbalanced (class imbalance ratio = 19.22), a cost-sensitive loss function was utilized to train the BDL-SP model in order to deal with the data imbalance problem. BDL-SP is trained in a supervised fashion, where the MM/MGUS target class label along with the feature matrix of 824×28 is provided as an input to the architecture. The network is trained until the loss reduces and saturates. Five-fold cross-validation is performed that led to the training of five BDL-SP classifiers, one for each fold of test data. Next the ShAP algorithm is used on these five trained BDL-SP classifiers to obtain the top genomic features and significantly altered signaling pathways as explained in the next subsection. The setting of layers of BDL-SP and the hyperparameters values are shown in **Table 1**.

Quantitative benchmarking of BDL-SP model with traditional machine learning classifiers: We have benchmarked the proposed BDL-SP model with six baseline ML models (random forest, decision tree, logistic regression, XGBoost, CatBoost, and SVM from scikit-sklearn [51]). The conventional cost-blind machine learning models do not account for the imbalanced classes in the data and tend to make decisions favoring the majority class resulting in misclassification. In the case of medical diagnosis, such misclassification can lead to erroneous direction of strategic treatment, causing patients to suffer. In our study, there were 95% samples of MM class and 5% samples of MGUS class, which made the data highly imbalanced (class imbalance ratio = 19.22). Therefore, we have used cost-sensitive ML models to account for the class imbalance in our data and model benchmarking. During training, the cost-sensitive loss function penalizes the mistake in identifying each MGUS sample (minority class) more compared to the mistake in identifying each MM sample (majority class). This ensures that the classifier is not biased to the majority class and learns to identify the samples of both the classes. These baseline models are trained with the traditional data pre-processing pipeline using principal component analysis (PCA). Each baseline ML model was trained exhaustively with five-fold cross-validation, where the confusion matrix of the hold-out set was kept separate for each fold. The final confusion matrix was obtained by adding the confusion matrices of all five hold-out sets and the performance metrics were calculated for each ML model.

Qualitative application-aware post-hoc benchmarking of BDL-SP model using ShAP: ShAP (SHapley Additive exPlanations) is an algorithm

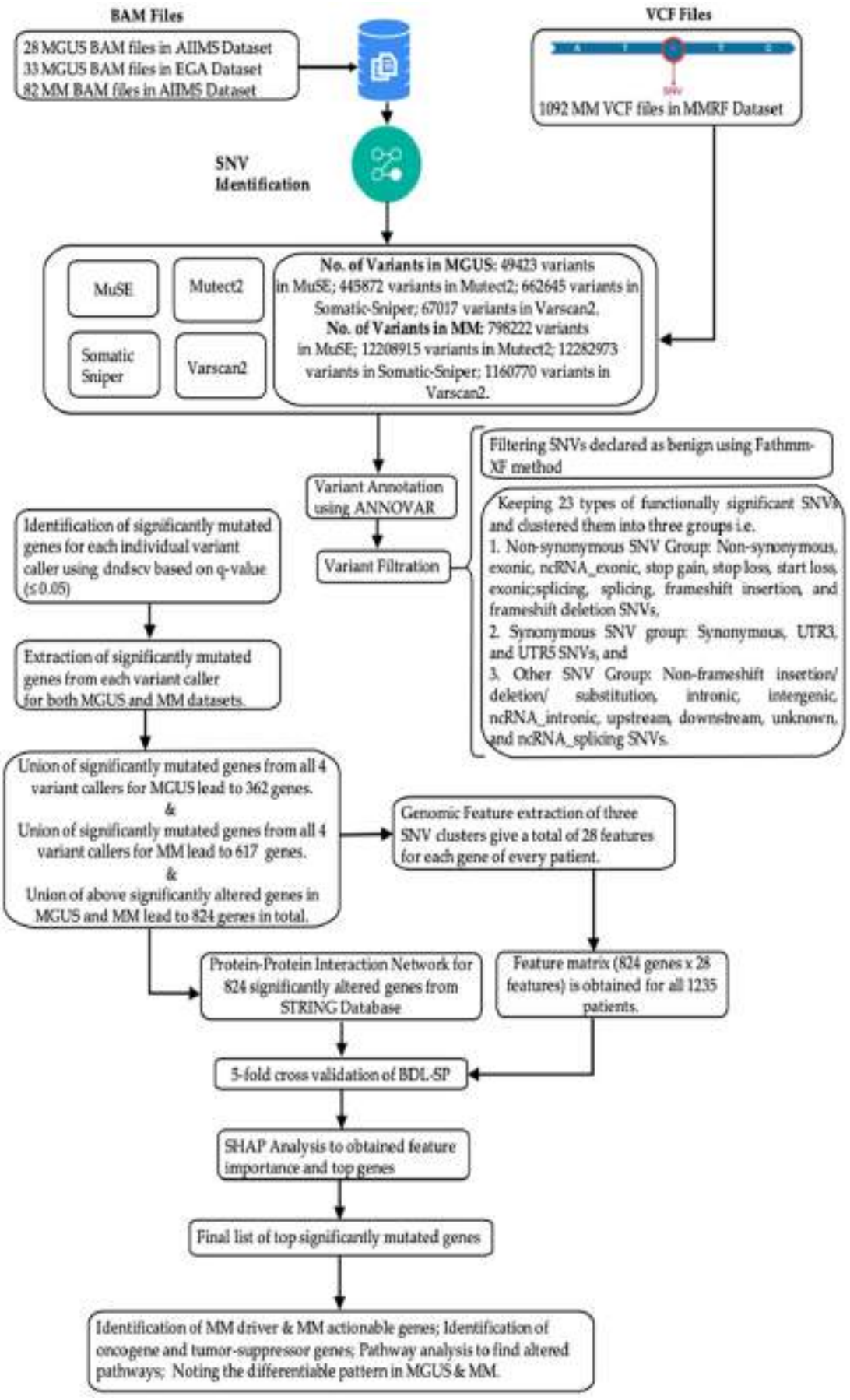

BAM Files
28 MGUS BAM files in AIIMS Dataset
33 MGUS BAM files in EGA Dataset
82 MM BAM files in AIIMS Dataset
VCF Files
1092 MM VCF files in MMRF Dataset
SNV Identification
MuSE
Mutect2
Somatic Sniper
Varscan2
No. of Variants in MGUS: 49423 variants in MuSE; 445872 variants in Mutect2; 662645 variants in Somatic-Sniper; 67017 variants in Varscan2.
No. of Variants in MM: 798222 variants in MuSE; 12208915 variants in Mutect2; 12282973 variants in Somatic-Sniper; 1160770 variants in Varscan2.
Variant Annotation using ANNOVAR
Filtering SNVs declared as benign using Fathmm-XF method
Variant Filtration
Keeping 23 types of functionally significant SNVs and clustered them into three groups i.e.
1. Non-synonymous SNV Group: Non-synonymous, exonic, ncRNA_exonic, stop gain, stop loss, start loss, exonic;splicing, splicing, frameshift insertion, and frameshift deletion SNVs,
2. Synonymous SNV group: Synonymous, UTR3, and UTR5 SNVs, and
3. Other SNV Group: Non-frameshift insertion/ deletion/ substitution, intronic, intergenic, ncRNA_intronic, upstream, downstream, unknown, and ncRNA_splicing SNVs.
Identification of significantly mutated genes for each individual variant caller using dndscv based on q-value (≤ 0.05)
Extraction of significantly mutated genes from each variant caller for both MGUS and MM datasets.
Union of significantly mutated genes from all 4 variant callers for MGUS lead to 362 genes.
&
Union of significantly mutated genes from all 4 variant callers for MM lead to 617 genes.
&
Union of above significantly altered genes in MGUS and MM lead to 824 genes in total.
Genomic Feature extraction of three SNV clusters give a total of 28 features for each gene of every patient.
Protein-Protein Interaction Network for 824 significantly altered genes from STRING Database
Feature matrix (824 genes x 28 features) is obtained for all 1235 patients.
5-fold cross validation of BDL-SP
SHAP Analysis to obtained feature importance and top genes
Final list of top significantly mutated genes
Identification of MM driver & MM actionable genes; Identification of oncogene and tumor-suppressor genes; Pathway analysis to find altered pathways; Noting the differentiable pattern in MGUS & MM.

Figure 2. AI-based workflow to infer differentiable genomic biomarkers to identify MGUS and MM using the whole-exome sequencing (WES) data.

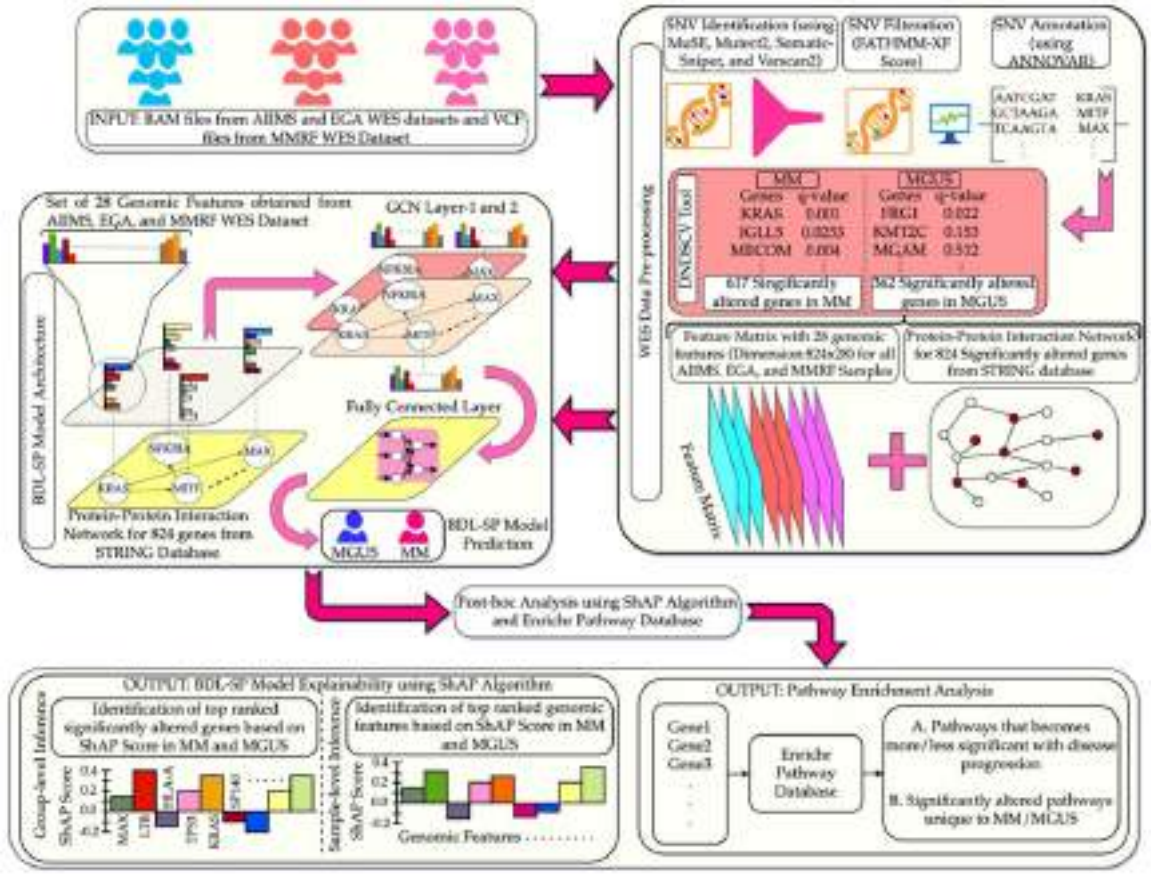


Figure 3. Infographic representation of the proposed AI-based BDL-SP model architecture and the application-aware post-hoc analysis for the identification of pivotal genomic biomarkers that distinguish MGUS from MM. The BAM files from EGA and AIIMS datasets and VCF files from MMRF dataset are processed to extract 824 significantly altered genes using the dndscv tool (as shown in the WES Data pre-processing block). The interaction among these 824 genes is extracted using the protein-protein interaction (PPI) network (from STRING database). A network of nodes and edges is set-up using this information, where each node denotes one of these 824 genes and each link implies that the two nodes/genes of that link were connected as per the PPI network. Each node is set up with its 28 genomic features extracted for the corresponding gene as explained earlier. This input layer is followed by two hidden layers of GCN, one fully connected layer, and a softmax layer at the output. Thus, each subject's WES data is analyzed and the feature vectors of all 824 genes are extracted. These are given as input along with the subject's MM/MGUS target class label to train the GCN in a supervised mode. Once the BDL-SP model is learned to distinguish MGUS from MM, the top genomic features and significantly altered signaling pathways were obtained from ShAP algorithm and the Enrichr Pathway Database.

that measures the significance of an attribute in the prediction of a model, scoring each attribute proportional to its contribution. Therefore, it was utilized for the post-hoc explainability of the BDL-SP model. The most-contributing genomic features and significantly altered genes at the group (i.e., either MGUS or MM) as well as at the individual sample-level were identified. Since five-fold validation was carried out during training, the ShAP algorithm was applied on each trained classifier to obtain the significant genomic attributes (both genes and genomic features) for each sample. Note that the ShAP score can either be positive or negative. Here, the positive ShAP score for an attribute indicates its contribution to the model's prediction toward the MGUS class (positive class), while the negative score indicates its

Table 1. Hyperparameters values and layer dimensions of the BDL-SP model architecture

GCN Architecture Attribute/Hyperparameter	Hyperparameter Value
No. of GCN Layers	2
GCN layer dimensions	Input sample dimension: 824×28 1st Layer (For each node): 28×7 2nd Layer (For each node): 7×1 Output dimension: 824×1
Output Linear Layer dimension	824×2 (number of classes = 2)
Activation Function	LeakyReLU (0.1)
Dropout	0.75
Cost function and Adjusted Cost for class imbalance	Cost Function: Cross-Entropy Loss Cost Adjusted: 20.0
GCN Weight Initialization	Uniform Xavier

contribution to the model's prediction toward the MM class (negative class). Therefore, the higher the magnitude of the ShAP score, the higher its impact on the model's positive class outcome. Moreover, only those samples were considered for extracting ShAP interpretability that were correctly predicted by at least one of the five classifiers.

Next, we devised the algorithms for the estimation of the best ShAP score 1) for all 824 significantly altered genes (**Figure 4A**) and 2) for all 28 genomic features (**Figure 4B**) at a sample-level to understand their contribution to the BDL-SP model's prediction. The pseudocodes with mathematical description for estimating the best ShAP scores for genes and genomic features are provided in **Table 2**, Algorithm-A, and Algorithm-B. The algorithms shown in **Figure 4A** and **4B** take the sample feature matrix as input and estimate the best ShAP scores for genes and genomic features at a sample-level. For each sample feature matrix, the corresponding sample class was predicted using all five trained classifiers of the BDL-SP model and the ShAP algorithm was applied only on those classifiers that predicted the sample's class correctly. Here, the ShAP score for all the genomic attributes were collected at the classifier-level and the sample-level. For each genomic attribute, the best ShAP score was first calculated at a classifier level and then the final best ShAP score was estimated among all classifiers at a sample-level. For each gene, we first collected the ShAP score of all 28 genomic features at a sample-level and then grouped them based on their positive and negative signs. Next, we compared the absolute value of the sum of ShAP scores of genomic features having the positive ShAP scores with the absolute sum of those having the negative ShAP scores. The ShAP score having the highest absolute value was considered as the best ShAP score for that gene and the classifier. This step was repeated for all those classifiers that predicted the sample's class correctly and the best ShAP score was saved for each of the classifiers. The ShAP score having the highest absolute value among all the classifiers was considered as the best ShAP score for a gene at a sample-level. For a better clarity of the steps employed in the estimation of the best ShAP scores of significantly altered genes and genomic features at a sample-level, one may refer to **Figure 4A** and Algorithm-A of **Table 2**, **Figure 4B** and Algorithm-B of **Table 2**, respectively.

Similarly, for each genomic feature, we first collected the ShAP score of all 824 genes at a sample-level and grouped them based on their positive and negative signs. Next, we compared the absolute value of the sum of ShAP scores of genes having positive scores with the sum of ShAP scores of genes having negative ShAP scores. The ShAP score having the highest absolute value was considered as the best ShAP score for a genomic feature and the classifier. We repeated the above step for all the classifiers that predicted the sample's class correctly and saved the best ShAP score for each of the classifiers. The ShAP score having the highest absolute value among all the classifiers was considered as the best ShAP score for that genomic feature at a sample-level. Once the best ShAP scores were obtained for all the genes and all the genomic attributes, the top ranked genes and the top ranked genomic attri-

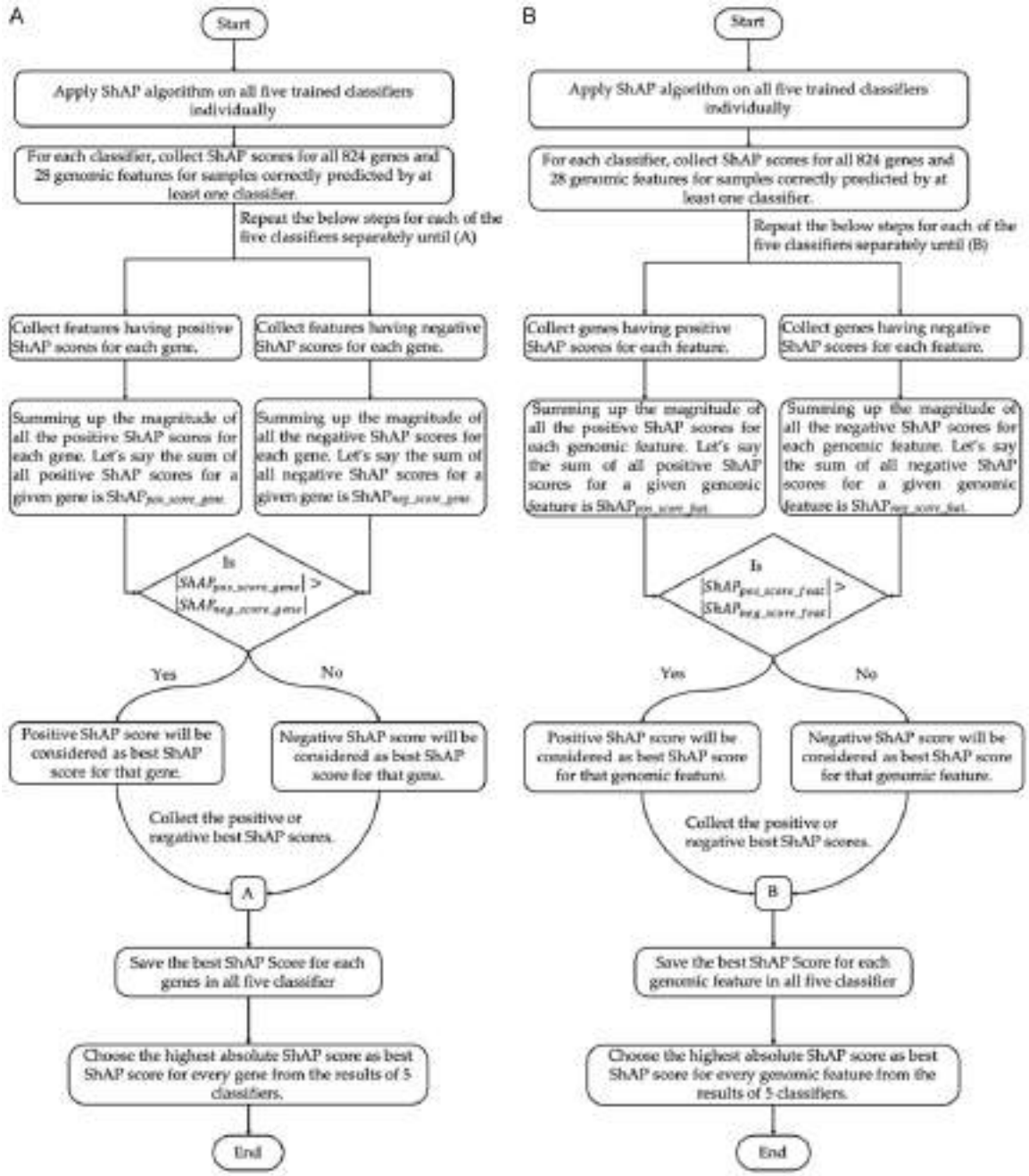


Figure 4. Flowchart showing steps for estimating the best ShAP score for (A) 824 significantly altered genes and (B) 28 genomic features at sample-level to reveal their contribution to the BDL-SP model prediction.

butes were identified at the group-level and at the sample-level.

Further, the top-ranked significantly altered genes revealed by BDL-SP were also compared with the multiple myeloma related studies to identify the previously reported significantly altered genes. We included information from multiple databases for model validation and post-hoc analysis at gene level analysis (OncoKB, COSMIC, IntoGen, and TargetDB databases). We downloaded a list of 1064 cancer genes from OncoKB [52] to deduce the oncogenes and tumor-suppressor genes in our top mutated genes. Further, 318 oncogenes and 320 tumor-suppressor genes obtained

Table 2. (A) Pseudo-codes of algorithm A for estimating the best ShAP score of 824 genes and (B) Algorithm B for estimating the best ShAP score of 28 genomic features at a sample level

Algorithm A: Estimate the Best ShAP Score (BSS) for each gene at a sample level

1. Fivefold classifiers = [List of five classifiers trained on each fold of test dataset]
2. CPC = [List of correct prediction classifiers i.e. classifiers that correctly predicted the sample's class]
3. SFM = Sample feature matrix
4. Genes = [List of 824 genes]
5. $GFPS_{g|c}$ = [List of genomic features having positive ShAP score for a gene "g" and classifier "c"]
6. $GFNS_{g|c}$ = [List of genomic features having negative ShAP score for a gene "g" and classifier "c"]
7. $CSG_{g|c}$ = Best ShAP score of gene "g" and classifier "c"
8. $ACGS[classifier]_{g}$ = [List of best ShAP scores of gene "g" for all the classifiers that correctly predicted the sample]
9. BSG_{g} = Best ShAP score of gene "g" among all classifiers
10. $LBSG_{genes}$ = List of the best ShAP score of all the genes among all the classifiers
11. function BSS gene (SFM)
12. for classifier in (Fivefold classifiers) do
    A. Predict the sample's class with the help of a classifier
    B. if classifier predict the sample class correctly then
        a. Append classifier in CPC list
        b. Apply ShAP algorithm on the classifier
        c. Collect the ShAP score for all 824 genes on their respective 28 GF for that classifier
13. for gene in Genes do
    A. for classifier in CPC do
        a. $GFPS_{gene|classifier} \leftarrow$ Collect features having positive ShAP score
        b. $GFNS_{gene|classifier} \leftarrow$ Collect features having negative ShAP score
        c. If $|\sum GFPS_{gene|classifier}| > |\sum GFNS_{gene|classifier}|$ then
        d. $CSG_{gene|classifier} \leftarrow GFPS_{gene|classifier}$
        e. else
        f. $CSG_{gene|classifier} \leftarrow GFNS_{gene|classifier}$
        g. $ACGS[classifier]_{gene} \leftarrow CSG_{gene|classifier}$
    B. $BSG_{gene} \leftarrow ACGS[argmax[|CSG| \text{ for } CSG \text{ in } ACSG]]$
    C. $LBSG_{genes}[gene] \leftarrow BSG_{gene}$
14. Output: $LBSG_{genes}$

Algorithm B: Estimate the Best ShAP Score (BSS) for each genomic feature (GF) at a sample level

1. Fivefold classifiers = [List of five classifiers trained on each fold of test dataset]
2. CPC = [List of correct prediction classifiers i.e. classifiers that correctly predicted the sample's class]
3. SFM = Sample feature matrix
4. Genomic Features = [List of 28 GFs]
5. $GPS_{gf|c}$ = [List of genes having positive ShAP score for a genomic feature and classifier]
6. $GNS_{gf|c}$ = [List of genes having negative ShAP score for a genomic feature and classifier]
7. $CSGF_{gf|c}$ = Best ShAP score of GF "gf" and classifier "c"
8. $ACGFS[classifier]_{gf}$ = [List of best ShAP scores of gene "gf" for all the classifiers that correctly predicted the sample]
9. $BSGF_{gf}$ = Best ShAP score of GF "gf" among all classifiers
10. $LBSGF_{gfs}$ = List of the best ShAP score of all GF among all classifiers
11. function BSS genomic feature (SFM)
12. for classifier in (Fivefold classifiers) do
    A. Predict the sample's class with the help of a classifier
    B. if classifier predict the sample class correctly then
        a. Append classifier in CPC list
        b. Apply ShAP algorithm on the classifier
        c. Collect the ShAP score for all 824 genes on their respective 28 GF for that classifier

13. for feature in Genomic features do
  A. for classifier in CPC do
    a. $GPS_{feature|classifier} \leftarrow$ Collect genes having positive ShAP score
    b. $GNS_{feature|classifier} \leftarrow$ Collect genes having negative ShAP score
    c. if $|\sum GPS_{feature|classifier}| > |\sum GNS_{feature|classifier}|$ then
    d. $CSGF_{gf|classifier} \leftarrow GPS_{feature|classifier}$
    e. else
    f. $CSGF_{gf|classifier} \leftarrow GNS_{feature|classifier}$
    g. $ACGFS[classifier]_{gf} \leftarrow CSGF_{gf|classifier}$
  B. $BSGF_{gf} \leftarrow ACGFS[argmax[|CSGF| \text{ for } CSGF \text{ in } ACGFS]]$
  C. $LBSGF_{gfs}[feature] \leftarrow BSGF_{gf}$
14. Output: $LBSGF_{gfs}$

from COSMIC database [55] were also used to deduce oncogenes and tumor-suppressor genes in our top-mutated genes. Similarly, we created a list of MM driver genes reported by [7, 53]. MM Driver genes were also extracted from IntoGen database [54] (https://www.intogen.org/) to infer MM drivers genes present in our top mutated gene list. Finally, a list of 180 actionable genes from the COSMIC database [55] and 135 actionable genes from the TargetDB database [56] was used to infer the actionable genes present in our top mutated gene list. The top ranked significantly altered genes were grouped in four categories based on their functional significance as follows: 1. Oncogenes (OGs); 2. Tumor-Suppressor genes (TSGs); 3. Onco-driver genes (ODGs); 4. Actionable genes (AGs).

The top-ranked significantly altered genes in each of the above gene categories were then collected at the group-level (MM/MGUS) and the sample-level. We also checked the role of genomic features on the disease classification in post-hoc interpretability analysis of the BDL-SP model.

Statistical analysis: We performed the unpaired Wilcoxon ranksum statistical analysis to study the impact of ethnicity in multiple myeloma. In this analysis, we first extracted the top significantly altered genes from the WES data of MGUS/MM patients of American (MMRF), European (EGA), and Indian (AIIMS) population using the top performing BDL-SP model. Next, for each sample, we computed the total number of significantly altered genes that belonged to the reported categories of OGs, TSGs, ODGs, and AGs of MM literature. Then, we performed a statistical comparison of the number of significantly altered genes of the reported category of OGs, TSGs, ODGs, and AGs on the groups of American (MMRF), European (EGA), and Indian (AIIMS) population to study the impact of ethnicity on individual gene category.

Gene pathway analysis: The significant genes identified by BDL-SP, which helped in differentiating MM from MGUS, were mapped back to the significant gene list obtained for MM and MGUS using the dndscv tool. Some genes were found to be common in both the groups, while some were found to be significantly mutated either in MGUS or in MM only. Pathway analysis was done on the top-500 genes obtained from the BDL-SP model. KEGG and Reactome pathways were found via Enrichr gene set enrichment analysis web server [57-59].

Results

Using the dndscv tool (as shown in **Figure 2**), 362 and 617 significantly altered genes were identified in MGUS and MM, respectively. Of these, 155 genes were common in MGUS and MM. The complete list of all 824 genes is shown in Table S1 of supplementary material. We then inferred the important genes that were accountable for distinguishing MGUS from MM as obtained through our graph-based BDL-SP model.

Comparative performance of BDL-SP and standard ML models

Using our AI-based workflow of BDL-SP (**Figures 2** and **3**), we trained the BDL-SP model with a 5-fold cross-validation and compared its performance with six standard cost-sensitive machine learning models. Results of the BDL-

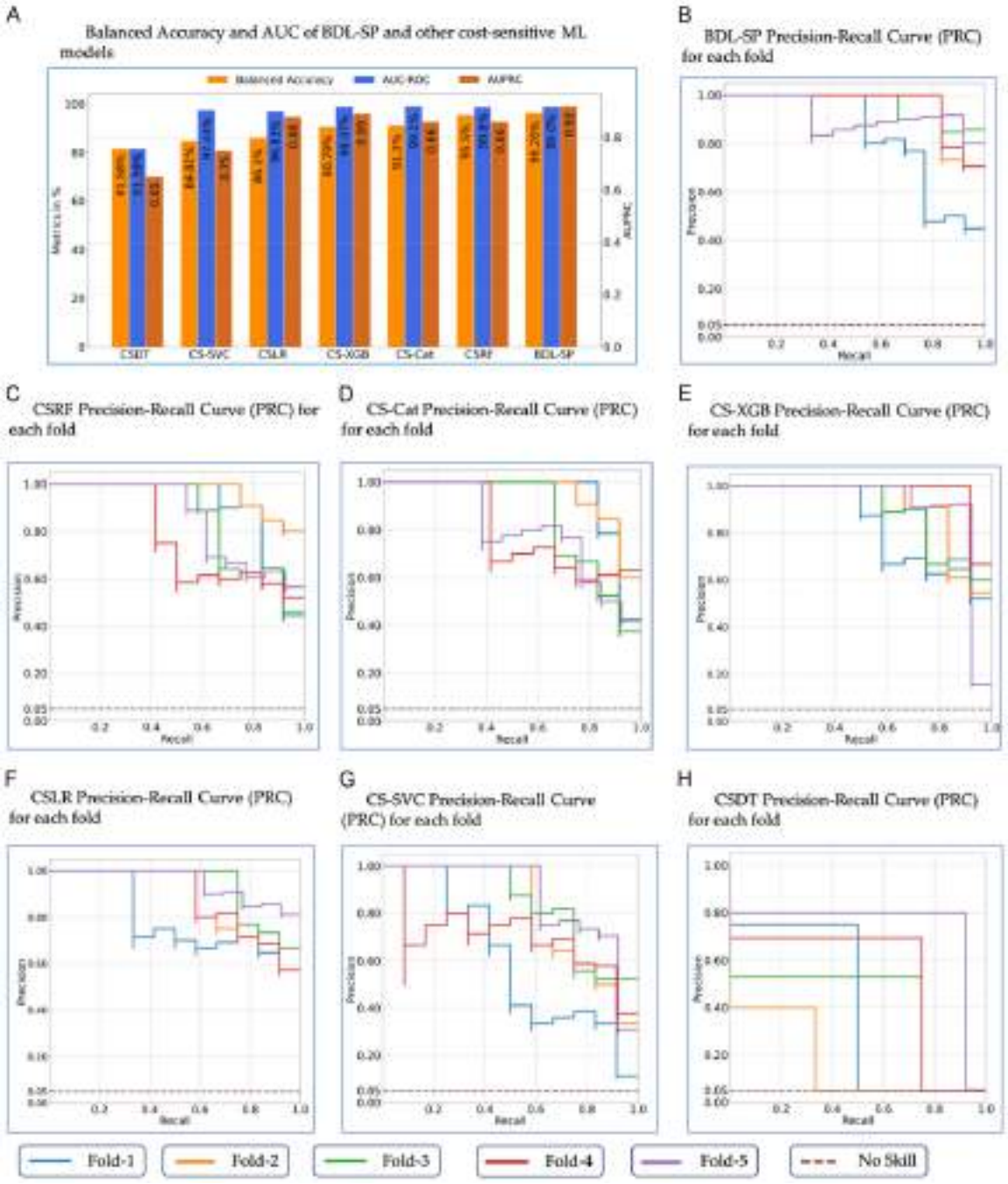


Figure 5. (A) The benchmarking of the performance of BDL-SP with six cost-sensitive ML models on the metrics of balanced accuracy, AUC, and AUPRC (Area under Precision-Recall Curve). Precision-Recall Curves (PRC) for all five folds of (B) BDL-SP, (C) CSRF, (D) CS-Cat, (E) CS-XGB, (F) CSLR, (G) CS-SVC, and (H) CSDT. No skill line is also shown in each of the AUPRC plots that represent the inability of the classifier to correctly classify a sample. The full form of the abbreviation used in this figures are as follows: CSDT = Cost-Sensitive Decision Tree, CS-SVC = Cost-Sensitive Support Vector Machine, CSLR = Cost-Sensitive Logistic Regression, CS-XGB = Cost-Sensitive XGBoost, CS-Cat = Cost-Sensitive CatBoost, and CSRF = Cost-Sensitive Random Forest.

SP model and all the six cost-sensitive classifiers are presented in **Figure 5**. The proposed BDL-SP model outperformed the rest of the models in terms of the balanced accuracy and AUPRC (area under precision-recall curve), while the area under the curve (AUC) was highest (and equal) for the top three models. BDL-SP model performed best with a balanced

Table 3. Types of four different gene categories (OG, TSG, ODG, and AG) and their counts in 824 significantly altered genes

Gene type based on functionality	Total number of previously reported genes present in our list of 824 significantly altered genes
Oncogenes (OGs)	31
Tumor-suppressor genes (TSGs)	43
Both oncogene and driver gene (ODGs)	10
Actionable genes (AGs)	19

Table 4. Counts of previously reported 4 categories of genes as found in the post-hoc analysis based top-250 and top-500 genes of the top-3 models (BDL-SP, CS-RF, and CS-Cat)

Top Gene	BDL-SP Model (Top-performing model)				CS-RF Model (Second best model)				CS-Cat Model (Third best model)			
	OG	TSG	ODG	AG	OG	TSG	ODG	AG	OG	TSG	ODG	AG
Top-250	**20**	**21**	**7**	**11**	7	10	1	4	6	5	1	4
Top-500	**27**	**37**	**10**	**17**	7	10	1	4	6	5	1	4

The number of previously reported genes (OG/TSG/ODG/AG) obtained in each category (top-250/top-500) using the best performing model are highlighted in bold.

accuracy of 96.26%. Cost-sensitive Random Forest (CS-RF) performed the next best with a balanced accuracy of 95.5%, and cost-sensitive Catboost (CS-Cat) performed the third best with a balanced accuracy of 91.3% (**Figure 5A**). All these three models reported an AUC value of 0.99. BDL-SP model also outperformed other models on AUPRC scoring the highest AUPRC of 0.92, while the AUPRC of both CSRF and CS-Cat model was 0.86 (**Figure 5B-D**). This is to note that AUPRC is one of the most important quantitative metrics and is more relevant compared to AUC on the unbalanced data. BDL-SP outperformed the other models on AUPRC with a great margin. This shows that, quantitatively, BDL-SP performed best, with the CS-RF model being the second best model.

BDL-SP identified the maximum number of minority class samples, i.e., 60 out of 61 MGUS samples, and 1087 MM samples out of a total of 1153 MM samples. The second-best model was CS-RF that identified 59 out of 61 MGUS samples and 1086 out of 1153 MM samples. The third best performing model was CS-Cat that identified 52 out of 61 MGUS samples and 1121 out of 1153 MM samples. Thus, again BDL-SP outperformed other models on minority class detection, CS-RF performed next to this model. Since the performance of CS-RF was close to the leading BDL-SP model on metrics other than AUPRC, we performed post-hoc interpretability benchmarking of the top-three performing models (BDL-SP, CS-RF, and CS-Cat). In post-hoc benchmarking, we utilized the ShAP algorithm and tabulated the top-250 and top-500 genes identified by the top-three trained models to understand the reasons for the models' predictions. Then, the top-ranked genes were further analyzed to identify previously reported oncogenes (OGs), tumor-suppressor genes (TSGs), both oncogenes and driver genes (ODGs), and actionable genes (AGs) in MM. As demonstrated later in this Section with the post-hoc interpretability analysis results, we observed that BDL-SP identified the maximum number of the previously reported genes in top-250 and top-500 genes.

Out of 824 significantly altered genes identified from the workflow shown in **Figure 2**, there were 31 oncogenes (OGs) (e.g. *KRAS, LTB, CARD11, NOTCH1*, etc.), 43 tumor-suppressor genes (TSGs) (e.g. *HLA-A/B/C, TRAF3, TP53, SDHA*, etc.), ten genes that were both oncogenes and driver genes (*KRAS, LTB, NRAS, FGFR3, BRAF*), and 19 actionable genes (e.g. *MITF, ARID1B, ARID2, RPTOR*, etc.) (**Table 3**). This full list of genes is provided in Table S1 of supplementary material. For each of the top-three models, we have considered only those genes in the top-250 or top-500 gene list that have a non-zero ShAP score in the post-hoc explainability analysis. The total counts of previously reported genes as found in the top-250 and top-500 genes of the top-three models is shown in **Table 4**.

From **Table 4**, we observed that BDL-SP model identified 20 out of 31 oncogenes in top-250 and 27 out of 31 oncogenes in the top-500 gene list, while CS-RF and CS-Cat could identify only 7 and 5 oncogenes in top-250 and top-500 gene list, respectively. Similarly, out of 43 TSGs, BDL-SP model identified 21 and 37 TSGs in the top-250 and top-500 gene list, while CS-RF and CS-Cat identified only 10 and 5 TSGs, respectively, in the top-250 and top-500 gene list. Further, the BDL-SP model identified 7 and all ten ODGs, while CS-RF and CS-Cat could identify only one ODG in top-250 and top-500 significantly altered genes. Finally, the BDL-SP model identified 11 and 17 actionable genes in top-250 and top-500 genes, respectively, while CS-RF and CS-Cat could identify only 4 actionable genes in top-250 and top-500 significantly altered genes. The post-hoc benchmarking of the top-three models is shown in **Table 4** and the list of OGs, TSGs, ODGs, and AGs in the top-250 and top-500 significantly altered gene list of BDL-SP, CS-RF, and CS-Cat models is provided in **Table 5**. Since BDL-SP model identified the highest number of previously reported OGs, TSGs, ODGs, and AGs, this model can be inferred as the best performing model and was used subsequently for inferring the top significantly altered genes, genomic features, and altered signaling pathways to identify the pivotal genomic biomarkers to distinguish MM and MGUS. This analysis shows that one can obtain similar quantitative results with two or more different ML models, but one should choose the model that is more interpretable with reference to the application domain.

Pathway analysis on the top 500 genes obtained from the BDL-SP model

On comparing the top-500 significantly altered genes obtained from the BDL-SP model (that helped in differentiating MM from MGUS) to the significant gene list obtained for MM and MGUS using the dndscv tool, 301 genes were observed to be statistically significantly mutated only in the MM cohort, 101 genes were observed to be statistically significantly mutated only in the MGUS cohort, while 98 genes were observed to be statistically significantly mutated in both MM and MGUS cohorts. The set of 301 genes that were found to be significantly mutated only in the MM cohort included several important OGs, ODGs, TSGs, and AGs such as *BCL7A, BRAF, CARD11, CYLD, DIS3, EGR1, FAM46C, IGLL5, KRAS, KMT2D, NRAS, MECOM,* etc. Similarly, the set of 101 genes significantly mutated only in the MGUS cohort included *APC, FAM47B, MGAM, NOTCH1, TYRO3*, etc. The set of 98 common genes observed to be significantly mutated in MGUS and MM cohorts included *AMER1, FANCD2, HLA-B, KMT2C, PABPC1, TRRAP*, etc. The complete list of top significantly altered genes only in MM, only in MGUS, and common in both MM and MGUS is provided in Table S7 of supplementary material.

Enrichr and Reactome were used to infer the KEGG and Reactome pathways altered by 399 MM and 199 MGUS genes. A total of 5 KEGG pathways inferred from Enrichr were signifi cantly altered in MGUS (Table S2 of supple mentary material) and 108 KEGG pathways were significantly altered in MM (Table S3 of supplementary material). Similarly, a total of 10 Reactome pathways inferred from Enrichr were significantly altered in MGUS (Table S2 of supplementary material) and 134 Reactome pathways inferred from Enrichr were significantly altered in MM (Table S3 of supplementary material). Further, we grouped the significantly altered pathway into four categories based on the variations in their significance with disease progression from MGUS to MM as:

1. *Category-1*: Pathways that become more significant with disease progression from MGUS to MM.

2. *Category-2*: Pathways that become less significant with disease progression from MGUS to MM.

3. *Category-3*: Significantly altered pathways observed only in MM and and not observed in MGUS.

4. *Category-4*: Significantly altered pathways observed only in MGUS and not observed in MM.

The complete list of significantly altered pathways for the above mentioned four categories are provided in Tables S4 and S5 of supplementary material. In Category-1 of significantly altered pathways, 05 KEGG and 09 Reactome pathways became more significant as the disease progressed from MGUS to MM (**Figure 6**).

Table 5. List of 4 categories of previously reported genes as found in the post-hoc analysis based top-250 and top-500 genes of the top-3 models (BDL-SP, CS-RF, and CS-Cat)

(A) List of oncogenes (OGs) and actionable genes (AGs) in top-250 and top-500 genes

Top Genes	BDL-SP Model (Top-performing model)		CS-RF Model (Second best model)		CS-Cat Model (Third best model)	
	OG	AG	OG	AG	OG	AG
Previously reported oncogenes and actionable genes in MM and MGUS as found ranked in top-250 during the post-hoc analysis of the model	MUC16, FGFR3, PABPC1, BIRC6, MUC4, IRS1, PGR, MGAM, VAV1, ABL2, MITF, TP53, RPTOR, NRAS, NOTCH1, BRAF, TCL1A, LTB, CARD11, KRAS	NRAS, TYRO3, NOTCH1, FGFR3, BRAF, ARID2, NF1, MITF, TP53, KRAS, RPTOR	TCL1A, LTB, RPTOR, ABL2, TAL1, VAV1, NOTCH1	RPTOR, NF1, NFKBIA, NOTCH1	TCL1A, MGAM, ABL2, VAV1, PGR, BRD4	NFKBIA, APC, BRD4, BRAF
Previously reported oncogenes and actionable genes in MM and MGUS as found ranked in top-500 during the post-hoc analysis of the model	MUC16, FGFR3, PABPC1, BIRC6, MUC4, KMT2D, IRS1, PGR, MECOM, MGAM, VAV1, TRRAP, BRD4, ABL2, TAL1, MITF, TP53, RPTOR, NRAS, NOTCH1, BRAF, TCL1A, LTB, CARD11, MACC1, TERT, KRAS	NRAS, APC, TYRO3, NOTCH1, RB1, ARID1B, FGFR3, BRAF, FANCD2, BRD4, ARID2, NF1, MITF, TP53, NFKBIA, KRAS, RPTOR	TCL1A, LTB, RPTOR, ABL2, TAL1, VAV1, NOTCH1	RPTOR, NF1, NFKBIA, NOTCH1	TCL1A, MGAM, ABL2, VAV1, PGR, BRD4	NFKBIA, APC, BRD4, BRAF

(B) List of tumor-suppressor genes (TSGs) and both oncogenes and driver genes (ODGs) in top-250 and top-500 genes

Top Genes	BDL-SP Model (Top-performing model)		CS-RF Model (Second best model)		CS-Cat Model (Third best model)	
	TSG	ODG	TSG	ODG	TSG	ODG
Previously reported TSGs and ODGs in MM and MGUS as found ranked in top-250 during the post-hoc analysis of the model	HLA-A, SP140, ARID2, PABPC1, CYLD, HLA-C, SAMHD1, SIRPA, SDHA, IRF1, NF1, MITF, TP53, ATP2B3, DIS3, KMT2C, NOTCH1, LTB, HLA-B, TRAF3, EGR1	NRAS, FGFR3, BRAF, LTB, PABPC1, TP53, KRAS	NCOR1, LTB, CYLD, NFKBIA, NF1, EGR1, TRAF3, NOTCH1, SDHA, KMT2C	LTB	NCOR, CYLD, NFKBIA, APC, MAX	BRAF
Previously reported TSGs and ODGs in MM and MGUS as found ranked in top-500 during the post-hoc analysis of the model	KMT2B, AMER1, RB1, ARID1B, FANCD2, HLA-A, CMTR2, SP140, ARID2, PABPC1, CYLD, MAX, HLA-C, SAMHD1, NCOR1, KMT2D, SIRPA, TERT, SDHA, IRF1, NF1, WNK2, MITF, ATP2B3, TP53, DIS3, ZFHX3, KMT2C, APC, NOTCH1, LTB, HLA-B, ACVR1B, NFKBIA, TRAF3, MYH11, EGR1	NRAS, FGFR3, TRRAP, BRAF, LTB, PABPC1, TP53, KRAS	NCOR1, LTB, CYLD, NFKBIA, NF1, EGR1, TRAF3, NOTCH1, SDHA, KMT2C	LTB	NCOR, CYLD, NFKBIA, APC, MAX	BRAF

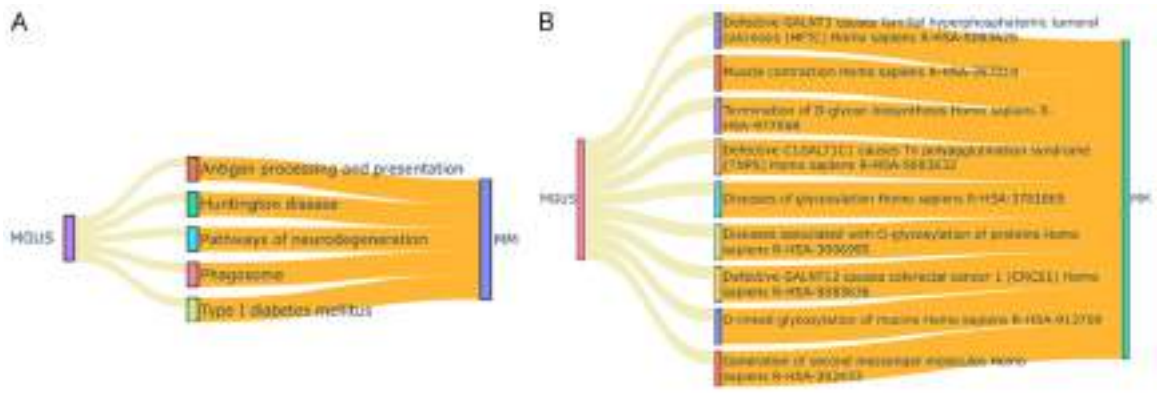


Figure 6. Pathway enrichment analysis of the top-genes obtained from BDL-SP model. A. KEGG Pathways that gained more significance during progression from MGUS to MM. B. Reactome Pathways gained more significance during the progression from MGUS to MM. Here in both of the figures, pale golden and orange ribbon means significant p-adjusted value (<0.05); orange refers to more significant and pale golden color refers to less significant.

In Category-2, no pathway became less significant with disease progression in KEGG and in Reactome. In Category-3, 103 KEGG pathways and 125 Reactome pathways were observed as significantly altered only in MM and not in MGUS (**Figures 7** and **8**). We further observed that 14 out of 103 KEGG pathways and 14 out of 125 Reactome pathways had no overlapping genes with the set of 199 significantly altered genes in MGUS. Finally, in Category-4, no KEGG pathway, but one Reactome pathway was observed as significantly altered only in MGUS and not in MM (**Figure 9**). Further, we observed that several signaling pathways such as Calcium signaling, B-cell receptor signaling, MAPK signaling pathway, regulation of actin cytoskeleton, etc. were significantly altered only in MM (p-adjusted value >0.05) and were not observed to be significantly altered in MGUS. The KEGG pathways that were significantly involved in disease progression from MGUS to MM with highlighted top-ranking genes identified by BDL-SP are shown in **Figure 10**.

Explainability of the BDL-SP model using ShAP algorithm

We utilized the ShAP algorithm for post-hoc model explainability and to rank genomic attributes based on their contribution to the model prediction. Here, each genomic attribute was assigned a ShAP score based on their contribution to each class (MM/MGUS) and has been ranked at the group-level (MM versus MGUS) and sample-level accordingly. We conducted the ShAP analysis for post-hoc explainability of the trained model in three different ways as explained in the subsequent sections.

Ranking of genes at the group-level from the explainability analysis of BDL-SP using ShAP: Based on the best ShAP score estimated for each genomic attribute using the algorithms shown in **Figure 4A** and **4B**, we ranked all the significantly altered genes at the group-level (MM/MGUS) to identify the top genes that significantly contributed to the model's prediction. The gene ranking of all 824 genes at group-level is shown in the beeswarm plot in Table S6 of supplementary material. In the beeswarm plot, each sample is represented as a dot, and the color of each dot corresponds to the best ShAP score of the gene. We have also highlighted all the previously reported genes of high relevance in MM in the beeswarm plot. In our analysis *KIR3DL2, IGLL5*, and *FCGR2A* are observed to be the top three genes based on their best ShAP scores in MGUS and MM samples from among the 824 significantly altered genes. Several previously reported driver genes in MM such as *IGLL5, HLA-A, KRAS, LTB, TP53, EGR1, FGFR3, NFKBIA, IRF1, NRAS*, etc. are observed in these top-ranked genes. Similarly, the previously reported oncogenes such as *CARD11, NOTCH1, VAV1, IRS1, MGAM, ABL2*, etc., and tumor-suppressor genes such as *HLA-B, HLA-C, SDHA*, etc. are observed in the top-ranked genes in our analysis. Also, many actionable genes are observed among the top genes, such as *KRAS, NOTCH1, TP53, FGFR3, ARID1B*, etc.

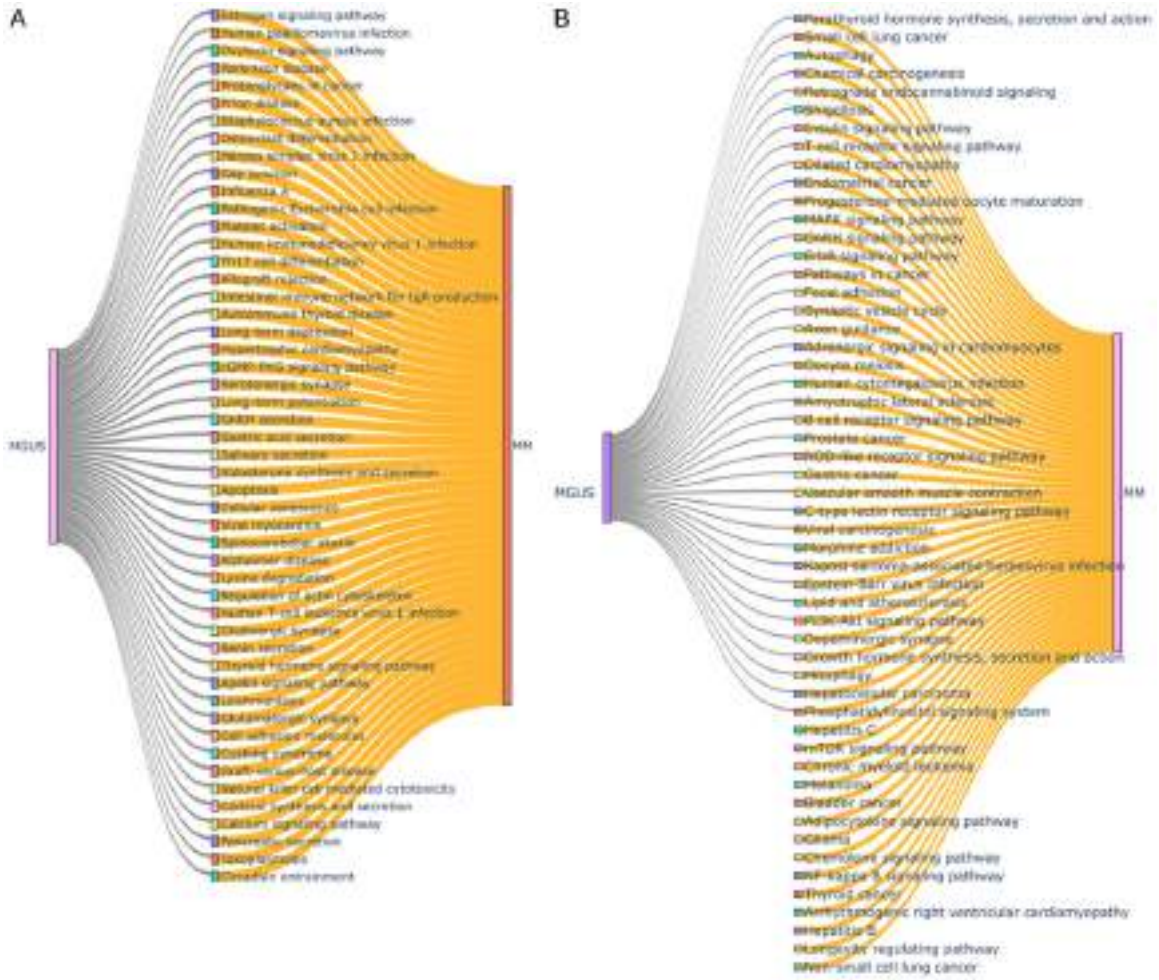


Figure 7. A, B. Pathway enrichment analysis of the top-genes obtained from BDL-SP model: KEGG Pathways that are uniquely significant in MM. In the above figure, orange ribbon means significant p-adjusted value (<0.05) and gray color refers to non-significant (p-adjusted value >0.05). There were a total of 108 KEGG pathways observed as significantly altered. Due to the large number of altered pathways, the above river plot was splitted into two parts to get more clarity.

Ranking of genes at the sample-level from the explainability analysis of BDL-SP using ShAP: In the sample-level analysis, we ranked genes found significantly altered in a sample according to their best ShAP scores estimated using the algorithm shown in **Figure 4A** and Algorithm A of **Table 2**. We observed that several previously reported OGs, TSGs, ODGs, and AGs were found in the top-ranked gene list of each sample. On assessing the ShAP scores of top significantly altered genes among all MM and MGUS samples, we observed that the mean ± standard deviation of the 100th ranked gene's ShAP score for all MM and MGUS samples is 0.0174±0.0037 and 0.0171±0.0040, respectively. Further, the ShAP score reduced to a considerably low value as we moved to a lower rank. Hence, we considered the top-100 significantly altered genes from all MM and MGUS samples based on their best ShAP scores for further analysis. The violin distribution plots for four gene groups of previously reported genes for all MM versus MGUS samples, only MGUS samples of EGA and AIIMS datasets, and only MM samples of MMRF and AIIMS datasets are shown in **Figure 11A-C**, respectively.

Analysis in MM & MGUS samples with ethnicity: We performed the statistical comparison of the disease stages (MM/MGUS) across American, European, and Indian populations (as mentioned in Section-2.1) on the basis of the number of previously reported genes in four gene groups using unpaired Wilcoxon ranksum

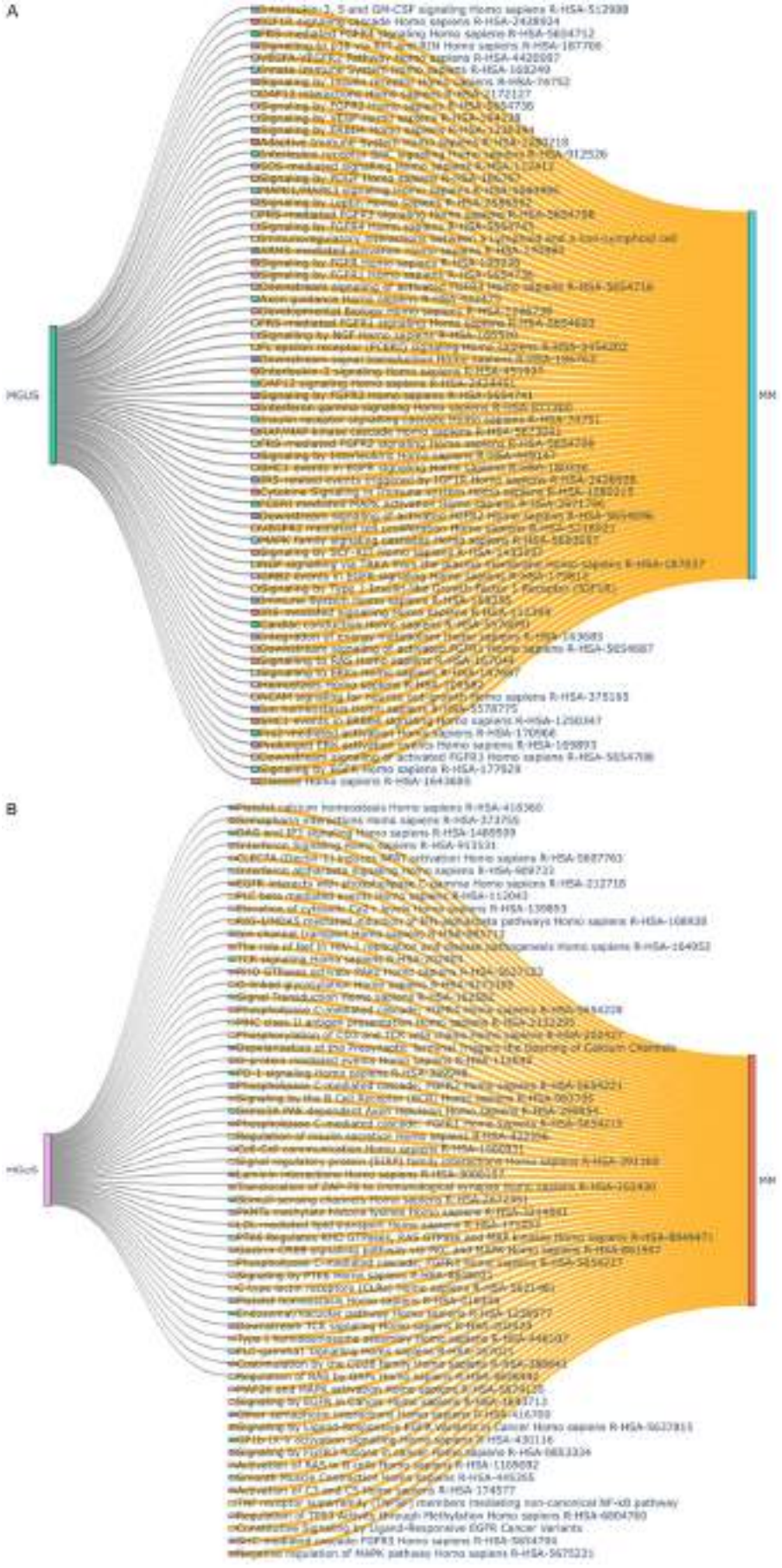

A
MGUS
MM
B
MGUS
MM

Figure 8. A. Pathway enrichment analysis of the top-genes obtained from BDL-SP model: Reactome Pathways that are uniquely significant in MM. In the above figure, orange ribbon means significant p-adjusted value (<0.05) and gray color refers to non-significant (p-adjusted value >0.05). There were a total of 134 Reactome pathways observed as significantly altered. Due to the large number of altered pathways, the above river plot was splitted into two parts to get more clarity. B. Pathway enrichment analysis of the top-genes obtained from BDL-SP model: Reactome Pathways that are uniquely significant in MM. In the above figure, orange ribbon means significant p-adjusted value (<0.05) and gray color refers to non-significant (p-adjusted value >0.05). There were a total of 134 Reactome pathways observed as significantly altered. Due to the large number of altered pathways, the above river plot was splitted into two parts to get more clarity.

Figure 9. Pathway enrichment analysis of the top-genes obtained from BDL-SP model: Reactome Pathways that are uniquely significant in MGUS. In the above figure, orange ribbon means significant p-adjusted value (<0.05) and gray color refers to non-significant (p-adjusted value >0.05).

test. We observed that the number of genes in OG, ODG, and AG gene groups are significantly different between the disease stages (MM and MGUS) in the analysis of combined data of different geographic populations (**Figure 11A**). Further, the medians of the number of genes in the OG, TSG, and AG gene groups were observed to be higher than the respective medians in the precursor stage (MGUS) (**Figure 11A**). Similarly, comparing the number of genes in all four gene groups between the MGUS samples of Indian (AIIMS dataset) and European (EGA dataset) population, the number of genes in OG, ODG, and AG gene groups were observed to be significantly different (**Figure 11B**). On the contrary, the number of genes in all four gene groups were not found to be statistically and significantly different in MM samples of Indian (AIIMS dataset) and American population (MMRF dataset) (**Figure 11C**). These observations indicate that ethnicity might be playing a significant role in the disease development and thus, ethnicity-specific analysis can be helpful in further gaining in-depth insights into the disease biology of the premalignant stage of MM (MGUS).

Genomic feature ranking at a sample-level using ShAP analysis: Besides identifying the top-significantly altered genes in MM and MGUS, we also ranked the genomic features based on their contribution in the model prediction. A set of 28 genomic features (**Figure 1**) were used to train the BDL-SP model. These genes were ranked on the basis of their ShAP scores. The algorithm for estimating the best ShAP score for each genomic feature is shown in **Figure 4B**. We observed that the total number of SNVs, total number of SNVs in the *Other SNV group* (as shown in **Figure 1**), and VAF's standard deviation of SNVs in the *Other SNV group* were the top three genomic features, while VAF's standard deviation of SNVs in the *Non-synonymous SNV group*, VAF's standard deviation of SNVs in the *Synonymous SNV group*, and AD's standard deviation of SNVs in the *Non-synonymous SNV group* were the least contributing genomic features. The beeswarm plot for genomic feature ranking from BDL-SP model post-hoc analysis using ShAP is shown in **Figure 12**.

Discussion

It is well established that MM evolves through premalignant stages driven by the acquisition of multiple genomic aberrations [60]. Though a few studies have analyzed the progression from MGUS to MM [10, 13], a limited amount of information is available on the notable biomarkers responsible for this transformation. However, if known apriori, appropriate treatment at the MGUS stage can help control the progression of MGUS to MM, thereby preventing the complications associated with MM, reducing morbidity, and increasing the overall survival of these patients. Thus, it is crucial to unravel the genomic features responsible for the malignant transformation of MGUS to MM.

In this work, we addressed the challenge of extracting relevant MM and MGUS differentiat-

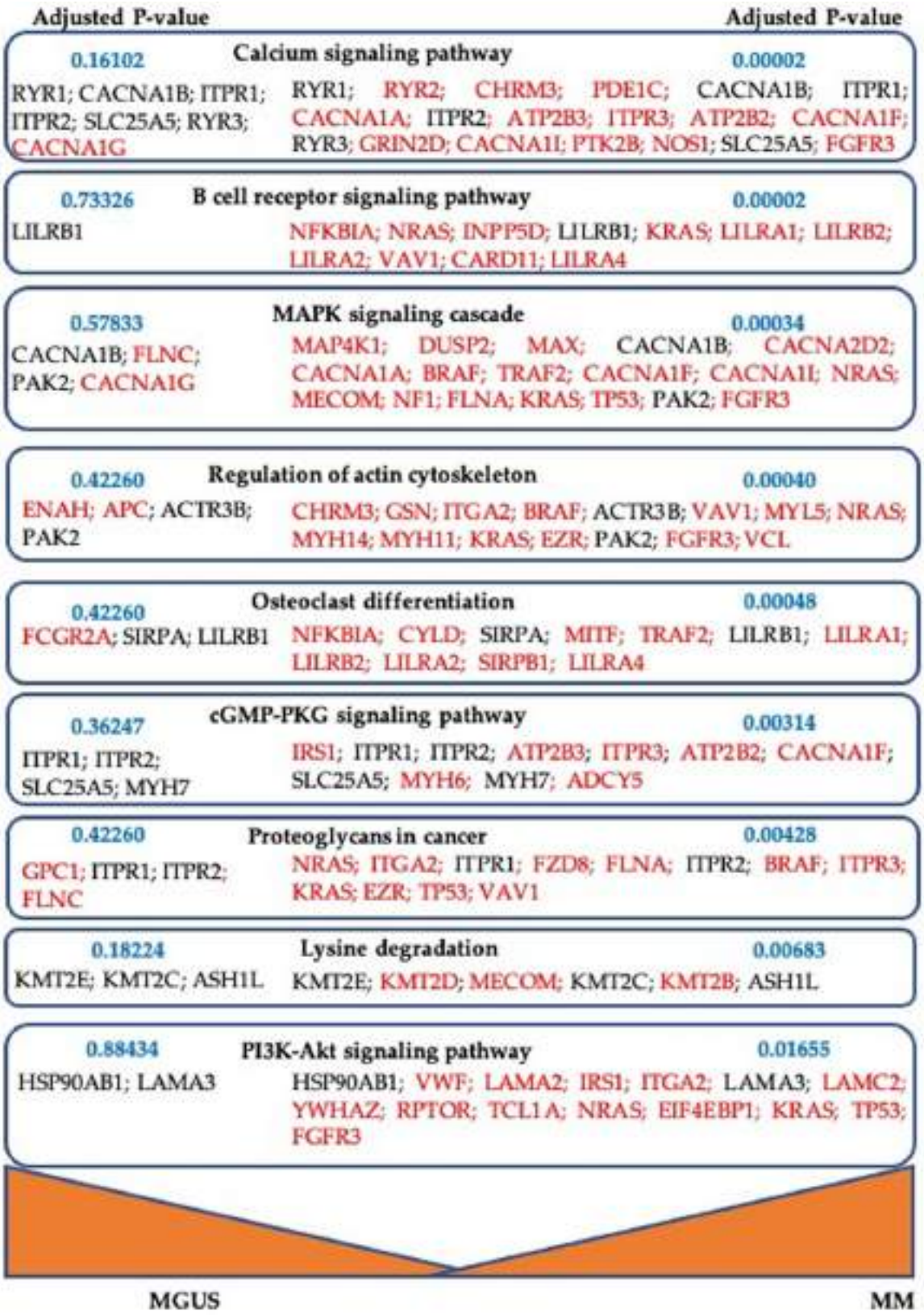
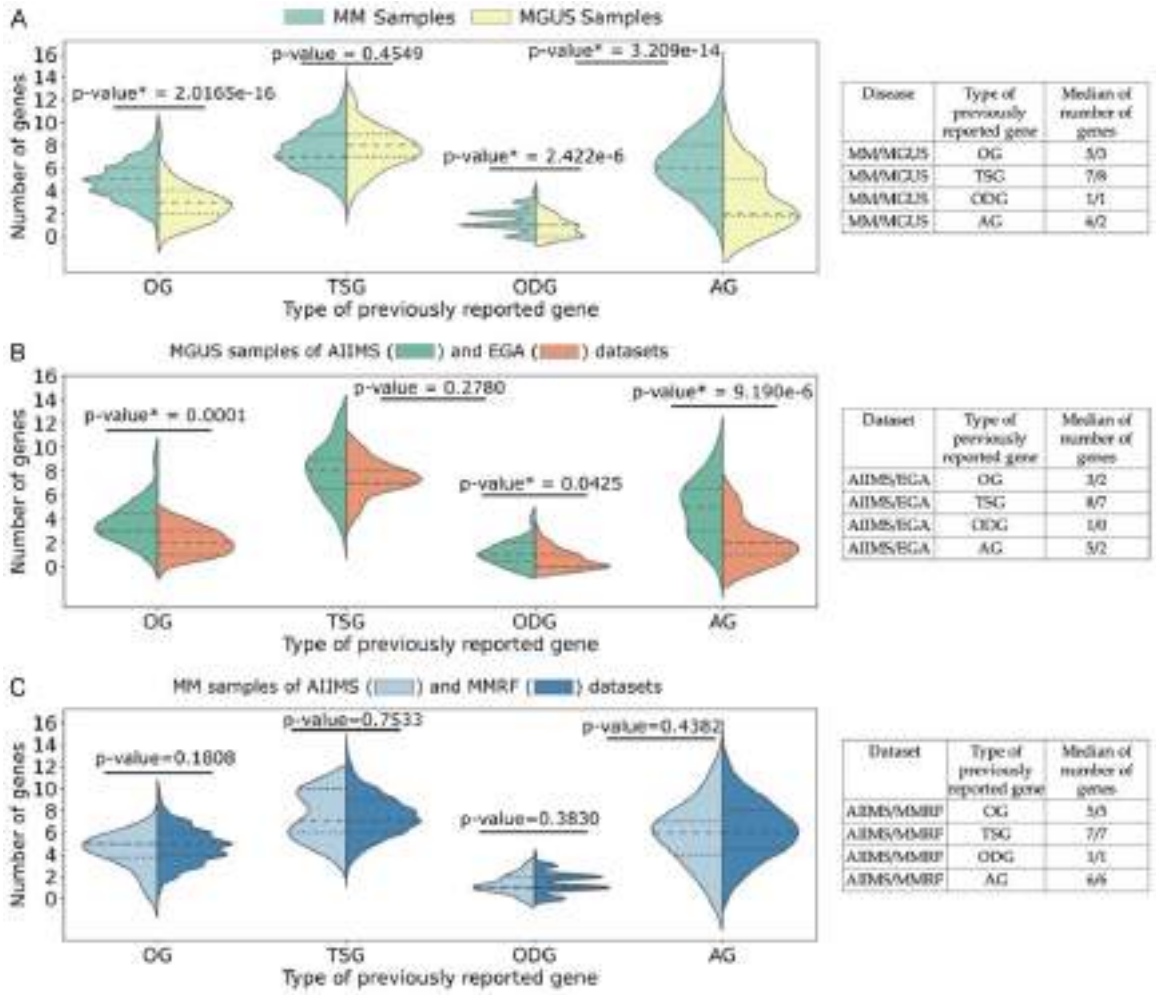


Figure 10. KEGG pathways found to be significantly involved in progression of MGUS to MM. Top genes that were identified by post-hoc analysis of BDL-SP using ShAP algorithm as significantly mutated either in MGUS only or in MM only (acting as differentiators of MGUS and MM) are shown in red colored font.



Disease	Type of previously reported gene	Median of number of genes
MM/MGUS	OG	5/3
MM/MGUS	TSG	7/8
MM/MGUS	ODG	1/1
MM/MGUS	AG	4/2

Dataset	Type of previously reported gene	Median of number of genes
AIIMS/EGA	OG	3/2
AIIMS/EGA	TSG	8/7
AIIMS/EGA	ODG	1/0
AIIMS/EGA	AG	5/2

Dataset	Type of previously reported gene	Median of number of genes
AIIMS/MMRF	OG	5/3
AIIMS/MMRF	TSG	7/7
AIIMS/MMRF	ODG	1/1
AIIMS/MMRF	AG	6/6

Figure 11. A. The distribution of the number of previously reported genes in four gene groups (OGs, TSGs, ODGs, and AGs) found significantly altered and ranked in top-100 across all MM and MGUS samples (combined dataset of MMRF, EGA, and AIIMS samples). B. The distribution of the number of previously reported genes found significantly altered and ranked in top-100 across all MGUS samples in EGA and AIIMS datasets. C. The distribution of the number of previously reported genes found significantly altered and ranked in top-100 across all MM samples in MMRF and AIIMS datasets (OG: Oncogenes, TSG: Tumor-Suppressor Genes, ODG: Both oncogenes and driver genes, AG: Actionable Genes). The *P*-value shown with each violin plot was estimated using unpaired Wilcoxon ranksum statistical test to check whether the number of genes in a particular gene group is significantly different from their respective counts in the other group. The gene group having *P*-value with superscript "*" (star) symbol represents that the number of genes in that gene group are significantly different compared to the other group. The table on the right of each figure shows the median of the number of genes in each gene group for disease stages (MM/MGUS) and datasets (MMRF, EGA, and AIIMS). Note: To have a better view of the violin plots, refer to the colored version of this figure.

ing genomic attributes from the pool of a large mutational information available for each patient. Our proposed BDL-SP based workflow has been successful in accomplishing this task. In the pre-processing of the data, we identified significantly mutated genes for each variant caller and then took the union of them so that we do not miss any important gene. Thus, a large cohort size and an ensemble of four variant callers enabled us to obtain generalizable mutational information, driver genes, and altered pathway information. Recently, graph-based learning has been extensively explored for deciphering crucial information such as disease progression, identification of novel biomarkers for targeted drug therapy, etc. in genomics. For example, the graph-based model was used to learn the temporal graphs of diagnosis (Dx), procedure (Px) and prescription (Rx) of multiple myeloma patients from the sequen-

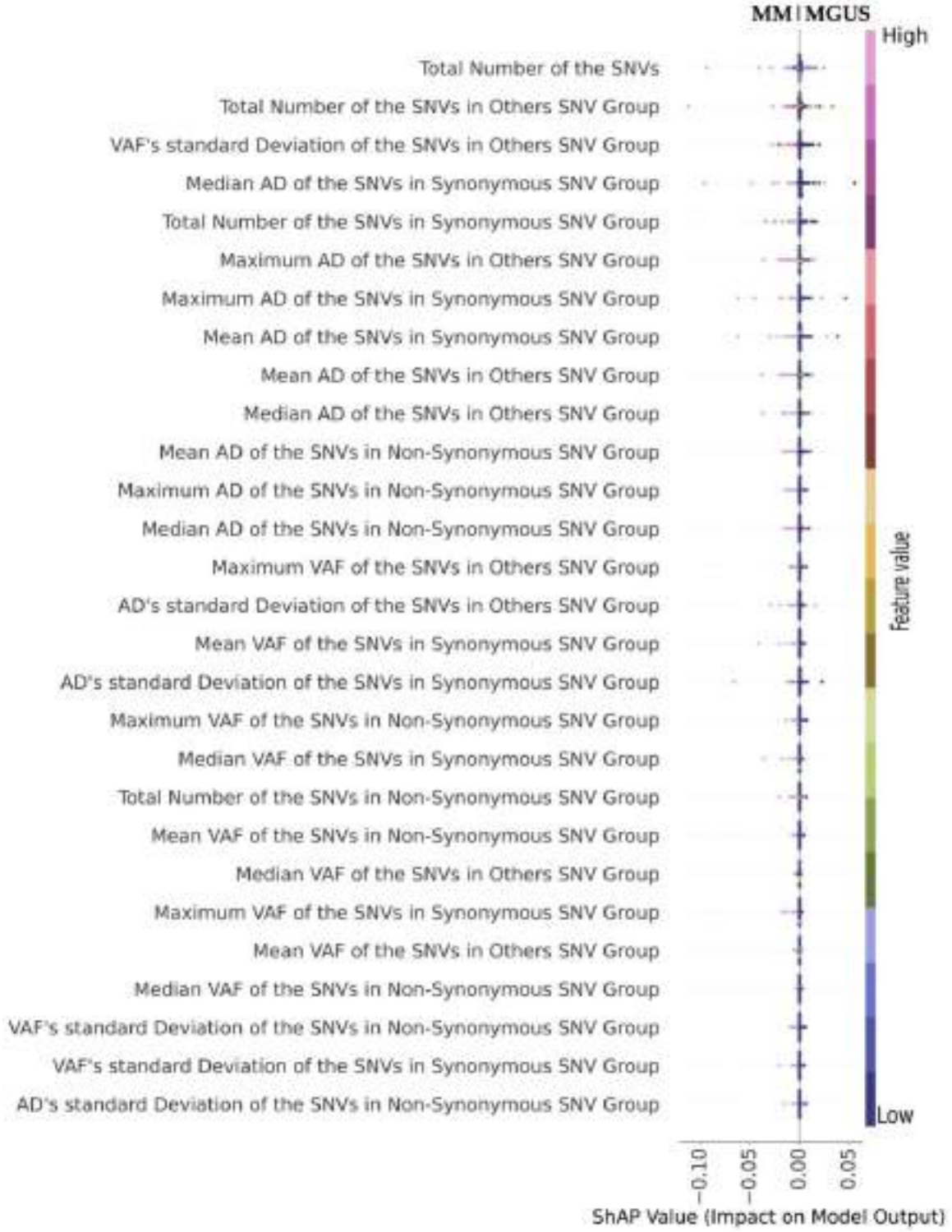


Figure 12. Genomic feature ranking based on the BDL-SP model's post-hoc explainability in MM and MGUS using ShAP algorithm. Each genomic feature is ranked according to their best ShAP score estimated using the algorithm shown in **Figure 4** and **Table 2**. The negative ShAP score represents the contribution of the genomic feature towards MM, while the positive ShAP score represents the contribution of the genomic feature towards MGUS. Further, each dot in the individual scatter plot of the genomic feature represents a sample and the color of dot represents the value of that genomic feature with the color-codes as follows: the dark blue color represents low and the pink color represents high value of the genomic feature. Note: Refer to the colored version of this figure for a clear view of the sample distribution for each genomic feature.

tial electronic health records (EHR) and predict a patient's response to treatment [61]. Till now, graph-based learning approaches have not yet been explored to identify the underlying difference between MM and its precursor stage (MGUS). In our BDL-SP model, we have used the connectionist model of graph-based learning to learn genomic mutational profiles (as node features) that were extracted from the WES datasets of AIIMS, EGA, and MMRF. We additionally utilized the gene-gene interaction information from PPI network to identify the pivotal biomarkers that can differentiate MM from MGUS.

Our proposed AI-based BDL-SP workflow is innovative in multiple ways as explained below:

1. The identification of pivotal biomarkers using WES datasets of MM of three populations (American, European, and Indian) increases the robustness of the workflow by enhancing its ability to assess the granular-level insights of mutational profiles from multiple datasets of different geographic locations/ethnicities.

2. Because of the pathogenic nature of deleterious SNVs, only deleterious SNVs were considered for identifying the significantly altered genes in the proposed workflow. We observed that the total number of SNVs were reduced considerably after variant filtration of benign SNVs using the FATHMM-XF [49] method.

3. An analysis of the genomic mutational profile along with the gene-gene interaction information enables this workflow to look at interdependencies between genes, making it a complete bio-inspired workflow.

4. The proposed workflow includes quantitative (using performance metrics) as well as an exhaustive qualitative (post-hoc interpretability analysis of the trained models) benchmarking. It also shows that multiple ML models behaving closely on the quantitative metrics may differ hugely in the qualitative analysis. Thus, application-aware interpretability analysis as carried out in this workflow (ShAP on genes and genomic features) can help in choosing the right model and increase the confidence of the doctors on the trained AI model.

The complete list of top significantly altered genes identified by the best three performing models (BDL-SP, CS-RF, and CS-Cat) is provided in **Table 5**. Of all, our proposed BDL-SP model identified the highest number of previously reported OGs, TSGs, ODGs, and AGs compared to the other standard ML methods. This shows that our GCN-based BDL-SP workflow is indeed capable of extracting the differentiating genomic features robustly that are otherwise difficult to obtain. Many of the top-ranking genes in the present study included known cancer drivers (*IGLL5, HLA-A, KRAS, LTB*, etc.), oncogenes (*KRAS, NRAS, FGFR3, BRAF*, etc.), tumor-suppressor genes (*HLA-A, LTB, TRAF3, EGR1, SAMHD1, DIS3, ARID2, CYLD, SP140*, etc.) and actionable genes (*KRAS, TP53, NF1, NFKBIA, ARID2*, etc.) having high relevance in MM. Interestingly, some TSGs (*HLA-B/C, NOTCH1, SDHA, MITF, ARID1B, FANCD2, KMT2D, APC, CMTR2* and *AMER1*) and oncogenes (*CARD11, NOTCH1, VAV1, IRS1, MGAM, ABL2, TCL1A, PGR, MITF, RPTOR, TERT, BRD4, MECOM*, and *TAL1*) that are so far not reported as drivers in MM, were also listed in the top ranking genes of BDL-SP. Further focused studies are required to validate the above finding and to check the functional status and other characteristics of these genes before classifying them as MM drivers.

Pathway analysis on MM and MGUS genes revealed that the MM related pathways such as MAPK, cGMP-PKG, B-cell receptor, etc. were not significantly altered in MGUS (adj *P*-value ≥0.05) and became significant in MM (adj *p*-value ≤0.05). We observed that several OGs, TSGs, ODGs, and AGs associated with the significantly altered pathways were found significantly altered only in the MM cohort and not in the MGUS cohort (See **Figure 13**). Here, the additional alterations in several previously reported genes such as *BRAF, FGFR3, IRS1, MAX, KRAS*, etc. assisted the malignant progression of MGUS to MM. Our pathway analysis also demonstrated that some pathways that lost their statistical significance from MGUS to MM were actually related to the other cancer types.

However, the results in our study are unique because we have demonstrated that these pathways are selectively and significantly dysregulated in MM compared to its precursor stage of MGUS due to a distinct set of genes that are differentially mutated in the two dis-

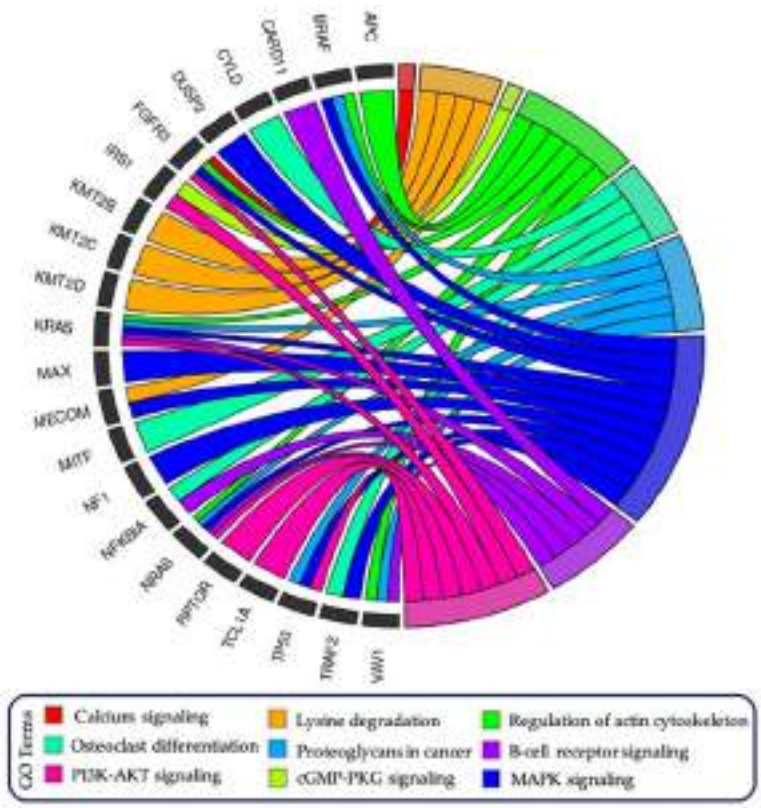


Figure 13. GOChord plot reveals the association of driver/TSG/Onco/Actionable genes associated with important pathways. The gene *KMT2C* was observed to be significantly mutated in MGUS and MM, while the gene *APC* was mutated only in MGUS. All other genes were observed to be significantly mutated in MM only.

eased stages. These observations warrant further investigations to decipher if any of these differentiating genes could become druggable targets especially during the early phase of MGUS. Some of the key genes and pathways that are selectively altered at the MM stage, such as *FGFR3, BRAF* and *MAP kinase* pathways, are actionable and hence, targeted therapy for them is under evaluation in clinical trials [62]. *FGFR3* is a partner gene in t(4;14) that has been observed as a significantly altered gene in all datasets of MM and MGUS. However, the poor prognostic impact of *FGFR3* has been linked to activating mutations in the *FGFR3* gene rather than the fusion event which exerts its influence via activation of *WHSC/MMSET* genes and is responsive to proteasome inhibitors [63, 64]. Besides, single case reports demonstrating efficacy of *BRAF* inhibitors in relapsed refractory MM with *BRAF* mutations, a recent report on NCI-Match trial shows promising results for *BRAF* inhibitors, Dabrafenib and Trametinib, in patients harboring tumors with *BRAF V600E* mutations including MM [65]. Many structural variations (SVs) observed in MM such as *IgH translocations, 1q gain, 1p del* are also observed at the MGUS stage. However, *C-MYC* alterations, which are predominantly structural variations, are secondary events and are seen in nearly 40% of MM patients [66]. Lack of analysis of structural variants is one of the limitations of this study. However, we did observe mutations in *MAX* at the MM stage, which is known to dimerize with *C-MYC* and influence the transcription of multiple genes and thus, the pathogenesis of MM [67].

The frequently observed complex genomic traits that can drive the disease progression from MGUS to MM can be 3'UTR/5'UTR mutations [12], copy number variants, structural variants [1, 60], and loss of the ability of the dysfunctional immune environment to control virulent cell clones [68]. Akin to higher levels of disease load in MM compared to its precursor states, measurable disease load, increased number of non-synonymous mutations in MM compared with MGUS [9-12] and increased levels of deregulated cytokines in relapsed refractory MM compared to treatment naïve MM has been reported [69].

In addition, Mikalasova et al. In MGUS, a positive correlation between the increasing chromosome changes and the somatic gene mutations, and absence of *MYC translocation* and *TP53 deletions* or mutations has been observed [11]. From the genomic profile analysis of paired MGUS-MM and SMM-MM samples, it has been observed that as the disease progressed, the number of NS mutations actually decreased in 70% samples. This observation is in contrast to reports on unpaired samples, where an increase in the NS mutations has

been reported from MGUS to MM [9-12]. Further, the comparisons of unpaired MGUS/SMM and MM samples have shown the mutational similarity of MGUS/SMM with MM [13]. Based on this observation, we hypothesize that the progression is associated with an altered landscape of acquired mutations, rather than an increased total mutational burden.

The post-hoc explainability of the BDL-SP model using ShAP algorithm revealed the top genomic attributes (genomic features and significantly altered genes) at both the group- and sample-levels. At group-level, all the 824 significantly altered genes were ranked according to their ShAP score using the algorithm shown in **Figure 4** (Table S6 of supplementary material) and top-500 genes were further compared with the literature (Table S7 of supplementary material). Several significantly altered genes found in our analysis were previously reported as driver genes in [7, 53, 54], oncogenes and TSGs in [52], and actionable genes in [55, 56], while some genes such as *KIR3DL2, FCGR2A, LILRB1/2, KIR2DL1/4* etc. were novel that contributed significantly in disease classification (See Table S6 of supplementary material). The *KIR* framework genes (*KIR3DL2/2DL4*) were among the top significantly altered genes with highest ShAP scores. The KIR gene complex on chromosome 19 encodes a series of inhibitory or activatory *KIR* genes expressed on *NK cells* [70-72]. These receptors serve as *HLA* ligands and modulate *NK* cell immune function against tumors [70]. A few activating genes in the *KIR* family (*KIR2DS4* and *KIR2DS5*) have been shown to have higher prevalence in MM patients [70] than healthy people. The *KIR*s have also been reported to influence the efficacy of therapies such as that of isatuximab [71]. The second topmost gene with the highest ShAP score was *IGLL5*. Again, the *IGLL5* gene undergoes point mutations and *IGLL5/IGH* translocations in MM [73]. Point mutations are largely mutually exclusive of *RAS* mutations and associated with greater risk of disease progression. Similarly, other genes such as *HLA-A/B/C, FCGR2A* and *LILRB1/2* reported in previous studies are also shown to have a significant role in MM [74-77]. Given the crucial role of these top immune related genes highlighted by the ShAP ranking in our study suggests their potential role as drivers of progression and disease stratifying biomarkers.

We have also highlighted the impact of ethnicity (**Figure 11**) among three groups of American (MMRF), European (EGA), and Indian (AIIMS) population. The number of OGs, ODGs, and AGs were significantly different in the MM samples of American and Indian population, and MGUS samples of European and Indian population (**Figure 11A**). Also, the median of the number of OGs and AGs increased with the disease progression from MGUS to MM. This increase could be due to the increasingly active participation of OGs and AGs in disease progression from MGUS to MM. Similarly, the number of OGs, ODGs, and AGs are significantly different in the MGUS samples of Indian population and MGUS samples of European population (**Figure 11B**). Here, we also observed the increasingly active participation of OGs, ODGs, and AGs in MGUS samples of Indian and European population. On the other hand, the number of previously reported genes (OGs, ODGs, TSGs, and AGs) present in the MM samples of the American and Indian population were not found to be statically different (**Figure 11C**). These observations indicate that the impact of ethnicity on disease biology can not be overlooked and might be an important factor during the initial phase or development phase of MM. Further analysis of ethnicity-specific information to infer the responsible prognostic factor for disease development and progression is strongly suggested.

The sample-wise gene-ranking highlighted their contribution at the individual sample-level. The study in [11] showed that the transition from MGUS to MM is due to the acquisition of mutations in critical driver genes and oncogenes. Interestingly, we have observed that not only driver genes and oncogenes, but several TSGs and actionable genes were also altered significantly in MGUS (**Figure 11**). Further, the role of oncogenes increased as disease progressed from MGUS to MM. On comparing the top contributing genomic features in MGUS and MM samples, we observed that the genomic features related to the *Synonymous SNVs group* (a group of UTR3, synonymous, and UTR5 type SNVs) and the *Other SNVs group* (a group of Non-frameshift insertion/deletion/substitution, intronic, intergenic, ncRNA_intronic, upstream, downstream, unknown, and ncRNA_splicing SNVs) contributed largely in disease classification as compared to the genomic fea-

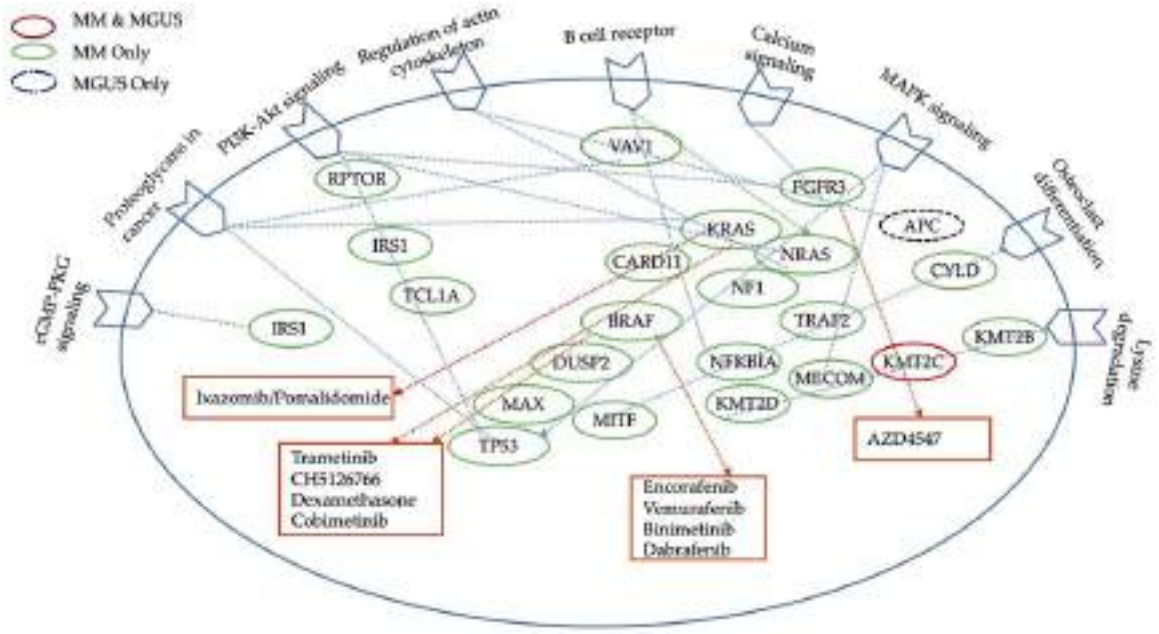


Figure 14. Important pathways significantly altered in MM. Drugs used for pathway-directed therapies associated with mutations in genes are also shown with red colored text-boxes and arrows.

tures of the *Non-synonymous SNVs group* (**Figure 12**).

Although the role of synonymous SNVs are unclear in MM, recent studies have observed these synonymous SNVs as significant contributors in multiple cancer types [78-82]. Further exploration of differentially affected biological pathways may provide the pathogenic link between MM, its precursor (MGUS or SMM), and overt disease stages so as to find appropriate targeted therapy to halt the progression from precursor to stage to MM (**Figure 14**). We have shown in the current study that the incremental accumulation of key mutations tilts the balance of biological pathways in favor of progression from the state of MGUS to MM in a large cohort of unpaired MGUS-MM samples. Some of these pathways are actionable and, targeting them may enable us to reverse the balance in favor of a controlled and relatively indolent clinical course. Further, AI-based workflow has assisted in successfully differentiating MGUS from MM. We have shown in our study that our trained machine learning (ML) classifiers are able to identify pivotal genomic biomarkers helpful in distinguishing MM and MGUS, thereby, leading to a better understanding of malignant transformation of MGUS to MM and prognostication.

Conclusion

MGUS and MM share many common features such as genomic biomarkers and structural variants, although MGUS has a relatively less complex genomic profile than MM. Therefore, it is a challenging task to distinguish MM from MGUS. In our proposed work, we have presented an innovative, bio-inspired AI-based workflow BDL-SP to identify pivotal genomic biomarkers to distinguish MGUS from MM. The proposed graph convolutional network based BDL-SP model is able to extract discriminative genomic biomarkers for identifying MM and MGUS samples. BDL-SP outperformed all the baseline ML-based models. Further, using the application-aware interpretability analysis of the trained AI model, we have demonstrated a way to identify the best AI model from among the multiple machine learning or deep learning models that may have performed similarly on the quantitative metrics on the available data. In the post-hoc interpretability benchmarking, BDL-SP outperformed all the baseline models by identifying the highest number of previously reported genes such as *KRAS, BRAF, LTB, NRAS, FGFR3, NF1, NFKBIA, ARID2, RB1, HLA-A, TP53, SP140, TRAF3, EGR1, IRF1, SAMHD1, DIS3, CYLD, KMT2B/C, MAX, ZFHX3 and NCOR1,* that are of high relevance in MM.

Further, some of the genes that acted as differentiable biomarkers included TSGs (*HLA-B/C, NOTCH1, SDHA, MITF, ARID1B, FANCD2, KMT2D, APC, CMTR2*, and *AMER1*) and oncogenes (*CARD11, NOTCH1, VAV1, IRS1, MGAM, ABL2, TCL1A, PGR, MITF, RPTOR, TERT, BRD4, MECOM*, and *TAL1*) that have not yet been identified as MM drivers. These require validation by future studies before being declared as MM drivers. We further validated our findings by performing pathway analysis on the top mutated genes. It was inferred from the pathway analysis that several signaling pathways such as Calcium signaling pathway, B-Cell receptor signaling pathway, PI3K-Akt signaling pathway, MAPK signaling pathway, etc. are selectively and more significantly dysregulated with disease progression. Additional mutations in driver genes, critical oncogenes, tumor-suppressor genes, and actionable genes fostered the transformation of benign MGUS to MM. Similarly, the genomic mutation associated with the Synonymous SNV group (synonymous SNVs, UTR3, and UTR5) were found to be the most significantly contributing biomarker differentiating MM from MGUS. These observations may hold great significance from a therapeutic point of view. We observed that the number of oncogenes, driver genes, and actionable genes in the MGUS samples of European and Indian populations were statistically different. Although no population specific differences were observed in our analysis of the MM data consisting of the American and Indian population, the results on MGUS data indicates that the impact of ethnicity on the disease biology of MM should be further explored.

Acknowledgements

This work was supported by grant from Department of Biotechnology, Govt. of India [Grant: BT/MED/30/SP11006/2015] and Department of Science and Technology, Govt. of India [Grant: DST/ICPS/CPS-Individual/2018/279(G)]. Authors acknowledge dbGaP (Project# 18964) for providing authorized access to the MM datasets (phs000748 and phs000348). We also acknowledge EGA (EGAD00001001901) for providing authorized access to the MGUS data. Authors would also like to thank the Centre of Excellence in Healthcare, IIIT-Delhi for their support in this research. Authors acknowledge the valuable insights provided by Dr. Sriram K. (Department of Computational Biology, IIIT-Delhi) and Akanksha Farswan (SBI Lab, Department of ECE, IIIT-Delhi) that greatly improved the AI-based workflow.

Voluntary written informed consent was obtained from all the study individuals.

Disclosure of conflict of interest

None.

Address correspondence to: Anubha Gupta, SBI Lab, Department of Electronics and Communication Engineering & Centre of Excellence in Healthcare, Indraprastha Institute of Information Technology-Delhi (IIIT-D), New Delhi 110020, India. ORCID: 0000-0002-7752-1926; E-mail: anubha@iiitd.ac.in; Ritu Gupta, Laboratory Oncology Unit, Dr. B.R.A. IRCH, All India Institute of Medical Sciences (AIIMS), Ansari Nagar, New Delhi 110029, India. ORCID: 0000000153644086; E-mail: drritugupta@gmail.com; drritu.laboncology@aiims.edu

References

[1] Manier S, Salem K, Glavey SV, Roccaro AM and Ghobrial IM. Genomic aberrations in multiple myeloma. Cancer Treat Res 2016; 169: 23-34.

[2] Chapman MA, Lawrence MS, Keats JJ, Cibulskis K, Sougnez C, Schinzel AC, Harview CL, Brunet JP, Ahmann GJ, Adli M, Anderson KC, Ardlie KG, Auclair D, Baker A, Bergsagel PL, Bernstein BE, Drier Y, Fonseca R, Gabriel SB, Hofmeister CC, Jagannath S, Jakubowiak AJ, Krishnan A, Levy J, Liefeld T, Lonial S, Mahan S, Mfuko B, Monti S, Perkins LM, Onofrio R, Pugh TJ, Rajkumar SV, Ramos AH, Siegel DS, Sivachenko A, Stewart AK, Trudel S, Vij R, Voet D, Winckler W, Zimmerman T, Carpten J, Trent J, Hahn WC, Garraway LA, Meyerson M, Lander ES, Getz G and Golub TR. Initial genome sequencing and analysis of multiple myeloma. Nature 2011; 471: 467-472.

[3] Bolli N, Avet-Loiseau H, Wedge DC, Van Loo P, Alexandrov LB, Martincorena I, Dawson KJ, Iorio F, Nik-Zainal S, Bignell GR, Hinton JW, Li Y, Tubio JM, McLaren S, O'Meara S, Butler AP, Teague JW, Mudie L, Anderson E, Rashid N, Tai YT, Shammas MA, Sperling AS, Fulciniti M, Richardson PG, Parmigiani G, Magrangeas F, Minvielle S, Moreau P, Attal M, Facon T, Futreal PA, Anderson KC, Campbell PJ and Munshi NC. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. Nat Commun 2014; 5: 2997.

[4] Walker BA, Boyle EM, Wardell CP, Murison A, Begum DB, Dahir NM, Proszek PZ, Johnson DC, Kaiser MF, Melchor L, Aronson LI, Scales M, Pawlyn C, Mirabella F, Jones JR, Brioli A, Mikulasova A, Cairns DA, Gregory WM, Quartilho A, Drayson MT, Russell N, Cook G, Jackson GH, Leleu X, Davies FE and Morgan GJ. Mutational spectrum, copy number changes, and outcome: results of a sequencing study of patients with newly diagnosed myeloma. J Clin Oncol 2015; 33: 3911-20.

[5] Lohr JG, Stojanov P, Carter SL, Cruz-Gordillo P, Lawrence MS, Auclair D, Sougnez C, Knoechel B, Gould J, Saksena G, Cibulskis K, McKenna A, Chapman MA, Straussman R, Levy J, Perkins LM, Keats JJ, Schumacher SE, Rosenberg M; Multiple Myeloma Research Consortium, Getz G and Golub TR. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer Cell 2014; 25: 91-101.

[6] Farswan A, Jena L, Kaur G, Gupta A, Gupta R, Rani L, Sharma A and Kumar L. Branching clonal evolution patterns predominate mutational landscape in multiple myeloma. Am J Cancer Res 2021; 11: 5659-5679.

[7] Walker BA, Mavrommatis K, Wardell CP, Ashby TC, Bauer M, Davies FE, Rosenthal A, Wang H, Qu P, Hoering A, Samur M, Towfic F, Ortiz M, Flynt E, Yu Z, Yang Z, Rozelle D, Obenauer J, Trotter M, Auclair D, Keats J, Bolli N, Fulciniti M, Szalat R, Moreau P, Durie B, Stewart AK, Goldschmidt H, Raab MS, Einsele H, Sonneveld P, San Miguel J, Lonial S, Jackson GH, Anderson KC, Avet-Loiseau H, Munshi N, Thakurta A and Morgan GJ. Identification of novel mutational drivers reveals oncogene dependencies in multiple myeloma. Blood 2018; 132: 587-597.

[8] Kaur G, Jena L, Gupta R, Farswan A, Gupta A and Sriram K. Correlation of changes in subclonal architecture with progression in the MMRF CoMMpass study. Transl Oncol 2022; 23: 101472.

[9] Mikulasova A, Smetana J, Wayhelova M, Janyskova H, Sandecka V, Kufova Z, Almasi M, Jarkovsky J, Gregora E, Kessler P, Wrobel M, Walker BA, Wardell CP, Morgan GJ, Hajek R and Kuglik P. Genomewide profiling of copy-number alteration in monoclonal gammopathy of undetermined significance. Eur J Haematol 2016; 97: 568-575.

[10] Walker BA, Wardell CP, Melchor L, Brioli A, Johnson DC, Kaiser MF, Mirabella F, Lopez-Corral L, Humphray S, Murray L, Ross M, Bentley D, Gutiérrez NC, Garcia-Sanz R, San Miguel J, Davies FE, Gonzalez D and Morgan GJ. Intraclonal heterogeneity is a critical early event in the development of myeloma and precedes the development of clinical symptoms. Leukemia 2014; 28: 384-390.

[11] Mikulasova A, Wardell CP, Murison A, Boyle EM, Jackson GH, Smetana J, Kufova Z, Pour L, Sandecka V, Almasi M, Vsianska P, Gregora E, Kuglik P, Hajek R, Davies FE, Morgan GJ and Walker BA. The spectrum of somatic mutations in monoclonal gammopathy of undetermined significance indicates a less complex genomic landscape than that in multiple myeloma. Haematologica 2017; 102: 1617-1625.

[12] Farswan A, Gupta A, Jena L, Ruhela V, Kaur G and Gupta R. Characterizing the mutational landscape of MM and its precursor MGUS. Am J Cancer Res 2022; 12: 1919-1933.

[13] Dutta AK, Fink JL, Grady JP, Morgan GJ, Mullighan CG, To LB, Hewett DR and Zannettino ACW. Subclonal evolution in disease progression from MGUS/SMM to multiple myeloma is characterised by clonal stability. Leukemia 2019; 33: 457-468.

[14] Mosquera Orgueira A, González Pérez MS, Díaz Arias JÁ, Antelo Rodríguez B, Alonso Vence N, Bendaña López Á, Abuín Blanco A, Bao Pérez L, Peleteiro Raíndo A, Cid López M, Pérez Encinas MM, Bello López JL and Mateos Manteca MV. Survival prediction and treatment optimization of multiple myeloma patients using machine-learning models based on clinical and gene expression data. Leukemia 2021; 35: 2924-2935.

[15] Venezian Povoa L, Ribeiro CHC and Silva ITD. Machine learning predicts treatment sensitivity in multiple myeloma based on molecular and clinical information coupled with drug response. PLoS One 2021; 16: e0254596.

[16] Farswan A, Gupta A, Gupta R, Hazra S, Khan S, Kumar L and Sharma A. AI-supported modified risk staging for multiple myeloma cancer useful in real-world scenario. Transl Oncol 2021; 14: 101157.

[17] Farswan A, Gupta A, Sriram K, Sharma A, Kumar L and Gupta R. Does ethnicity matter in multiple myeloma risk prediction in the era of genomics and novel agents? Evidence from real-world data. Front Oncol 2021; 11: 720932.

[18] Farswan A and Gupta A. TV-DCT: method to impute gene expression data using DCT based sparsity and total variation denoising. ICASSP 2019; 1244-1248.

[19] Farswan A, Gupta A, Gupta R and Kaur G. Imputation of gene expression data in blood cancer and its significance in inferring biological pathways. Front Oncol 2020; 9: 1442.

[20] Anzar I, Sverchkova A, Stratford R and Clancy T. NeoMutate: an ensemble machine learning framework for the prediction of somatic mutations in cancer. BMC Med Genomics 2019; 12: 63.

[21] Hsu YC, Hsiao YT, Kao TY, Chang JG and Shieh GS. Detection of somatic mutations in exome

sequencing of tumor-only samples. Sci Rep 2017; 7: 15959.
[22] Pounraja VK, Jayakar G, Jensen M, Kelkar N and Girirajan S. A machine-learning approach for accurate detection of copy number variants from exome sequencing. Genome Res 2019; 29: 1134-1143.
[23] Huang T, Li J, Jia B and Sang H. CNV-MEANN: a neural network and mind evolutionary algorithm-based detection of copy number variations from next-generation sequencing data. Front Genet 2021; 12: 700874.
[24] Hill T and Unckless RL. A deep learning approach for detecting copy number variation in next-generation sequencing data. G3 (Bethesda) 2019; 9: 3575-3582.
[25] Li Y and Luo Y. Performance-weighted-voting model: an ensemble machine learning method for cancer type classification using whole-exome sequencing mutation. Quant Biol 2020; 8: 347-358.
[26] Collier O, Stoven V and Vert JP. LOTUS: a single- and multitask machine learning algorithm for the prediction of cancer driver genes. PLoS Comput Biol 2019; 15: e1007381.
[27] Han Y, Yang J, Qian X, Cheng WC, Liu SH, Hua X, Zhou L, Yang Y, Wu Q, Liu P and Lu Y. DriverML: a machine learning algorithm for identifying driver genes in cancer sequencing studies. Nucleic Acids Res 2019; 47: e45.
[28] Zeng Z, Mao C, Vo A, Li X, Nugent JO, Khan SA, Clare SE and Luo Y. Deep learning for cancer type classification and driver gene identification. BMC Bioinformatics 2021; 22 Suppl 4: 491.
[29] Luo P, Ding Y, Lei X and Wu FX. deepDriver: predicting cancer driver genes based on somatic mutations using deep convolutional neural networks. Front Genet 2019; 10: 13.
[30] Kipf TN and Welling M. Semi-supervised classification with graph convolutional networks. ArXiv Preprint 2016; ArXiv: 1609.02907.
[31] Bruna J, Zaremba W, Szlam A and LeCun Y. Spectral networks and locally connected networks on graphs. ArXiv Preprint 2013; ArXiv: 1312.6203.
[32] Schulte-Sasse R, Budach S, Hnisz D and Marsico A. Integration of multiomics data with graph convolutional networks to identify new cancer genes and their associated molecular mechanisms. Nat Mach Intell 2021; 3: 513-526.
[33] Barabási AL, Gulbahce N and Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Genet 2011; 12: 56-68.
[34] Goh KI, Cusick ME, Valle D, Childs B, Vidal M and Barabási AL. The human disease network. Proc Natl Acad Sci U S A 2007; 104: 8685-8690.
[35] Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, Jensen LJ and von Mering C. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res 2021; 49: D605-D612.
[36] Lundberg SM and Lee SI. A unified approach to interpreting model predictions. NuraIPS 2017; 30.
[37] Ahmad MA, Eckert C and Teredesai A. Interpretable machine learning in healthcare. ACM-BCB 2018; 559-560.
[38] Du M, Liu N and Hu X. Techniques for interpretable machine learning. CACM 2019; 63: 68-77.
[39] Keats JJ, Craig DW, Liang W, Venkata Y, Kurdoglu A, Aldrich J, Auclair D, Allen K, Harrison B, Jewell S, Kidd PG, Correll M, Jagannath S, Siegel DS, Vij R, Orloff G, Zimmerman TM; Mmrf CoMMpass Network, Capone W, Carpten J and Lonial S. Interim analysis of the MMRF CoMMpass Trial, a longitudinal study in multiple myeloma relating clinical outcomes to genomic and immunophenotypic profiles. Blood 2013; 122: 532.
[40] Lappalainen I, Almeida-King J, Kumanduri V, Senf A, Spalding JD, Ur-Rehman S, Saunders G, Kandasamy J, Caccamo M, Leinonen R, Vaughan B, Laurent T, Rowland F, Marin-Garcia P, Barker J, Jokinen P, Torres AC, de Argila JR, Llobet OM, Medina I, Puy MS, Alberich M, de la Torre S, Navarro A, Paschall J and Flicek P. The European genome-phenome archive of human data consented for biomedical research. Nat Genet 2015; 47: 692-695.
[41] Fan Y, Xi L, Hughes DS, Zhang J, Zhang J, Futreal PA, Wheeler DA and Wang W. MuSE: accounting for tumor heterogeneity using a sample-specific error model improves sensitivity and specificity in mutation calling from sequencing data. Genome Biol 2016; 17: 178.
[42] Benjamin D, Sato T, Cibulskis K, Getz G, Stewart C and Lichtenstein L. Calling somatic SNVs and indels with Mutect2. BioRxiv 2019; 861054.
[43] Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L and Wilson RK. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 2012; 22: 568-576.
[44] Larson DE, Harris CC, Chen K, Koboldt DC, Abbott TE, Dooling DJ, Ley TJ, Mardis ER, Wilson RK and Ding L. SomaticSniper: identification of somatic point mutations in whole genome sequencing data. Bioinformatics 2012; 28: 311-317.

[45] Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, Banks E, Garimella KV, Altshuler D, Gabriel S and DePristo MA. From FastQ data to high confidence variant calls: the genome analysis toolkit best practices pipeline. Curr Protoc Bioinformatics 2013; 43: 11.10.1-11.10.33.

[46] Li H and Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009; 25: 1754-1760.

[47] McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M and DePristo MA. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010; 20: 1297-1303.

[48] Wang K, Li M and Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010; 38: e164.

[49] Rogers MF, Shihab HA, Mort M, Cooper DN, Gaunt TR and Campbell C. FATHMM-XF: accurate prediction of pathogenic point mutations via extended features. Bioinformatics 2018; 34: 511-513.

[50] Martincorena I, Raine KM, Gerstung M, Dawson KJ, Haase K, Van Loo P, Davies H, Stratton MR and Campbell PJ. Universal patterns of selection in cancer and somatic tissues. Cell 2017; 171: 1029-1041.

[51] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V and Vanderplas J. Scikit-learn: machine learning in python. J Mach Learn Res 2011; 12: 2825-2830.

[52] Chakravarty D, Gao J, Phillips SM, Kundra R, Zhang H, Wang J, Rudolph JE, Yaeger R, Soumerai T, Nissan MH, Chang MT, Chandarlapaty S, Traina TA, Paik PK, Ho AL, Hantash FM, Grupe A, Baxi SS, Callahan MK, Snyder A, Chi P, Danila D, Gounder M, Harding JJ, Hellmann MD, Iyer G, Janjigian Y, Kaley T, Levine DA, Lowery M, Omuro A, Postow MA, Rathkopf D, Shoushtari AN, Shukla N, Voss M, Paraiso E, Zehir A, Berger MF, Taylor BS, Saltz LB, Riely GJ, Ladanyi M, Hyman DM, Baselga J, Sabbatini P, Solit DB and Schultz N. OncoKB: a precision oncology knowledge base. JCO Precis Oncol 2017; 2017: PO.17.00011.

[53] Maura F, Bolli N, Angelopoulos N, Dawson KJ, Leongamornlert D, Martincorena I, Mitchell TJ, Fullam A, Gonzalez S, Szalat R, Abascal F, Rodriguez-Martin B, Samur MK, Glodzik D, Roncador M, Fulciniti M, Tai YT, Minvielle S, Magrangeas F, Moreau P, Corradini P, Anderson KC, Tubio JMC, Wedge DC, Gerstung M, Avet-Loiseau H, Munshi N and Campbell PJ. Genomic landscape and chronological reconstruction of driver events in multiple myeloma. Nat Commun 2019; 10: 3835.

[54] Martínez-Jiménez F, Muiños F, Sentís I, Deu-Pons J, Reyes-Salazar I, Arnedo-Pac C, Mularoni L, Pich O, Bonet J, Kranas H, Gonzalez-Perez A and Lopez-Bigas N. A compendium of mutational cancer driver genes. Nat Rev Cancer 2020; 20: 555-572.

[55] Tate JG, Bamford S, Jubb HC, Sondka Z, Beare DM, Bindal N, Boutselakis H, Cole CG, Creatore C, Dawson E, Fish P, Harsha B, Hathaway C, Jupe SC, Kok CY, Noble K, Ponting L, Ramshaw CC, Rye CE, Speedy HE, Stefancsik R, Thompson SL, Wang S, Ward S, Campbell PJ and Forbes SA. COSMIC: the catalogue of somatic mutations in cancer. Nucleic Acids Res 2019; 47: D941-D947.

[56] De Cesco S, Davis JB and Brennan PE. TargetDB: a target information aggregation tool and tractability predictor. PLoS One 2020; 15: e0232644.

[57] Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW and Ma'ayan A. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res 2016; 44: W90-W97.

[58] Xie Z, Bailey A, Kuleshov MV, Clarke DJB, Evangelista JE, Jenkins SL, Lachmann A, Wojciechowicz ML, Kropiwnicki E, Jagodnik KM, Jeon M and Ma'ayan A. Gene set knowledge discovery with enrichr. Curr Protoc 2021; 1: e90.

[59] Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR and Ma'ayan A. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics 2013; 14: 128.

[60] Lopez-Corral L, Sarasquete ME, Beà S, García-Sanz R, Mateos MV, Corchete LA, Sayagués JM, García EM, Bladé J, Oriol A, Hernández-García MT, Giraldo P, Hernández J, González M, Hernández-Rivas JM, San Miguel JF and Gutiérrez NC. SNP-based mapping arrays reveal high genomic complexity in monoclonal gammopathies, from MGUS to myeloma status. Leukemia 2012; 26: 2521-2529.

[61] Wei G, Zhou L, Lu LL and Romano M. Sequential EHR-based dynamic graph network for multiple myeloma detection and feature interaction investigation. J Clin Oncol 2022; 40: e13591.

[62] National Cancer Institute (NCI). Targeted therapy directed by genetic testing in treating patients with advanced refractory solid tumors, lymphomas, or multiple myeloma (The MATCH Screening Trial). NLM Identifier: NCT0246-5060; 2020.

[63] Gupta R, Kaur G, Kumar L, Rani L, Mathur N, Sharma A, Dahiya M, Shekhar V, Khan S, Mookerjee A and Sharma OD. Nucleic acid based risk assessment and staging for clinical practice in multiple myeloma. Ann Hematol 2018; 97: 2447-2454.

[64] Benard B, Christofferson A, Legendre C, Aldrich J, Nasser S, Yesil J, Auclair D, Liang W, Lonial S and Keats JJ. FGFR3 mutations are an adverse prognostic factor in patients with t(4;14)(p16;q32) multiple myeloma: an mmrf commpass analysis. Blood 2017; 130: 3027.

[65] Salama AKS, Li S, Macrae ER, Park JI, Mitchell EP, Zwiebel JA, Chen HX, Gray RJ, McShane LM, Rubinstein LV, Patton D, Williams PM, Hamilton SR, Armstrong DK, Conley BA, Arteaga CL, Harris LN, O'Dwyer PJ, Chen AP and Flaherty KT. Dabrafenib and trametinib in patients with tumors with BRAFV600E mutations: results of the NCI-MATCH trial subprotocol H. J Clin Oncol 2020; 38: 3895-3904.

[66] Misund K, Keane N, Stein CK, Asmann YW, Day G, Welsh S, Van Wier SA, Riggs DL, Ahmann G, Chesi M, Viswanatha DS, Kumar SK, Dispenzieri A, Gonzalez-Calle V, Kyle RA, O'Dwyer M, Rajkumar SV, Kortüm KM, Keats JJ; MMRF CoMMpass Network; Fonseca R, Stewart AK, Kuehl WM, Braggio E and Bergsagel PL. MYC dysregulation in the progression of multiple myeloma. Leukemia 2020; 34: 322-326.

[67] Garcia SB, Ruiz-Heredia Y, Da Via M, Gallardo M, Garitano-Trojaola A, Zovko J, Raab MS, Sonneveld P, Braggio E, Stewart AK and Einsele H. Role of MAX as a tumor suppressor driver gene in multiple myeloma. Blood 2017; 130: 4347.

[68] Ohmine K and Uchibori R. Novel immunotherapies in multiple myeloma. Int J Hematol 2022; 115: 799-810.

[69] Jasrotia S, Gupta R, Sharma A, Halder A and Kumar L. Cytokine profile in multiple myeloma. Cytokine 2020; 136; 155271.

[70] Hoteit R, Bazarbachi A, Antar A, Salem Z, Shammaa D and Mahfouz R. KIR genotype distribution among patients with multiple myeloma: higher prevalence of KIR 2DS4 and KIR 2DS5 genes. Meta Gene 2014; 2: 730-736.

[71] Sun H, Martin TG, Marra J, Kong D, Keats J, Macé S, Chiron M, Wolf JL, Venstrom JM and Rajalingam R. Individualized genetic makeup that controls natural killer cell function influences the efficacy of isatuximab immunotherapy in patients with multiple myeloma. J Immunother Cancer 2021; 9: e002958.

[72] Mahaweni NM, Ehlers FAI, Bos GMJ and Wieten L. Tuning natural killer cell anti-multiple myeloma reactivity by targeting inhibitory signaling via KIR and NKG2A. Front Immunol 2018; 9: 2848.

[73] White BS, Lanc I, O'Neal J, Gupta H, Fulton RS, Schmidt H, Fronick C, Belter EA Jr, Fiala M, King J, Ahmann GJ, DeRome M, Mardis ER, Vij R, DiPersio JF, Levy J, Auclair D and Tomasson MH. A multiple myeloma-specific capture sequencing platform discovers novel translocations and frequent, risk-associated point mutations in IGLL5. Blood Cancer J 2018; 8: 35.

[74] Lozano E, Díaz T, Mena MP, Suñe G, Calvo X, Calderón M, Pérez-Amill L, Rodríguez V, Pérez-Galán P, Roué G, Cibeira MT, Rosiñol L, Isola I, Rodríguez-Lobato LG, Martin-Antonio B, Bladé J and Fernández de Larrea C. Loss of the immune checkpoint CD85j/LILRB1 on malignant plasma cells contributes to immune escape in multiple myeloma. J Immunol 2018; 200: 2581-2591.

[75] Kang X, Kim J, Deng M, John S, Chen H, Wu G, Phan H and Zhang CC. Inhibitory leukocyte immunoglobulin-like receptors: immune checkpoint proteins and tumor sustaining factors. Cell Cycle 2016; 15: 25-40.

[76] Beksac M, Gragert L, Fingerson S, Maiers M, Zhang MJ, Albrecht M, Zhong X, Cozen W, Dispenzieri A, Lonial S and Hari P. HLA polymorphism and risk of multiple myeloma. Leukemia 2016; 30: 2260-2264.

[77] Kassem S, Diallo BK, El-Murr N, Carrié N, Tang A, Fournier A, Bonnevaux H, Nicolazzi C, Cuisinier M, Arnould I, Sidhu SS, Corre J, Avet-Loiseau H, Teillaud JL, van de Velde H, Wiederschain D, Chiron M, Martinet L and Virone-Oddos A. SAR442085, a novel anti-CD38 antibody with enhanced antitumor activity against multiple myeloma. Blood 2022; 139: 1160-1176.

[78] Chu D and Wei L. Nonsynonymous, synonymous and nonsense mutations in human cancer-related genes undergo stronger purifying selections than expectation. BMC Cancer 2019; 19: 359.

[79] Sharma Y, Miladi M, Dukare S, Boulay K, Caudron-Herger M, Groß M, Backofen R and Diederichs S. A pan-cancer analysis of synonymous mutations. Nat Commun 2019; 10: 2569.

[80] Soussi T, Taschner PE and Samuels Y. Synonymous somatic variants in human cancer are not infamous: a plea for full disclosure in databases and publications. Hum Mutat 2017; 38: 339-342.

[81] Teng H, Wei W, Li Q, Xue M, Shi X, Li X, Mao F and Sun Z. Prevalence and architecture of post transcriptionally impaired synonymous mutations in 8,320 genomes across 22 cancer types. Nucleic Acids Res 2020; 48: 1192-1205.

[82] Supek F, Miñana B, Valcárcel J, Gabaldón T and Lehner B. Synonymous mutations frequently act as driver mutations in human cancers. Cell 2014; 156: 1324-1335.

Translational Oncology 14 (2021) 101157

Contents lists available at ScienceDirect

Translational Oncology

journal homepage: www.elsevier.com/locate/tranon

Original Research

AI-supported modified risk staging for multiple myeloma cancer useful in real-world scenario

Akanksha Farswan [a], Anubha Gupta [a,*], Ritu Gupta [b,*], Saswati Hazra [a], Sadaf Khan [b], Lalit Kumar [c], Atul Sharma [c]

[a] SBILab, Department of ECE, Indraprastha Institute of Information Technology-Delhi, New Delhi 110020, India
[b] Laboratory Oncology Unit, Dr. B.R.A. IRCH, AIIMS, New Delhi 110029, India
[c] Department of Medical Oncology, Dr. B.R.A. IRCH, AIIMS, New Delhi, India

ARTICLE INFO

Keywords:
Machine learning
Risk stratification of multiple myeloma
J48 decision tree
BIRCH clustering
Hazard ratios
Hematological malignancy

ABSTRACT

Introduction: : An efficient readily employable risk prognostication method is desirable for MM in settings where genomics tests cannot be performed owing to geographical/economical constraints. In this work, a new Modified Risk Staging (MRS) has been proposed for newly diagnosed Multiple Myeloma (NDMM) that exploits six easy-to-acquire clinical parameters i.e. age, albumin, β2-microglobulin (β2M), calcium, estimated glomerular filtration rate (eGFR) and hemoglobin.

Materials and Methods: : MRS was designed using a training cohort of 716 NDMM patients of our inhouse MM Indian (MMIn) cohort and validated on MMIn (n=354) cohort and MMRF (n=900) cohort. K-adaptive partitioning (KAP) was used to find new thresholds for the parameters. Risk staging rules, obtained via training a J48 classifier, were used to build MRS.

Results: : New thresholds were identified for albumin (3.6 g/dL), β2M (4.8 mg/L), calcium (11.13 mg/dL), eGFR (48.1 mL/min), and hemoglobin (12.3 g/dL) using KAP on the MMIn dataset. On the MMIn dataset, MRS outperformed ISS for OS prediction in terms of C-index, hazard ratios, and its corresponding p-values, but performs comparable in prediction of PFS. On both MMIn and MMRF datasets, MRS performed better than RISS in terms of C-index and p-values. A simple online tool was also designed to allow automated calculation of MRS based on the values of the parameters.

Discussion: : Our proposed ML-derived yet simple staging system, MRS, although does not employ genetic features, outperforms RISS as confirmed by better separability in KM survival curves and higher values of C-index on both MMIn and MMRF datasets.

Funding: : Grant: BT/MED/30/SP11006/2015 (Department of Biotechnology, Govt. of India), Grant: DST/ICPS/CPS-Individual/2018/279(G) (Department of Science and Technology, Govt. of India), UGC-Senior Research Fellowship.

Introduction

Staging of disease in oncology practice has been a useful tool for risk stratification as it helps in identifying patients requiring intense therapy upfront and/or a higher monitoring frequency during the follow-up periods. The first staging system for multiple myeloma (MM) was proposed by Salmon and Durie in 1975 that divided patients into three risk categories with differential overall survival [1]. Subsequently, in 2005 an International staging system (ISS) based on two simple laboratory parameters of serum albumin and beta2-microglobulin (β2M) was proposed by Greipp PR and colleagues [2]. Serum albumin reflected the normalcy of the protein compartment and serum β2M reflected the tumor burden. With the development of novel agents such as immunomodulators (IMIDs) and proteasome inhibitors (PSI) for treatment of MM, the landscape of responses and survival changed drastically [3,4]. In addition, the advances in molecular biology allowed investigators to look closely at the genomic changes in MM and, especially, in subgroups of patients with poor outcome. This led to the inclusion of cytogenetic aberrations into the staging system used for MM and thereby, emerged the Revised-ISS (RISS) [5]. The survival data used for developing RISS

* Corresponding authors.
E-mail addresses: anubha@iiitd.ac.in (A. Gupta), drritu.laboncology@aiims.edu (R. Gupta).

https://doi.org/10.1016/j.tranon.2021.101157
Received 10 April 2021; Received in revised form 8 June 2021; Accepted 15 June 2021
Available online 8 July 2021

consisted predominantly of patients who were treated with immunomodulatory agents. As the new class of drugs, i.e., the PSI made their way into the treatment of MM, some of the cytogenetic aberrations like t (4;14) included in the RISS seem to lose their poor prognostic impact [6]. From an academic and research perspective, it is desirable to characterize subset of patients with poor clinical outcome to develop effective therapies but in clinical practice, it is desirable to have a staging system that is based on clinical and laboratory parameters that are easily accessible in healthcare setting across the globe.

In recent times, data analytics including advanced machine learning methods are being used to extract valuable information from medical records. Machine learning algorithms have been shown to be useful in devising risk stratification system in type-2 diabetic patients by Ricci and colleagues [7], in cardiovascular disorders by Ahuja and Schaar [8], in prostate cancer by Varghese [9], in resected gastric cancer[10] and in nasopharyngeal carcinoma [11]. In this study, we used machine learning methods to develop a new risk stratification system, namely, Modified Risk Staging (MRS) for MM using six easy-to-acquire laboratory parameters: albumin, β2M, calcium, eGFR, hemoglobin along with age. The model was developed on a training dataset of patients with newly diagnosed multiple myeloma (NDMM) and validated on two test datasets. Rigorous comparison of the proposed risk staging model with ISS and RISS was undertaken to check its efficacy on the predictions of progression free survival (PFS) and overall survival (OS).

Methods

Study population

The computerized database search on June 28, 2019 with keyword 'ICD C90' returned 1675 entries of patients registered at the Institute Rotary Cancer centre, All India Institute of Medical Sciences (AIIMS). A total of 253 patients had plasma cell dyscrasia other than MM, 132 patients were lost to follow up after a single visit (n=111) or before first response could be assessed (n=21), and 121 patients' records had inadequate clinical and/or laboratory parameters. Patients who died within 16 weeks of diagnosis were labelled as early deaths (n=99) and were excluded from the staging algorithms. Remaining cohort of 1070 Indian patients of MM, referred to as the MMIn cohort, was evaluated in this study (Supplementary Fig. S1). An independent cohort of 900 patients of MM enrolled in Multiple Myeloma Research Foundation (MMRF) repository, for which the clinical and laboratory data is available publicly, was used for validation.

Clinical and laboratory characteristics

The clinical, laboratory, and radiological data was obtained from the medical case files. A subset of patients, for whom the molecular data (n=627) was available, were assigned RISS as described previously [12]. Treatment response was assessed as per the International uniform response criteria for multiple myeloma [13]. Progression free survival (PFS) was calculated from the date of diagnosis until progression or death. Overall survival (OS) was calculated from the date of diagnosis until death due to any cause or was censored at last follow-up. Clinical and laboratory features of the patients are given in Supplementary Table S1.

Design strategy

Patients (n=1070) in MMIn were randomly split in the ratio of 67:33 as training (n=716) and test cohorts (n=354). The test cohort did not have any missing value. In the training cohort, 41 patients (5.7% of 716 patients) had one or two missing values that were imputed with the median value of the parameters. The training data was used to develop the proposed staging system called Modified Risk Staging (MRS) and the test dataset was used to evaluate the correctness of the MRS. The staging system was then validated on MMRF data. No missing imputation was applied on the test cohort or the MMRF dataset. Complete MRS design strategy, shown in Fig. 1, is explained below.

Initially seven parameters, i.e., albumin, β2M, calcium, eGFR, hemoglobin, lactate dehydrogenase (LDH), and age were evaluated for designing MRS. β2M and LDH levels are reflective of tumor burden and serum albumin, hemoglobin, calcium and creatinine are reflective of the bone and renal homeostasis. eGFR was calculated from creatinine concentration using MDRD eGFR equation [14]. LDH values were brought to a common scale by multiplying each entry by 280 and dividing it by the upper limit of LDH provided for that particular entry. For each parameter, patients were initially divided into high-risk and low-risk groups using the well-established cut-offs of these parameters. Established thresholds for albumin and β2M are derived from ISS [2] and for eGFR, calcium, hemoglobin are derived from revised IMWG criteria [15]. Log-rank test on the Kaplan-Meier curves yielded significant p-values for all the parameters except LDH which was, therefore, not used further (Table 1). Next, K-adaptive partitioning (KAP) [16] algorithm was used to find new threshold values for the six parameters. KAP was performed on the training patients' parameters yielding two threshold values for each parameter, one from PFS and the other from OS analysis. The threshold with lower p-value of the two was chosen as the new cut-off for each parameter.

For the cumulative integration of parameters into risk staging, weights were assigned to each parameter using their respective hazard ratios (HR) for PFS and OS obtained from the univariate Cox-proportional hazard test on the training data (Supplementary Table S2). For each parameter, the highest of the two HR values obtained from PFS and OS was chosen and normalized using 'minmax' scaling in the range of 1 to 4. The scaled HR values were assigned as the respective weights of each of the parameters (Table 1). This captures the relative impact of each parameter on the patients' survival. Next, a new score for each patient was calculated by adding the weights of all those parameters that had values (in the respective patient) beyond the threshold defined for the high-risk group. These patient scores were used to compute an adjacency matrix of 716 rows and 716 columns (columns are features), where each row corresponds to one patient and each entry in the row is the absolute difference between the score of that patient with each of the 716 patients including self. BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) Clustering, an unsupervised ML method, was applied on the adjacency matrix to cluster the patients of the training dataset into three risk groups [17]. Each cluster of patients was assigned one label: Stage-1 (low-risk), Stage-2 (intermediate-risk), or Stage-3 (high-risk). Initially assigned risk stages via BIRCH clustering on training patients were used as ground truth labels. BIRCH is an unsupervised clustering algorithm that works on the entire data. It does not provide rules that can be employed on any prospective subject to determine its risk stage. Hence, there was a need to obtain rules of risk staging. At the same time, a supervised classifier cannot be trained initially, because there is no risk stage-label. A novel methodology is employed, wherein the risk stage labels provided by BIRCH were used as ground truth class labels (risk stages) on the training data to train a J48 classifier (a rule-based supervised decision-tree classifier). The trained J48 classifier provided the rules in terms of laboratory parameters and age for the identification of risk groups, labeled as MRS-1 (low risk), MRS-2 (intermediate-risk), and MRS-3 (high-risk) (Fig. 2). The risk stage assigned by the J48 tree was considered the actual risk class for each patient. All the patients in the test dataset were also assigned to one of the MRS groups using the J48 rules. These MRS groups were then analyzed for OS and PFS, and compared with those obtained with the ISS and RISS.

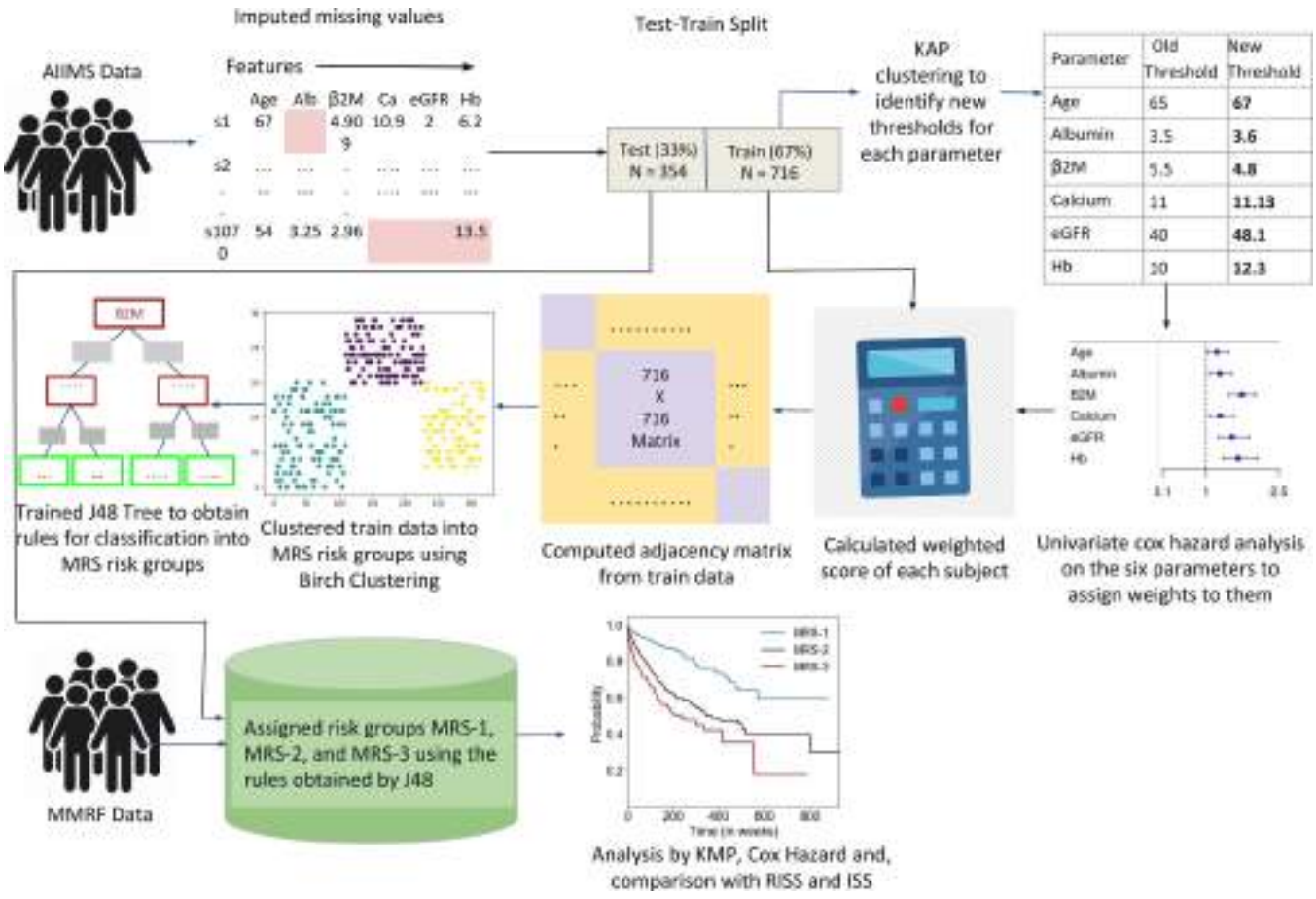


Fig. 1. Workflow for the development of Modified Risk staging (MRS) system for Multiple Myeloma.

Table 1
Comparison of established and proposed cut-offs for laboratory parameters for stratification of patients for progression free survival (PFS) and overall survival (OS) in training data (n=716 patients) using Kaplan Meier analysis and the weights assigned to the laboratory parameters for calculation of score.

Parameter	Established Threshold value	Proposed Threshold value	PFS		OS		Weights assigned after univariate Cox Hazard analysis
			p-value with established threshold	*p*-value with proposed threshold	*p*-value with established threshold	*p*-value with proposed threshold	
Age	>65	**>67**	0.19	**0.016**	8.7e-4	**1.86e-5**	**2.85**
Albumin	≤ 3.5	**≤ 3.6**	0.59	**0.089**	0.06	**2.7e-3**	**1**
β2M	≥ 5.5	**≥ 4.8**	**5.41e-6**	5.81e-6	5.6e-6	**8.78e-8**	**2.85**
Calcium	≥ 11	**≥ 11.13**	0.011	**6.5e-3**	**0.02**	0.029	**1.07**
eGFR	≤ 40	**≤ 48.1**	0.03	**0.012**	7.9e-3	**1.2e-3**	**1.14**
Hb	≤ 10	**≤ 12.3**	0.012	**4.88e-4**	0.027	**1.2e-4**	**4.0**
LDH	< 280	< 95	0.66	0.96	0.52	0.43	–

*Most significant *p*-values under each category are highlighted in bold.

Results

Clinical and laboratory characteristics of myeloma patients

The baseline demographic and laboratory features of patients are given in Supplementary Table S1; the training and the test cohorts were comparable in baseline demographic, laboratory, and clinical parameters (Supplementary Table S1). All the patients received novel agents (IMIDs:thalidomide or lenalidomide and/or PSI i.e. bortezomib) either as primary or maintenance therapy and dexamethasone. 56.5% of patients received triplet regimen. With a median follow up period of 166 weeks (range: 14–961 weeks), 626 patients had progressed and 372 died; the median PFS and OS of the entire cohort was 117 weeks and 166 weeks, respectively.

Patients in the training cohort (n=716) were initially stratified into high-risk and low-risk groups based on the new thresholds deduced by KAP algorithm that yielded better separability (lower *p*-values) between the two groups (Table 1). Weights assignment to each parameter on a scale of 1 to 4 led to the highest weight of 4 to hemoglobin and 2.85 to β2M and age, indicating them to be the important prognostic factors for risk stratification (Table 1). The trained J48 classifier yielded the rules for the risk stage assignment (Fig. 2A) as well as ten-fold cross-validation accuracy of 96.5% and the weighted-average ROC area of 97.5%. The patients were assigned labels, MRS-1 (Low-risk), MRS-2 (Intermediate-risk) and MRS-3 (High-risk) using the rules obtained from the J48 classifier. An online version of the MRS calculator (Supplementary Fig. S2) has also been developed. It calculates the risk stage of the patient based on the values of the six parameters, age, albumin, β2M, calcium,

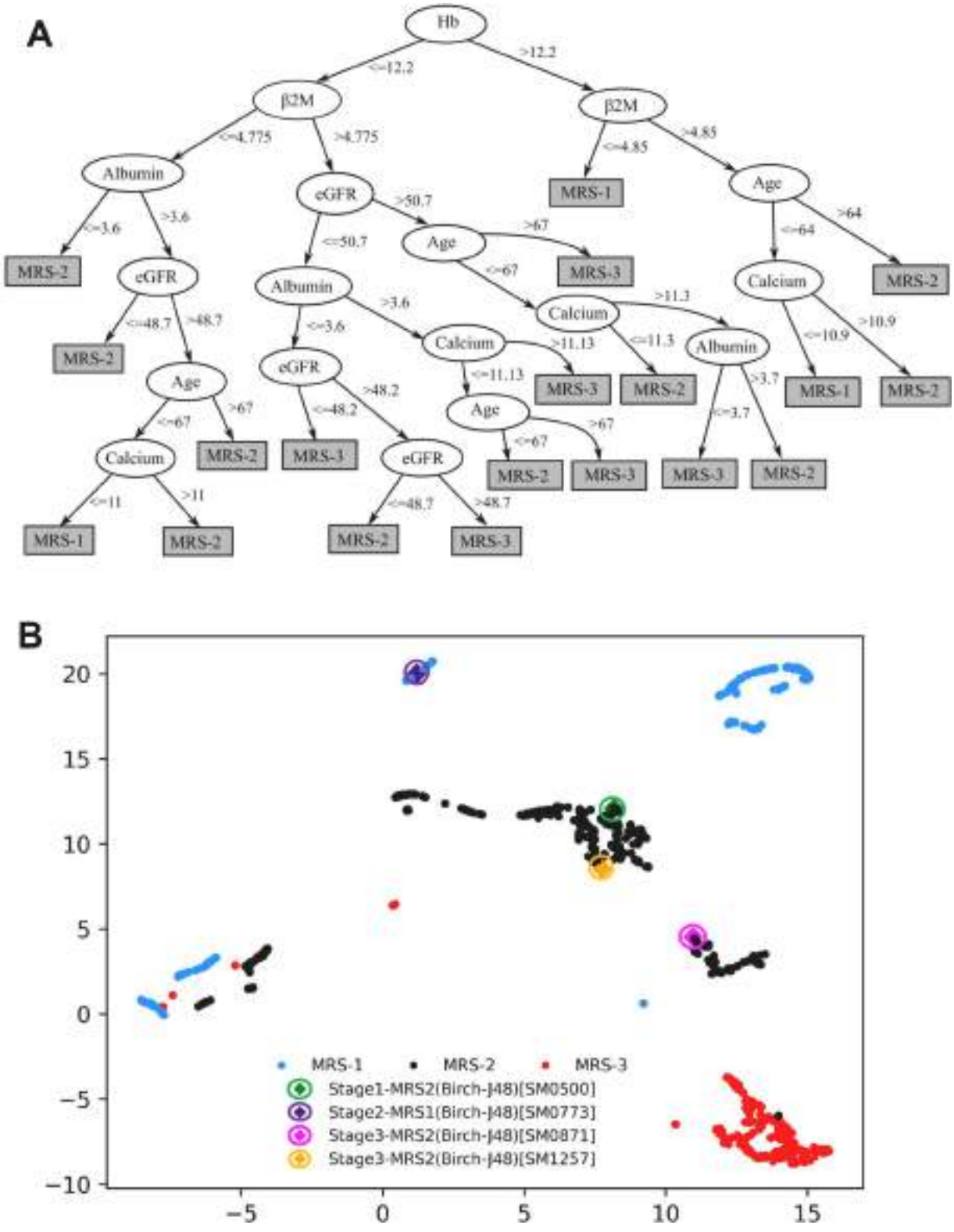


Fig. 2. **A**- Hierarchical rule based tree structure to assign data samples to MRS-1, MRS-2 and MRS-3 groups. Parameters: Age: Age; Alb: Albumin; β2M: beta2-macroglobulin; Ca: Calcium; eGFR: estimated glomerular filtration rate and Hb: hemoglobin. **B**- UMAP scatter plot of training data depicting the three labels identified by J48 classifier rules and the four mismatched patients (highlighted in circles).

eGFR and hemoglobin. It also displays median PFS and OS in weeks for the patient depending on his assigned risk group. Median PFS and OS have been calculated on the combined data (n=1970 patients) of MMIn and MMRF cohort.

Results on the training (n=716), test (n=354) and complete MMIn cohort (n=1070)

Largest proportion of training cohort *(n=716)* were assigned to MRS-2 (n=332, 46.36%), followed by MRS-3 (n=199, 27.80%) and MRS-1

(*n*=185, 25.84%). Results of the median PFS on MRS groups (*p*=6.28e-6) and ISS groups (*p*=1.25e-5) as well as of median OS on MRS groups (*p*=8.15e-10) and ISS groups (*p*=2.03e-5) show better performance of MRS than ISS (lower *p*-values; Supplementary Table S3). Similar findings were obtained on the test cohort (*n*=354; Supplementary Table S3). Univariate Cox analysis of the entire patient cohort (*n*=1070, Supplementary Table S4), revealed increased risk of progression and mortality for age>67 years, albumin≤3.6, β2M≥4.8, calcium≥11.13, eGFR≤48.1 and hemoglobin≤12.3. Using MRS, the largest proportion of patients were placed in MRS-2 (*n*=511, 47.76%) followed by MRS-1 (*n*=281, 26.26%) and MRS-3 (*n*=278, 25.98%). KM survival analysis of MRS groups indicated statistically significant difference in PFS between MRS-1 and MRS-2 groups (*p*=0.0012) and between MRS-2 and MRS-3 groups (*p*=0.0055). For ISS, the difference was significant between ISS-2 and ISS-3 groups (*p*=1.118e-6), but not between ISS-1 and ISS-2 groups (*p*=0.46). For RISS, there was statistically significant difference between RISS-2 and RISS-3 (*p*=5.6e-7) but not between RISS-1 and RISS-2 (*p*=0.96). KM survival analysis of MRS groups further revealed statistically significant difference in OS between MRS-1 and MRS-2 groups (*p*=5.9e-9) and between MRS-2 and MRS-3 groups (*p*=0.001). For ISS, the difference in OS was significant between ISS-2 and ISS-3 groups (*p*=3.12e-6) but not between ISS-1 and ISS-2 groups (*p*=0.118) and. For RISS, there was statistical difference in OS between RISS-2 and RISS-3 groups (*p*=8.32e-9), but was not significant between RISS-1 and RISS-2 groups (*p*=0.2) (Fig. 3). Results of multivariate Cox hazards model are also shown in Supplementary Table S5.

The C-Statistic computed on MRS and ISS demonstrate slightly better performance of MRS than ISS with respect to PFS and OS. C-Statistic for MRS was 0.57 (HR=1.34, 95% CI=1.20–1.5, *p*=1.9e-7) for PFS and 0.63 (HR=1.79, 95% CI=1.54–2.06, *p*=5.17e-15) for OS as compared to 0.57 (HR=1.36, 95% CI=1.23–1.52, *p*=9.9e-9) and 0.60 (HR=1.56, 95% CI=1.35–1.8, *p*=9.22e-10) for ISS (Fig. 3). RISS was available for only 627 patients, hence, MRS stages were determined separately for these patients. C-statistic was better for MRS with values of 0.57 (HR=1.38, 95% CI=1.17–1.62, *p*=9e-5) for PFS and 0.63 (HR=1.90, 95% CI=1.53–2.35, *p*=6.3e-9) for OS as compared to 0.56 (HR=1.61, 95% CI=1.29–2.00, *p*=2e-5) and 0.60 (HR=2.27, 95% CI=1.72–3.00, *p* = 8.73e-9) for RISS (Fig. 3).

Results on the MMRF cohort

MRS was further evaluated by comparing it with ISS and RISS using the MMRF dataset. ISS was available for 900 patients and RISS was available for 703 patients. In MMRF cohort, majority of the patients were placed in MRS-2 (*n*=405, 45%) followed by MRS-1 (*n*=348, 38.67%) and MRS-3 (*n*=147, 16.33%). In the univariate Cox hazard analysis of the MMRF data, risk of progression and mortality was increased for age>67 years, β2M≥4.8, albumin≤3.6, hemoglobin≤12.3, eGFR≤48.1 and calcium≥11.13 (Supplementary Table S4). Results of the median PFS on MRS groups (*p*=3.11e-11), ISS groups (*p*=7.35e-12), and RISS groups (*p*=1.21e-6) as well as of median OS on MRS groups (*p*=6.00e-13), ISS groups (*p*=9.28e-14), and RISS groups (*p*=1.23e-9) show comparable performance of MRS than ISS and RISS (comparable *p*-values; Supplementary Table S3; Fig. 4). The risk of progression and that of mortality was increased for MRS 2vs1, MRS 3vs1, ISS 2vs1, ISS 3vs1, RISS 2vs1, and RISS 3vs1 (Fig. 4). The C-Statistic for MRS in MMRF data is 0.60 (HR=1.60, 95% CI=1.39–1.82, *p*=6.01e-12) for PFS and 0.65 (HR=2.09, 95% CI=1.71–2.56, *p*=8.50e-13) for OS as compared to 0.61 (HR=1.54, 95% CI=1.37–1.74, *p*=2.8e-12) and 0.667 (HR=2.04, 95% CI=1.68–2.47, *p*=2.3e-13) for ISS; 0.58 (HR=1.67, 95% CI=1.36–2.06, *p*=9.60e-7) for RISS and 0.62 (HR=2.38, 95% CI=1.76–3.23, *p*=1.75e-8) for RISS, respectively. Results of multivariate Cox hazards model are also shown in Supplementary Table S6. The 5-year OS for the complete MMIn data (*n*=1070) was 85.82% for MRS-1, 61.73% for MRS-2 and 48.78% for MRS-3 (Table 2). The difference in the percentages of the 5-year OS and median OS for different risk groups indicated that the groups were significant. A similar stratification was achieved when the MRS model was applied on the MMRF test dataset. The 5-year OS for MMRF data was 79.06% for MRS-1, 66.66% for MRS-2 and 41.91% which is quite comparable to that obtained in the MMIn data.

For MMRF data, 5-year OS was 57.48% for RISS-1, 70.49% for RISS-2 and 35.27% for RISS-3 (Table 2) which suggested some anomaly since RISS-1 should have a higher OS as compared to RISS-2. This anomaly may be because of assigning a much larger number of patients to RISS-2 having greater OS time as compared to RISS-1.

Discussion

The advent of immunomodulatory drugs and PSI has considerably improved treatment outcomes in MM and hence, the current risk stratifications based on ISS and RISS need to be relooked at. ISS is a simple model based on two laboratory parameters of serum-albumin and β2M, but is largely based on data from patients treated in the pre-IMID era and is not very informative on PFS in NDMM patients [2]. Since the length of PFS is an important predictor of long-term OS, a better model to assess PFS is desirable [12]. The RISS for MM takes into consideration the molecular abnormalities and is based on data from patients treated with either IMIDS or PI, or both and is informative on PFS as well as OS [5]. However, many molecular aberrations such as 1q gain and chromothripsis that adversely affect outcome in MM have been overlooked and t(4;14) included in RISS has lost significance in patients treated with triplet regimens [6,18]. The preferred frontline treatment for MM is triplet regimen consisting of an IMID, PI and steroid regardless of molecular aberrations or risk stratification. The targeted therapy in MM is reserved for patients with RAS and BRAF mutations in progressive disease and relapsed refractory setting, and the upfront molecular characterization adds to the cost of healthcare in clinical settings. It is, thus, desirable to develop simple risk staging models for MM to enable judicious use of healthcare resources reserving the molecular analysis for patients who progress or relapse on frontline therapy and in setting of clinical trials when a targeted therapy is intended to be used.

Performance of MRS as compared to RISS and ISS

On the MMIn dataset, the proposed MRS performed better than ISS in prediction of OS in terms of C-index, HR, and *p*-values, while the performance in PFS was comparable. The performance of MRS was superior to RISS in terms of C-index and *p*-values. On the MMRF dataset, the performance of MRS was superior to RISS in terms of C-index and *p*-values, but was comparable to ISS. The performance on MMRF dataset indicated that there may be nuances with respect to ethnicity and race as the population of MMIn and MMRF dataset is different in terms of ethnicity and race. The algorithm was trained on MMIn dataset using the KAP-proposed thresholds on MMIn dataset while the developed MRS model was only tested on the MMRF dataset.

Specific characteristics of different staging systems

Levels of albumin and β2M are significant in ISS staging method, while LDH along with albumin and β2M are significant in RISS staging method (Supplementary Table S8). LDH was not observed to be significant in our preliminary findings and was, therefore, excluded from MRS staging. eGFR although significant came lower in the tree. The possible reason could be that renal dysfunction gets reversed in a significant number of patients of MM treated with novel agents as opposed to alkylating agents used in the past to treat MM. Age remained a significant parameter of outcome in the new staging as well. In MRS, none of the deranged parameters individually dominated the underlying risk of death. For example, when the level of hemoglobin was less than 12.2 and β2M was greater than 4.775, then for eGFR levels less than 50.7 and albumin and calcium levels greater than 3.6 and 11.13 respectively, the patient was placed in the MRS-3 (High risk) group. However, even if the

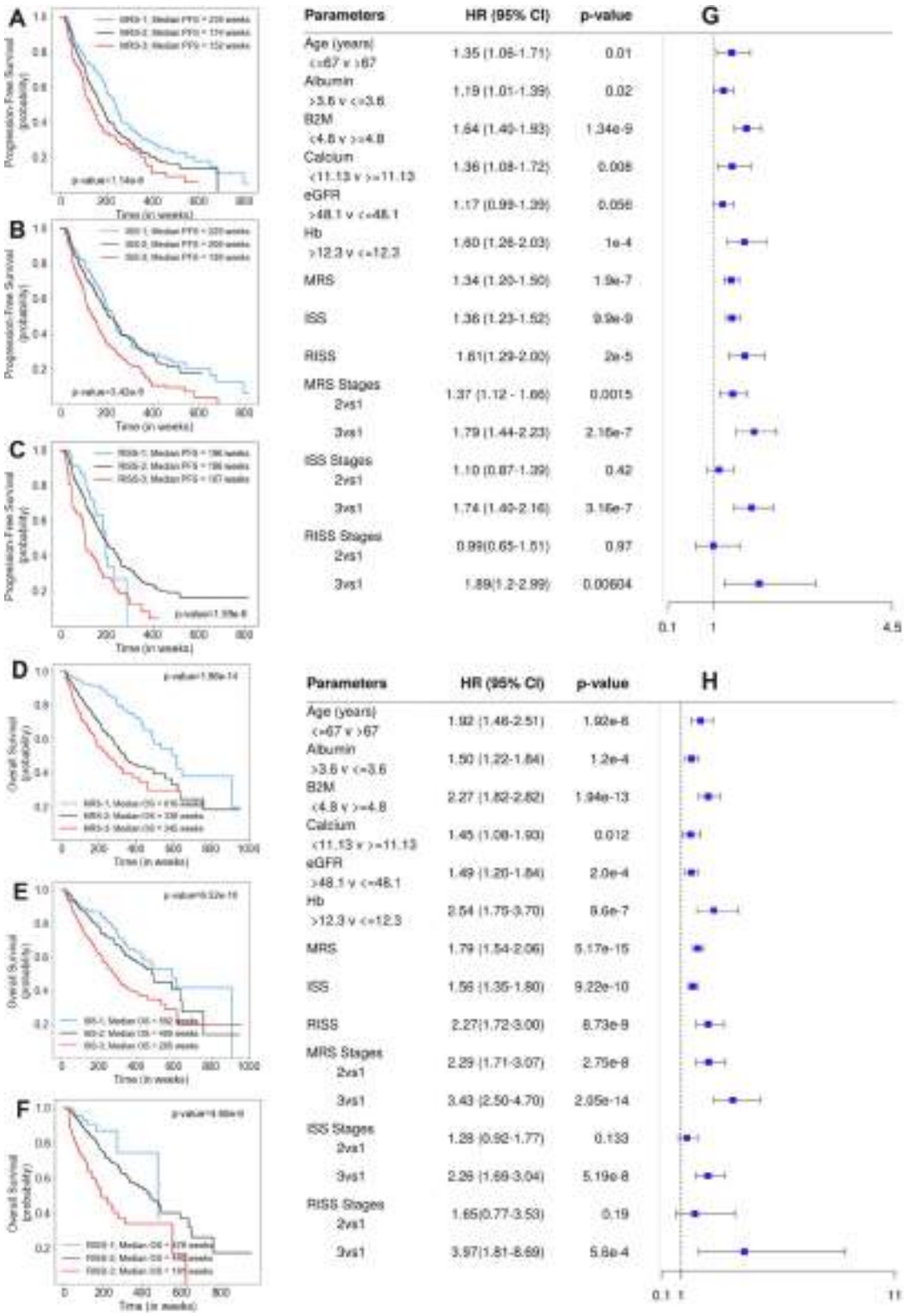
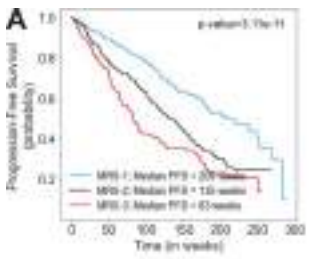
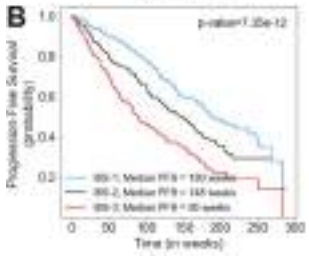
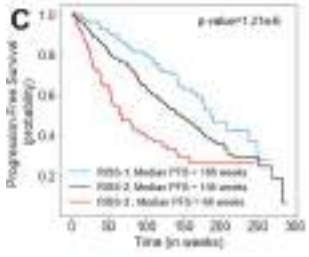
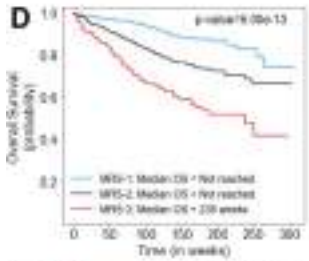


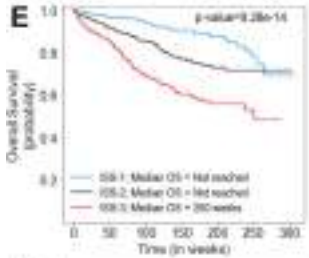
Fig. 3. A, B, C- Progression-Free Survival in patients with MM from MMIn cohort (n=1070) stratified by the proposed MRS (n=1070), ISS (n=1070) and RISS (n=627), respectively. **D, E, F**- Overall Survival in patients with MM from MMIn cohort (n=1070) stratified by the proposed MRS (n=1070), ISS (n=1070) and RISS (n=627), respectively. **G, H**- Cox Hazard Analysis of PFS and OS. Univariate analysis of parameters- Age, Albumin, β2M, Calcium, eGFR, Hb and different staging methods-MRS, ISS and RISS. Multivariate analysis of different groups of MRS, ISS and RISS.

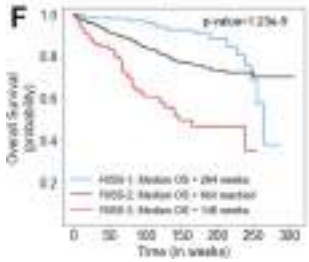






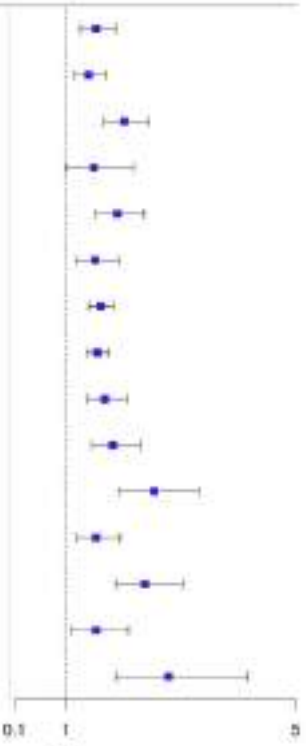






Parameters	HR (95% CI)	p-value
Age (years) <=67 v >67	1.52 (1.25-1.86)	3e-5
Albumin >3.6 v <=3.6	1.39 (1.14-1.68)	0.001
B2M <4.8 v >=4.8	2.01 (1.65-2.44)	1.5e-12
Calcium <11.13 v >=11.13	1.48 (0.99-2.18)	0.05
eGFR >48.1 v <=48.1	1.89 (1.51-2.36)	1.61e-8
Hb >12.3 v <=12.3	1.50 (1.17-1.92)	0.0013
MRS	1.60 (1.39-1.82)	6.01e-12
ISS	1.54 (1.37-1.74)	2.8 e-12
RISS	1.67(1.36-2.06)	9.6e-7
MRS Stages 2vs1	1.81 (1.43-2.28)	5.5e-7
3vs1	2.53 (1.92-3.30)	4.2e-11
ISS Stages 2vs1	1.52 (1.19-1.94)	6.4e-4
3vs1	2.37 (1.86-3.03)	2.94e-12
RISS Stages 2vs1	1.52(1.10-2.09)	9.9e-3
3vs1	2.79(1.86-4.15)	5.6e-7

G



Parameters	HR (95% CI)	p-value
Age (years) <=67 v >67	2.05 (1.54-2.75)	1e-6
Albumin >3.6 v <=3.6	2.01 (1.49-2.72)	1e-5
B2M <4.8 v >=4.8	2.71 (2.02-3.63)	2.65e-11
Calcium <11.13 v >=11.13	2.46 (1.56-3.89)	1e-4
eGFR >48.1 v <=48.1	2.51 (1.86-3.41)	2.9e-9
Hb >12.3 v <=12.3	1.56 (1.06-2.28)	0.02
MRS	2.09 (1.71-2.56)	8.5e-13
ISS	2.04 (1.68-2.47)	2.3e-13
RISS	2.38(1.78-3.23)	1.75e-8
MRS Stages 2vs1	2.21 (1.49-3.27)	7e-5
3vs1	4.44 (2.92-6.73)	2.45e-12
ISS Stages 2vs1	2.10 (1.38-3.18)	4.7e-4
3vs1	4.19 (2.81-6.25)	1.61e-12
RISS Stages 2vs1	1.82(1.06-3.15)	0.029
3vs1	5.02(2.74-9.19)	1.71e-7

H

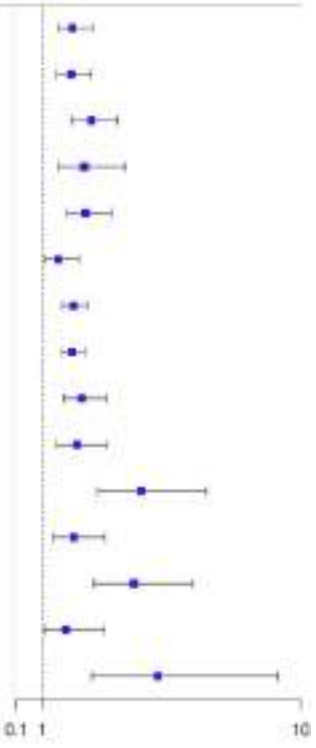


Fig. 4. **A, B, C**- Progression-Free Survival in patients with MM from MMRF cohort (n=900) stratified by the proposed MRS (n=900), ISS (n=900) and RISS (n=703), respectively. **D, E, F**- Overall Survival in patients with MM from MMRF cohort (n=900) stratified by the proposed MRS (n=900), ISS (n=900) and RISS (n=703), respectively. **G, H**- Cox Hazard analysis of PFS and OS. Univariate analysis of parameters- Age, Albumin, β2M, Calcium, eGFR, Hb and different staging methods- MRS, ISS and RISS. Multivariate analysis of different groups of MRS, ISS and RISS.

Table 2
Prediction of Progression-free survival and overall survival (in%) for MRS, ISS and RISS at 1, 2, 3, 4 and 5 years in MMIn (n=1070) and MMRF datasets (n=900).

MMIn dataset										
	Year	MRS (n = 1070)			ISS (n = 1070)			RISS (n = 627)		
		1	2	3	1	2	3	1	2	3
PFS	1	89.78	85.73	77.11	89.64	88.48	80.24	91.11	86.76	68.24
	2	77.00	71.45	58.46	80.33	73.89	62.68	84.15	72.39	53.41
	3	69.13	54.94	43.02	68.50	62.48	46.22	62.62	58.07	36.87
	4	55.66	40.28	33.60	53.22	49.98	33.48	33.46	45.28	27.44
	5	43.04	33.30	27.48	40.25	41.67	27.22	26.77	36.66	18.29
OS	1	96.75	93.19	86.74	93.64	95.88	89.94	97.78	95.56	82.27
	2	94.39	84.18	71.89	90.10	88.82	78.01	93.28	87.91	70.40
	3	91.72	76.52	63.62	87.84	82.63	69.45	90.76	81.80	59.32
	4	90.58	68.34	55.29	85.82	76.00	61.30	86.98	73.61	49.03
	5	85.82	61.73	48.78	80.94	71.91	52.79	86.98	68.65	40.39
MMRF dataset										
	Year	MRS (n = 900)			ISS (n = 900)			RISS (n = 703)		
		1	2	3	1	2	3	1	2	3
PFS	1	88.95	77.64	69.84	89.95	79.44	69.08	90.32	80.49	62.03
	2	77.56	59.90	42.20	77.11	61.86	46.24	79.57	61.73	38.13
	3	62.82	42.17	35.10	59.75	48.21	33.6	62.95	47.73	28.99
	4	51.49	28.09	23.16	47.7	32.95	22.35	46.41	33.74	26.36
	5	35.61	24.66	13.89	34.81	29.27	14.37	27.69	24.96	26.36
OS	1	96.93	90.90	82.81	96.76	91.61	85.29	98.07	91.17	83.21
	2	93.73	82.69	66.44	93.91	85.24	68.65	96.11	83.81	61.09
	3	88.43	76.18	59.20	90.73	77.16	60.79	92.86	77.28	49.78
	4	86.56	72.64	51.88	88.21	73.39	55.8	88.33	73.52	47.02
	5	79.06	66.66	41.91	74.73	71.92	48.93	57.48	70.49	35.27

eGFR level was greater than 50.7 but age was greater than 67, the patient was still placed in MRS-3. Hence, it was evident that poor outcome was associated with a combination of abnormally high or low levels of multiple prognostic factors. Thus, MRS staging does not rely on a single parameter but takes into consideration multiple parameters that are associated with hemodynamic systems as a whole.

In the MMRF dataset, 703 patients out of 900 had RISS labels. 91 out of these 703 patients (12.9%) were labeled as RISS-3. In these 91 patients, 43 were labeled as MRS-3, 43 as MRS-2 and 5 were labeled as MRS-1. The median OS for these 5 patients was 70 weeks and no death event was observed in any of these patients. In fact, there was no disease progression in these 5 patients in the first year of diagnosis. Hence, the staging provided by MRS is more accurate as it positioned these patients in the low-risk stage (MRS-1) contrary to the high-risk stage (RISS-3) provided by the RISS scheme. Similarly, 239 patients out of 900 (26.56%) patients were stratified as ISS-3. Out of these 239 patients, 125 were labeled MRS-3, 105 were labeled MRS-2 and 9 were labeled MRS-1. Median OS of these 9 patients was 70 weeks and no death event was observed in these patients. Further, there was no disease progression in the first year of diagnosis. MRS correctly placed the patients in a low-risk group as compared to ISS. Thus, it can be deduced that MRS helps in better filtering of the patients compared to RISS and ISS. Further, there were overall 147 patients (16.33%) that were stratified as MRS-3 in MMRF dataset. None of the 147 patients was present in the low-risk stages of ISS and RISS, thereby establishing the efficacy of MRS staging in identification of high risk patients.

Hierarchical rules in J48 tree and mismatched labels between BIRCH and J48 classifier

Hemoglobin and β2M were observed to be the most important poor prognostic factors in MRS staging followed by others. Hemoglobin had the greatest weight assigned based on hazard ratios obtained from the univariate Cox hazard analysis. It was present at the first level (highest node) for classification in the J48 tree, thus confirming the high prognostic value of β2M on PFS and OS. It was observed that no leaf node of the decision ended in MRS-3 stage for hemoglobin values greater than the threshold 12.2, while five leaf nodes ended in MRS-3 stage for β2M values lower than 12.2, thereby, indicating that the patients with higher disease load have lower hemoglobin. J48 tree utilized the values of hemoglobin at continuous scale and provided a single cut-off for hemoglobin as 12.2 in the decision rules (Fig. 2A) which is quite close to the proposed threshold value of 12.3 obtained via KAP and hence, justified the choice of our new threshold for hemoglobin. Lower levels of hemoglobin were predictors of poor outcome and were associated with high-risk patients, as evident from the hierarchical rules obtained from J48 classifier. Similarly, β2M levels of 4.775 or lower were associated with either low or intermediate risk as observed in the J48 classifier rules. J48 tree provided two cut-offs for β2M, 4.775 and 4.85 in the decision rules. These cut-offs were quite close to the proposed threshold for β2M, 4.8 and justified the choice of our new threshold for β2M. Apart from hemoglobin and β2M, the hierarchical rules in the J48 tree for other parameters also exhibited values comparable to the proposed thresholds.

The J48 assigned risk group label of four patients (0.5%) of the training set were found to have a mismatch with that assigned by the unsupervised BIRCH clustering. Training data was visualized using UMAP scatter plot. UMAP (Uniform Manifold Approximation and Projection) is a technique of dimensionality reduction mostly used for visualization of high dimensional data. It is evident from the plot that SM0500 (green), SM0773 (violet), SM0871 (magenta), and SM1257 (yellow) were assigned labels different from the labels obtained from BIRCH clustering. Further, there is a possibility of overfitting due to a smaller number of training samples (n=716), which was addressed by performing pruning on the J48 tree. Results in terms of p-values and hazard ratios in the test set of MMIn dataset and in the overall cohort of MMRF data is further suggestive of the reduced possibility of overfitting.

Sensitivity analysis of the J48 decision tree classifier

We first trained the J48 decision tree classifier using all the variables and observed the 10-fold cross validation accuracy. As already discussed, the ground truth risk stages were obtained via BIRCH clustering. We then trained the J48 classifier under the same settings six times, each

time excluding one of the variables and observed their classification accuracy. Highest classification accuracy was 96.5% when all the parameters were used for training. Classification accuracy was least affected by absence of calcium. However, classification accuracy decreased drastically (Supplementary Table S7) in absence of one of the variables- albumin, β2M, Hb, eGFR and age (ordered in rank from the highest impact to the lowest impact).

Conclusion

Overall, this work presents a new reliable and inexpensive staging system, namely, MRS that utilizes easily acquirable laboratory parameters. It is valuable for the settings where genomic tests cannot be performed owing to economical and/or geographical constraints. The thresholds, proposed by this study, of laboratory parameters via KAP produce distinct PFS and OS patterns that are quantified by minimum *p*-value with better separation of MRS groups compared with those obtained with established thresholds and hence, can be adopted. 10-fold cross validation classification accuracy and ROC area confirm that our hierarchical stratification model can correctly classify patients into different risk groups. Application of machine learning techniques in MRS has led to better prediction of the survival outcome and identified different risk groups with distinct characteristics. The study recommends training of machine learning models on larger datasets because that can provide efficient upfront prognostication that may be useful in selection of therapy of appropriate intensity especially in high-risk MM patients. Further, the performance on MMRF dataset indicates that there may be nuances with respect to ethnicity and race. MRS model is trained on MMIn dataset using the new cut-offs proposed via KAP and is only tested on the MMRF dataset. Both the dataset belong to populations of different ethnicity and race. Therefore, the impact of ethnicity and race on risk staging ML models can be explored in future.

Data availability/calculator availability

An online version of the MRS calculator (https://github.com/AkankshaFarswan/MRS_Calculator) (Supplementary Fig S2) is also available which provides the risk stage of the patient based on the values of the six parameters-age, albumin, β2M, calcium, eGFR and hemoglobin.

Author contributions statement

Akanksha Farswan: Methodology, Software, Formal analysis, Investigation, Writing- Original draft preparation. **Anubha Gupta**: Methodology, Investigation, Validation, Writing- Original draft preparation, Resources, Project management, Supervision. **Ritu Gupta**: Conceptualization, Investigation, Validation, Resources, Writing- Original draft preparation, Project management, Supervision. **Saswati Hazra**: Formal analysis, Writing- Original draft preparation. **Sadaf Khan**: Resources. **Lalit Kumar:** Resources. **Atul Sharma:** Resources.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This work was supported by grant from Department of Biotechnology, Govt. of India [Grant: BT/MED/30/SP11006/2015] and Department of Science and Technology, Govt. of India [Grant: DST/ICPS/CPS-Individual/2018/279(G)]. Akanksha Farswan would like to thank the University Grants Commission, Govt. of India for the UGC-Senior Research Fellowship. Authors acknowledge MMRF and dbGaP (Project #18964) for providing the dataset. These data were generated as part of the Multiple Myeloma Research Foundation Personalized Medicine Initiative. Authors would also like to thank centre of Excellence in Healthcare, IIIT-Delhi for support in their research.

Role of funding source

The funding bodies had no role in study design, data collection, data analysis, data interpretation or writing of the report. The corresponding authors had full access to all the data used in the study and had final responsibility for the decision to submit for publication.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tranon.2021.101157.

References

[1] B.G. Durie, S.E. Salmon, A clinical staging system for multiple myeloma correlation of measured myeloma cell mass with presenting clinical features, response to treatment, and survival, Cancer 36 (3) (1975) 842–854.
[2] P.R. Greipp, J.S. Miguel, B.G. Durie, et al., International staging system for multiple myeloma, J. Clin. Oncol. 23 (15) (2005) 3412–3420.
[3] M.A. Dimopoulos, S. Delimpasi, E. Katodritou, et al., Significant improvement in the survival of patients with multiple myeloma presenting with severe renal impairment after the introduction of novel agents, Ann. Oncol. 25 (1) (2014) 195–200.
[4] G. Fouquet, B. Pegourie, M. Macro, et al., Safe and prolonged survival with long-term exposure to pomalidomide in relapsed/refractory myeloma, Ann. Oncol. 27 (5) (2016) 902–907.
[5] A. Palumbo, H. Avet-Loiseau, S. Oliva, et al., Revised international staging system for multiple myeloma: a report from International Myeloma Working Group, J. Clin. Oncol. 33 (26) (2015) 2863.
[6] H. Avet-Loiseau, X. Leleu, M. Roussel, et al., Bortezomib plus dexamethasone induction improves outcome of patients with t(4;14) myeloma but not outcome of patients with del(17p), J. Clin. Oncol. 28 (30) (2010) 4630–4634.
[7] Ricci Beatrice, M. van der Schaar, J. Yoon, et al., Machine learning techniques for risk stratification of non-ST-elevation acute coronary syndrome: the role of diabetes and age, Circulation 136 (suppl_1) (2017) A15892.
[8] K. Ahuja, M. van der Schaar, Risk-stratify: confident stratification of patients based on risk. arXiv preprint arXiv:1811.00753, 2018.
[9] B. Varghese, F. Chen, D. Hwang, et al., Objective risk stratification of prostate cancer using machine learning and radiomics applied to multiparametric magnetic resonance images, Sci. Rep. 9 (1) (2019) 1570.
[10] E. Bria, G. De Manzoni, S. Beghelli, et al., A clinical–biological risk stratification model for resected gastric cancer: prognostic impact of Her2, Fhit, and APC expression status, Ann. Oncol. 24 (3) (2013) 693–701.
[11] E.P. Hui, W.F. Li, B.B. Ma, et al., Integrating post-radiotherapy plasma Epstein-Barr virus DNA and TNM stage for risk stratification of nasopharyngeal carcinoma to adjuvant therapy, Ann. Oncol. 31 (6) (2020) 769–779.
[12] R. Gupta, G. Kaur, L. Kumar, et al., Nucleic acid based risk assessment and staging for clinical practice in multiple myeloma, Ann. Hematol. 97 (12) (2018) 2447–2454.
[13] S. Kumar, B. Paiva, K.C. Anderson, et al., International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma, Lancet Oncol. 17 (8) (2016) e328–e346.
[14] C.M. Florkowski, J.S. Chew-Harris, Methods of estimating GFR–different equations including CKD-EPI, Clin. Biochem. Rev. 32 (2) (2011) 75.
[15] S.V. Rajkumar, Multiple myeloma: 2016 update on diagnosis, risk-stratification, and management, Am. J. Hematol. 91 (7) (2016) 719–734.
[16] S.H. Eo, H.J. Kang, S.M. Hong, H. Cho, K-adaptive partitioning for survival data, with an application to cancer staging. arXiv preprint arXiv:1306.4615, 2013.
[17] C. Goswami, S. Poonia, L. Kumar, et al., Staging system to predict the risk of relapse in multiple myeloma patients undergoing autologous stem cell transplantation, Front. Oncol. 9 (July 2019) 633, https://doi.org/10.3389/fonc.2019.00633.
[18] G. Kaur, R. Gupta, N. Mathur, et al., Clinical impact of chromothriptic complex chromosomal rearrangements in newly diagnosed multiple myeloma, Leuk. Res. 76 (2019) 58–64, https://doi.org/10.1016/j.leukres.2018.12.005.

frontiers in Oncology

ORIGINAL RESEARCH
published: 09 November 2021
doi: 10.3389/fonc.2021.720932

Does Ethnicity Matter in Multiple Myeloma Risk Prediction in the Era of Genomics and Novel Agents? Evidence From Real-World Data

Akanksha Farswan[1], Anubha Gupta[1], Krishnamachari Sriram[2], Atul Sharma[3], Lalit Kumar[3] and Ritu Gupta[4]**

[1] Signal Processing and Biomedical Imaging Lab (SBILab), Department of Electronics and Communication, Indraprastha Institute of Information Technology-Delhi, New Delhi, India, [2] Department of Computational Biology, Indraprastha Institute of Information Technology-Delhi, New Delhi, India, [3] Department of Medical Oncology, Dr. B.R.A. IRCH, AIIMS, New Delhi, India, [4] Laboratory Oncology Unit, Dr. Bhim Rao Ambedkar Institute Rotary Cancer Hospital, All India Institute of Medical Sciences (Dr. B.R.A. IRCH, AIIMS), New Delhi, India

OPEN ACCESS

Edited by:
Varsha Gandhi,
University of Texas MD Anderson Cancer Center, United States

Reviewed by:
Stefan Knop,
Julius Maximilian University of Würzburg, Germany
Antonio Giovanni Solimando,
University of Bari Aldo Moro, Italy

***Correspondence:**
Anubha Gupta
anubha@iiitd.ac.in
Ritu Gupta
drritugupta@gmail.com;
drritu.laboncology@aiims.edu

Specialty section:
This article was submitted to Hematologic Malignancies, a section of the journal Frontiers in Oncology

Received: 05 June 2021
Accepted: 20 October 2021
Published: 09 November 2021

Citation:
Farswan A, Gupta A, Sriram K, Sharma A, Kumar L and Gupta R (2021) Does Ethnicity Matter in Multiple Myeloma Risk Prediction in the Era of Genomics and Novel Agents? Evidence From Real-World Data. Front. Oncol. 11:720932. doi: 10.3389/fonc.2021.720932

Introduction: Current risk predictors of multiple myeloma do not integrate ethnicity-specific information. However, the impact of ethnicity on disease biology cannot be overlooked. In this study, we have investigated the impact of ethnicity in multiple myeloma risk prediction. In addition, an efficient and robust artificial intelligence (AI)-enabled risk-stratification system is developed for newly diagnosed multiple myeloma (NDMM) patients that utilizes ethnicity-specific cutoffs of key prognostic parameters.

Methods: K-adaptive partitioning is used to propose new cutoffs of parameters for two different datasets—the MMIn (MM Indian dataset) dataset and the MMRF (Multiple Myeloma Research Foundation) dataset belonging to two different ethnicities. The Consensus-based Risk-Stratification System (CRSS) is designed using the Gaussian mixture model (GMM) and agglomerative clustering. CRSS is validated *via* Cox hazard proportional methods, Kaplan–Meier analysis, and log-rank tests on progression-free survival (PFS) and overall survival (OS). SHAP (SHapley Additive exPlanations) is utilized to establish the biological relevance of the risk prediction by CRSS.

Results: There is a significant variation in the key prognostic parameters of the two datasets belonging to two different ethnicities. CRSS demonstrates superior performance as compared with the R-ISS in terms of C-index and hazard ratios on both the MMIn and MMRF datasets. An online calculator has been built that can predict the risk stage of a multiple myeloma (MM) patient based on the values of parameters and ethnicity.

Conclusion: Our methodology discovers changes in the cutoffs with ethnicities from the established cutoffs of prognostic features. The best predictor model for both cohorts was obtained with the new ethnicity-specific cutoffs of clinical parameters. Our study also

revealed the efficacy of AI in building a deployable risk prediction system for MM. In the future, it is suggested to use the CRSS risk calculator on a large dataset as the cohort size of the present study is 25% of the cohort used in the R-ISS reported in 2015.

Keywords: AI in cancer research, ML in cancer survival, risk stratification of multiple myeloma, GMM clustering in cancer, consensus clustering in cancer, hematological malignancy

INTRODUCTION

Multiple myeloma is a hematopoietic malignancy of plasma cells with an overall survival period ranging from 6 months to more than 10 years. The variability in the outcome of patients is an implication of the clinical and biological heterogeneity underlying multiple myeloma (MM). Substantial advances in tumor biology have made it possible to dissect the tumor heterogeneity present in MM, optimize patient treatment, and examine patient outcome. Multiple prognostic systems (1–5) have been described in MM that stratify patients into different risk groups. These risk groups further assist in identifying high-risk patients who may require intense therapy upfront and/or a higher monitoring frequency during the follow-up periods. The first staging system for MM was proposed in 1975 (1) followed by the development of the International Staging System (ISS) (2) in 2005 and a Revised ISS (R-ISS) (3) in 2015. The ISS utilizes serum albumin and beta2-microglobulin, while the R-ISS makes use of ISS, lactate dehydrogenase (LDH), and high-risk cytogenetic aberrations (HRCA). Currently, triplet combination therapy is the new standard of care in MM which has shifted many high-risk patients to standard-risk category, thereby justifying the need for a new risk-stratification system with the possibility of inclusion of more prognostic factors.

Although human physiological and genetic profile is known to vary across ethnic groups, the current MM risk-staging systems do not account for ethnicity-specific information that can have a huge impact on the risk score prediction. It is evident from the studies that African Americans experience two to three times higher incidence rates than Asians, Mexican-Americans, or Europeans (6). Recent studies have observed a significant variation in the overall survival of different groups belonging to distinct races/ethnicities since the introduction of novel treatment agents in MM (7–10). In a recent study, vitamin D deficiency at diagnosis was found to be a predictor of poor overall survival in MM (11). However, this was significant only for White Americans and not for African Americans even at lower cutoffs of deficiency (11). Similarly, HRCA, which is used to determine the intensity of frontline therapy, does not track with survival outcomes in African Americans (10), thereby highlighting the need for a race-specific risk-stratification system. Though ethnicity is an important prognostic factor in predicting the risk for MM (12), the variations in the clinical characteristics among the different ethnic groups have not been evaluated adequately. Therefore, it is desirable to have a staging system that includes the variations in the clinical characteristics of the patients pertaining to distinct ethnic groups. In addition, it should be based on clinical and laboratory parameters that are easily accessible in healthcare settings across the globe. Therefore, to address this concern, we first investigated the role of ethnicity in the differential clinical characteristics in the two independent cohorts of MMIn and MMRF patients with newly diagnosed multiple myeloma (NDMM) belonging to two separate ethnic groups. Furthermore, we proposed the Consensus based Risk-Stratification System (CRSS), an AI-enabled risk-stratification system, for NDMM that incorporates the ethnicity-specific cutoffs of the laboratory parameters like albumin, beta-2 microglobulin (β2M), calcium, estimated glomerular filtration rate (eGFR), hemoglobin, and age along with HRCA. The newly proposed ethnicity-aware AI-assisted CRSS method was shown to have superior performance as compared with R-ISS. In addition, we also interpreted our proposed model *via* SHapley Additive exPlanations (SHAP) (13) analysis to demonstrate the clinical significance of the risk stage predictions by CRSS. Our findings establish the significance of integrating ethnicity-specific information as well as the effectiveness of machine learning methods in devising a robust risk-staging model for MM.

MATERIALS AND METHODS

Datasets

A total of 1,675 entries were found in the computerized database search on June 28, 2019, with the keyword "ICD C90" registered at the Institute Rotary Cancer Centre, All India Institute of Medical Sciences (AIIMS). Patients with plasma cell dyscrasia other than MM (n = 253) or who were lost to follow-up after a single visit (n = 111) or before first response could be assessed (n = 21) or with inadequate clinical and/or laboratory parameters (n = 121) or with early deaths (n = 99) were excluded. The remaining 1,070 patients of MM belonging to the Indian population, referred to as MMIn, were evaluated in this study (**Figure S1**). Out of 1,070 patients, 41 patients had one or two missing values. There are several methods to impute missing values (14–17). However, in the MMIn dataset, missing values were imputed with the median value of the parameters. An independent cohort of 900 MM patients enrolled in the Multiple Myeloma Research Foundation (MMRF) repository was also used for developing the model. Clinical and laboratory data for the MMRF dataset, belonging to the American population, are available publicly. High-risk cytogenetic information was available for 384 out of 1,070 patients in the MMIn cohort and 800 out of a total of 900 patients in the MMRF which were further used for building the staging model.

Clinical and Laboratory Characteristics

The clinical, laboratory, and radiological data were obtained from the medical case files. The R-ISS could be assigned to a subset of patients (n = 627) as described previously (18). Response outcome was estimated following the international uniform response criteria for multiple myeloma (19). Progression-free survival (PFS) was computed from the date of diagnosis till the time of progression or death. Overall survival (OS) was computed from the date of diagnosis till death due to any cause or being censored at last follow-up. Baseline clinical and laboratory features of the patients are given in **Supplementary Table S1**.

Study Design

The complete design strategy of the consensus-based approach for developing the risk-stratification system (CRSS) is explained in this section (**Figure 1**). Data from both cohorts were separately used to develop the risk-staging models based on CRSS. Different clinical parameters were evaluated for developing the risk-staging system consisting of age, albumin, β2M, calcium, eGFR, hemoglobin, LDH, and HRCA which includes t(4;14), t(14;16), and del17. β2M and LDH levels are reflective of tumor burden and serum albumin, hemoglobin, calcium, and creatinine are reflective of the bone and renal homeostasis. eGFR was calculated from creatinine concentration using the MDRD eGFR equation (20). LDH values were brought to a common scale by multiplying each entry by 280 and dividing it by the upper limit of LDH provided for that particular entry in MMIn data. Description of the steps used in the consensus-based approach for developing the risk-staging model is given below:

Step 1: Dividing patients into two risk groups based on established thresholds of parameters. For each parameter, patients were initially divided into high-risk and low-risk groups using the well-established cutoffs of these parameters (21) as shown in **Table 1**. Established thresholds for albumin and β2M are derived from the ISS, and for eGFR, calcium, and hemoglobin, the thresholds are derived from the revised IMWG criteria (21).

Step 2: Finding new thresholds of parameters via KAP. The K-adaptive partitioning (22) (KAP) algorithm was used to find new threshold values for the parameters using complete data of MMIn (n = 1,070) and MMRF (n = 900). KAP was performed on the parameters of the patients yielding two threshold values for each parameter, one from PFS and the other from OS analysis. The cutoff which was close to the original value was chosen as the new cutoff for each parameter. Patients were again divided into high- and low-risk groups based on the proposed cutoffs. The proposed thresholds maximized the separation between high- and low-risk groups as compared with the established thresholds. This is evident from the lower p-values obtained from the log-rank test on the Kaplan–Meier curves for all the parameters. A complete list of the proposed thresholds for the MMIn and MMRF data is shown in **Table 1**.

Step 3: Cumulative integration of the prognostic impact of the parameters. The collective prognostic impact of the parameters was integrated into risk staging *via* creation of three different adjacency graphs using hazard ratios obtained from univariate Cox hazard analysis, p-values obtained from log-rank test on Kaplan–Meier curves, and ranks obtained from multivariate Cox hazard analysis.

Step 4: Creation of the first adjacency graph. The first adjacency graph was created using ranks obtained from the multivariate Cox hazard analysis. The parameter with the highest hazard value was given the highest rank, and the one with the lowest hazard value was given the lowest rank. The respective ranks served as the weights of each of the parameters and captured the relative impact of each parameter on the survival of patients. Next, the risk score for each patient was calculated by successive addition of the weights of all those parameters that had values (in the respective patient) greater than the cutoffs defined for the high-risk group. These patient scores were used to compute an adjacency graph of n rows and n columns (columns are features), where n is the number of patients. Each row corresponds to one patient and each entry in the row is the absolute difference between the score of that patient with each of the patients including self.

Step 5: Creation of the second and third adjacency graphs. For the second adjacency graph, hazard ratio values obtained from univariate Cox hazard analysis were used. For each parameter, the highest of the two HR values obtained from PFS and OS was chosen and normalized using "minmax" scaling. The scaled HR values were assigned as the respective weights of each of the parameters representing the impact of each parameter on the survival of patients. The third adjacency graph was created using p-values obtained by performing a log-rank test on Kaplan–Meier curves. For each parameter, the lower of the two p-values obtained from PFS and OS was chosen and normalized using "minmax" scaling. The scaled p-values were assigned as the respective weights of each of the parameters. Furthermore, the risk score for each patient was calculated by successive addition of the weights of all those parameters that had values (in the respective patient) greater than the cutoff defined for the high-risk group. The two different patient scores obtained from univariate hazard ratios and p-values were further used to compute two separate adjacency graphs of n rows and n columns (columns are features), where n is the number of patients. Each row corresponds to one patient and each entry in the row is the absolute difference between the score of that patient with each of the patients including self.

Step 6: Gaussian mixture model (GMM) clustering on the adjacency graphs. GMM-based clustering is an unsupervised clustering algorithm which was applied on the three adjacency graphs to obtain clustering labels.

Step 7: Creation of a consensus graph. The clustering outputs of the three different adjacency graphs were used to create a consensus graph (23) of size $n \times n$. The entry for the ith row and jth column in the consensus graph was determined by calculating the number of times ith and jth patients were assigned the same group. Diagonal entries were zero in this graph.

Step 8: Hierarchical clustering on the consensus graph. Agglomerative clustering was performed on the consensus graph to cluster the patients into three risk groups. Each

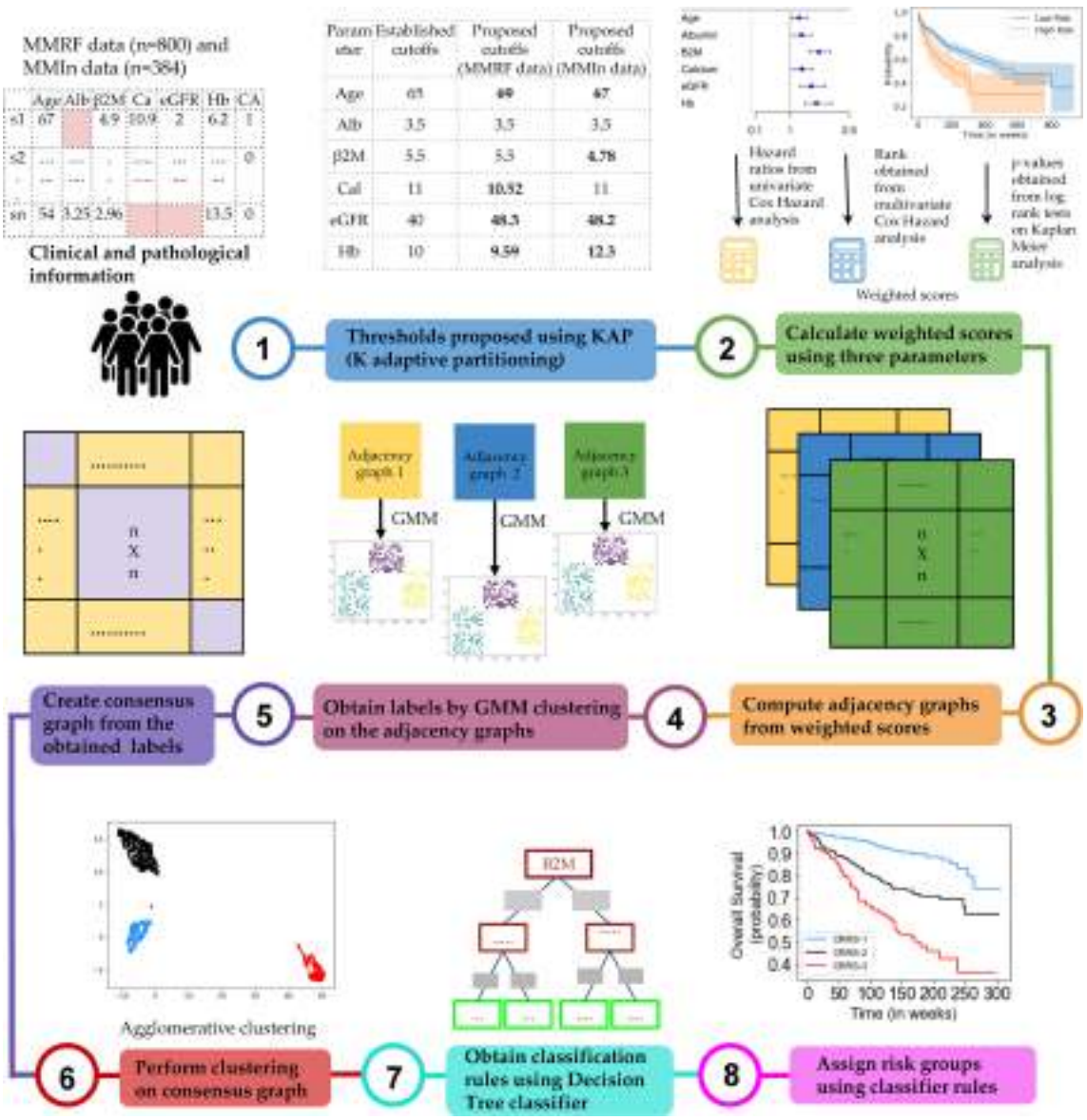


FIGURE 1 | Workflow for the development of the Consensus-based Risk-Stratification System (CRSS) for newly diagnosed multiple myeloma patients.

cluster of patients was assigned one label: stage 1 (low risk), stage 2 (intermediate risk), or stage 3 (high risk). The rationale behind using multiple clustering was to combine the results of the clustering outputs achieved from the different adjacency graphs and ensure the stability of the final clusters deduced from agglomerative clustering.

Step 9: Training a decision tree classifier. The staging labels obtained from agglomerative clustering served as ground-truth labels for training the supervised decision tree classifier. The trained decision tree classifier provided the rules in terms of the parameters for the identification of risk groups, labeled as CRSS-1 (low risk), CRSS-2 (intermediate risk), and CRSS-3 (high risk) (**Figures S2, S3**).

Step 10: Infer actual risk groups of the patients using decision tree classifier rules. Decision tree classifier rules were then used to identify the risk stages of the patients in both cohorts. The risk stage assigned by the decision tree classifier was considered the actual risk class for each patient.

Creation of Multiple Models on the Datasets

The CRSS method explained in **Figure 1** was used to create multiple models for the MMIn and MMRF datasets. Models A1, A2, and A3 were built for the MMIn data. Model A1 was built using established cutoffs of the parameters of albumin, β2M, LDH, and HRCA. Model A2 was built using the established

TABLE 1 | Comparison of established and proposed cutoffs for clinical and laboratory parameters for the stratification of patients for progression-free survival (PFS) and overall survival (OS) in MMIn and MMRF using Kaplan–Meier analysis.

Parameter	Established cutoff value	Proposed cutoff value	PFS		OS	
			p-value with established cutoff	*p*-value with proposed cutoff	*p*-value with established cutoff	*p*-value with proposed cutoff
MMIn (*n* = 1,070)						
Age (years)	>65	**>67**	0.11	**0.012**	5.84e-5	**1.25e-6**
Albumin (g/dl)	≤3.5	≤3.5	0.115	0.115	7.0e-4	7.0e-4
β2M (mg/L)	≥5.5	**≥4.78**	**8.15e-10**	9.32e-10	4.13e-10	**4.53e-14**
Calcium (mg/dl)	≥11	≥11	0.0078	0.0078	0.0037	0.0037
eGFR (ml/min/1.73m^2)	≤40	**≤48.2**	0.16	**0.04**	0.005	**1.5e-4**
Hb (g/dl)	≤10	**≤12.3**	0.0019	**8.56e-5**	0.0014	**3.75e-7**
MMRF (*n* = 900)						
Age (years)	>65	**>69**	3.23e-05	**1.98e-08**	1.06e-05	**1.58e-09**
Albumin (g/dl)	≤3.5	≤3.5	0.00017	0.00017	8.47e-07	8.47e-07
β2M (mg/L)	≥5.5	≥5.5	1.22e-10	1.22e-10	9.25e-13	9.25e-13
Calcium (mg/dl)	≥11	**≥10.52**	0.0077	**1.40e-04**	5.88e-06	**3.49e-06**
eGFR (ml/min/1.73m^2)	≤40	**≤48.3**	4.5e-05	**4.67e-09**	7.48e-06	**2.48e-10**
Hb (g/dl)	≤10	**≤9.59**	2.82e-06	**5.69e-09**	6.77e-06	**5.42e-07**

The proposed cutoffs were found using complete data of MMIn (n = 1,070) and MMRF (n = 900). Less than or equal to cutoff reveals the increased risk in the patient. ">65" shows that a patient with age greater than 65 years is at greater risk than a patient less than 65 years. "≤3.5" shows that a patient with albumin levels less than equal to 3.5 is at a greater risk than a patient with albumin levels greater than 3.5. It holds true for other parameters also in a similar manner. Bold values of the column "proposed cutoff value" signify the change in the value of the parameters from the existing cut-offs. p-values in bold signify that p-values became more significant with the proposed changes in cutoffs.

cutoffs of the parameters of albumin, age, calcium, eGFR, hemoglobin, β2M, and HRCA. Model A3 uses the same parameters as model A2, but with the newly proposed cutoffs of the parameters derived from the MMIn dataset. Similarly, models M1, M2, M3, and M4 were built for the MMRF data. Models M1 and M2 are equivalent to models A1 and A2, respectively. For model M3, the proposed cutoffs of parameters derived from the MMIn dataset were used for albumin, age, calcium, eGFR, hemoglobin, β2M, and HRCA. Model M4 is similar to model M3, but uses the proposed cutoffs of the parameters derived from the MMRF dataset.

RESULTS

Clinical and Laboratory Characteristics of Myeloma Patients

The baseline clinical and laboratory features of patients from the two cohorts were compared using unpaired Wilcoxon rank-sum test. The median values of all the parameters except albumin were found to be significantly different (p-value < 0.05, **Table S2**) in both cohorts thereby substantiating that the two populations are different. Novel agents (IMIDs: thalidomide or lenalidomide and/or PSI, i.e., bortezomib) either as primary or maintenance therapy were given to all the patients. Triplet therapy was rendered to 56.5% of the patients. With a median follow-up of 166 weeks (range: 14–961 weeks), 626 patients progressed (median PFS = 117 weeks) and 372 died (median OS = 166 weeks).

Results on the MMIn Dataset (*n* = 384)

Univariate Cox analysis of the entire patient cohort (n = 1,070, **Table S3**, **Figure 2**) revealed increased risk of progression and mortality for age >67 years, albumin ≤3.5, β2M ≥4.78, calcium ≥11, eGFR ≤48.2, and hemoglobin ≤12.3. Multivariate Cox hazard analysis was also performed to analyze the cumulative risk of the parameters (**Table S4**). Of the three models generated, model A3 based on ML-derived cutoffs for the prognostic parameters was the best with higher C-index and hazard ratio (**Table 2**). Using model A3, the patients were risk stratified and the largest proportion of patients were placed in CRSS-2 (n = 192, 50%) followed by CRSS-1 (n = 137, 35.68%) and CRSS-3 (n = 55, 14.32%). KM survival analysis of CRSS groups indicated statistically significant difference in PFS between CRSS-1 and CRSS-2 groups (median PFS: 213 vs. 138 weeks; p = 0.0003) and between CRSS-2 and CRSS-3 groups (median PFS: 138 vs. 100 weeks; p = 0.0026) (**Figure 2**). For R-ISS, there was a statistically significant difference in PFS between R-ISS2 and R-ISS3 (median PFS: 160 vs. 105 weeks; p = 0.01) but not between R-ISS1 and R-ISS2 (median PFS = 196 vs. 160 weeks; p = 0.31). Furthermore, for CRSS, there was statistically significant difference in OS between CRSS-1 and CRSS-2 groups (median OS = 495 vs. 249 weeks; p = 1.08e-8) as well as between CRSS-2 and CRSS-3 groups (median OS = 249 vs. 182 weeks; p = 0.02). For R-ISS, there was statistical difference in OS between R-ISS2 and R-ISS3 groups (median OS = 377 vs. 168 weeks; p = 1.86e-5) as well as between R-ISS1 and R-ISS2 groups (median OS = 478 vs. 377 weeks; p = 0.03).

C-statistic and hazard ratios computed on CRSS surpassed the C-index and hazard ratios obtained for R-ISS with respect to both PFS and OS (**Table 2**). C-statistic for CRSS was 0.60 [Akaike information criteria (AIC) = 2,171.49, Bayesian information criteria (BIC) = 2,175.43, HR = 1.80, 95% CI = 1.46–2.21, p < 5e-6] for PFS and 0.67 (AIC = 1,244.72, BIC = 1,248.67, HR = 2.43, 95% CI = 1.87–3.14, p < 5e-6) for OS, while C-statistic for R-ISS was 0.57 (AIC = 2,011.14, BIC = 2,015.01, HR = 1.43, 95% CI = 1.12–1.82, p = 4.18e-3) for PFS and 0.636

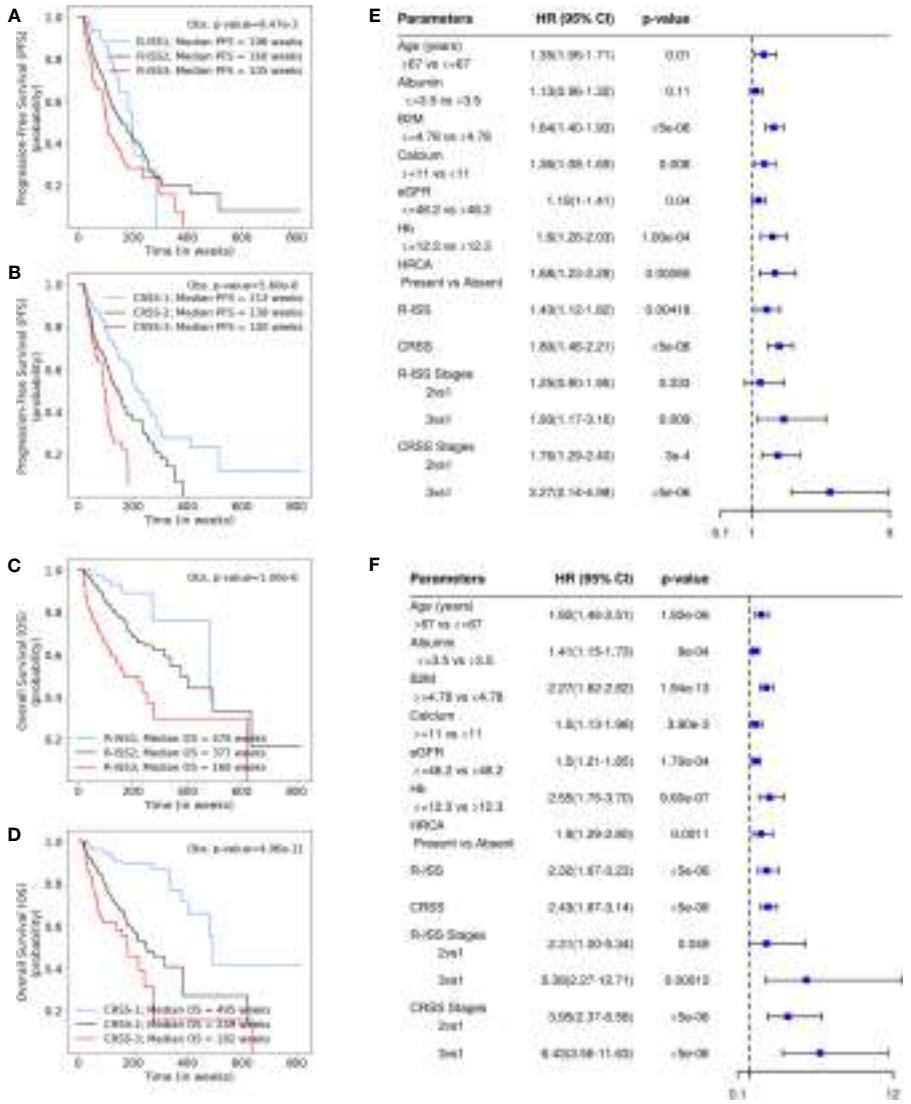


FIGURE 2 | **(A, B)** Progression-free survival in patients with multiple myeloma (MM) from the MMIn cohort (n = 1,070) stratified by the Revised International Staging System (R-ISS) (n = 355) and the proposed CRSS (n = 384), respectively. R-ISS1 is the low-risk stage, R-ISS2 is the intermediate-risk stage, and R-ISS3 is the high-risk stage. Median progression-free survival (PFS) for R-ISS1, R-ISS2, and R-ISS3 are 196, 160, and 105 weeks, respectively. The observed p-value obtained after performing a log-rank test on R-ISS is 9.47e-3. Similarly, CRSS-1 is the low-risk stage, CRSS-2 is the intermediate-risk stage, and CRSS-3 is the high-risk stage. Median PFS for CRSS-1, CRSS-2, and CRSS-3 are 213, 138, and 100 weeks, respectively. The observed p-value obtained after performing a log-rank test on CRSS is 5.60e-8. **(C, D)** Overall survival in patients with MM from the MMIn cohort (n = 1,070) stratified by the R-ISS (n = 355) and CRSS (n = 384), respectively. Median overall survival (OS) for R-ISS1, R-ISS2, and R-ISS3 are 478, 337, and 168 weeks, respectively. The observed p-value obtained after performing a log-rank test on R-ISS is 1.00e-6. Median OS for CRSS-1, CRSS-2, and CRSS-3 are 495, 249, and 182 weeks, respectively. The observed p-value obtained after performing a log-rank test on CRSS is 4.96e-11. **(E, F)** Univariate Cox hazard analysis on the prognostic factors—age, albumin, beta-2 microglobulin (β2M), calcium, estimated glomerular filtration rate (eGFR), hemoglobin, and high-risk cytogenetic abnormalities (HRCA)—for PFS and OS, respectively. Hazard ratios for all the parameters except HRCA were calculated on complete data (n = 1,070) for the MMIn dataset. Hazard ratio for HRCA and the risk-staging models were found using the data for which HRCA information was present (n = 384 for the MMIn dataset).

(AIC = 1,132.20, BIC = 1,136.07, HR = 2.32, 95% CI = 1.67–3.23, p < 5e-6) for OS.

Results on the MMRF Dataset (n = 800)

For the MMRF data, out of the four models generated, model M4 performed the best and had the highest C-index and hazard ratios as compared with the other models as well as R-ISS (**Table 2**). In the univariate Cox hazard analysis of the MMRF data, risk of progression and mortality was increased for age >69 years, β2M ≥5.5, albumin ≤3.5, hemoglobin ≤9.59, eGFR ≤48.3, and calcium ≥10.52 (**Table S3**, **Figure S4**). Multivariate Cox hazard analysis was also performed (**Table S4**). In the MMRF cohort, using the M4 model, the majority of the patients were placed in CRSS-2 (n = 452, 56.5%) followed by CRSS-3 (n = 174, 21.75%) and CRSS-1 (n = 174, 21.75%). Results of

TABLE 2 | Comparison of different models devised for the risk stratification of patients in the MMIn and MMRF cohorts with the R-ISS.

		PFS			OS		
		Hazard ratio	***p*-value**	**C-index**	**Hazard ratio**	***p*-value**	**C-index**
MMIn (*n* = 384)							
R-ISS (*n* = 355)		1.42	0.004	0.57	2.32	<5e-6	0.636
	2vs1	1.24	0.33		2.31	0.04	
	3vs1	1.92	0.009		5.37	0.00013	
Model A1		1.5	1.00e-5	0.594	2.03	<5e-6	0.646
	2vs1	1.53	0.007		2.13	0.0013	
	3vs1	2.26	2.00e-5		4.16	<5e-6	
Model A2		1.4	0.0001	0.579	1.74	1.00e-5	0.616
	2vs1	1.42	0.056		1.9	0.02	
	3vs1	1.98	0.00013		3.13	2.00e-5	
Model A3 (CRSS)		**1.8**	<5e-6	**0.6**	**2.43**	<5e-6	**0.67**
	2vs1	**1.76**	3.00e-4		**3.95**	<5e-6	
	3vs1	**3.27**	<5e-6		**6.43**	<5e-6	
MMRF (*n* = 800)							
R-ISS (*n* = 658)		1.61	0.00001	0.578	2.26	<5e-6	0.618
	2vs1	1.49	0.015		1.79	0.03	
	3vs1	2.6	0.00001		4.66	<5e-6	
Model M1		1.55	<5e-6	0.6	2.07	<5e-6	0.656
	2vs1	1.55	0.00042		2.06	0.00067	
	3vs1	2.4	<5e-6		4.3	<5e-6	
Model M2		1.62	<5e-6	0.6	2.36	<5e-6	0.657
	2vs1	1.44	0.01		2.12	0.0081	
	3vs1	2.54	<5e-6		5.22	<5e-6	
Model M3		1.54	<5e-6	0.604	2.2	<5e-6	0.679
	2vs1	1.87	<5e-6		2.95	<5e-6	
	3vs1	2.32	<5e-6		5.11	<5e-6	
Model M4 (CRSS)		**1.79**	<5e-6	**0.61**	**2.85**	<5e-6	**0.676**
	2vs1	**1.76**	8.10e-4		**4.1**	3.40e-4	
	3vs1	**3.19**	<5e-6		**10.61**	<5e-6	

Models were built using data for which high-risk cytogenetic information (HRCA) was available (n = 384 for MMIn and n = 800 for MMRF). R-ISS information was available for only 355 out of 384 patients in the MMIn dataset and 658 out of 800 patients in the MMRF dataset. The model with the best performance was A3 and M4 (in bold).
Model A1: beta-2 microglobulin (β2M), albumin, LDH, and CA [del17, t(4;14), t(14;16)] at existing cutoffs. Model A2: age, β2M, albumin, calcium, estimated glomerular filtration rate (eGFR), Hb, and HRCA using existing cutoffs. Model A3: age, β2M, albumin, calcium, eGFR, Hb, and HRCA using proposed cutoffs for MMIn data. Model M1: β2M, albumin, LDH, and HRCA at existing cutoffs. Model M2: age, β2M, albumin, calcium, eGFR, Hb, and HRCA using existing cutoffs. Model M3: age, β2M, albumin, calcium, eGFR, Hb, and HRCA using proposed cutoffs for MMIn data. Model M4: age, β2M, albumin, calcium, eGFR, Hb, and HRCA using proposed cutoffs for MMRF data.

the median PFS on CRSS groups (p = 8.64e-12) and R-ISS groups (p = 1.73e-5) as well as median OS on CRSS groups (p = 1.08e-15) and R-ISS groups (p = 6.57e-8) reveal the superior performance of the CRSS than the R-ISS (significant p-values; **Figure S4**).

C-statistic for CRSS in MMRF data is 0.61 (AIC = 4,126.07, BIC = 4,130.74, HR = 1.79, 95% CI = 1.52–2.12, p < 5e-6) for PFS and 0.676 (AIC = 1,819.95, BIC = 1,824.62, HR = 2.85, 95% CI = 2.19–3.71, p < 5e-6) for OS. C-statistic for R-ISS is 0.578 (AIC = 3,413.36, BIC = 3,416.49, HR = 1.61, 95% CI = 1.30–2.00, p = 1.00e-5) for PFS and 0.618 (AIC = 1,586.78, BIC = 1,591.27, HR = 2.26, 95% CI = 1.65–3.11, p < 5e-6) for OS (**Table 3**).

The 5-year OS for the MMIn (n = 384) was 89.79% for CRSS-1, 47.91% for CRSS-2, and 31.36% for CRSS-3 (**Table 3**). Overall, there is a substantial difference in the percentages of the 5-year OS and median OS for different risk groups which indicate that the groups were significant. A similar stratification was achieved when the CRSS model was applied on the MMRF test dataset. The 5-year OS for MMRF data was 94.78% for CRSS-1, 65.69% for CRSS-2, and 46.91% for CRSS-3 which is quite comparable to that obtained in the MMIn data. Higher values of C-index and hazard ratios as well as lower values of partial AIC and BIC on both datasets were indicative of the superior performance of our AI-based CRSS method as compared with R-ISS.

Statistical Analysis on the Parameters Used in CRSS

The Kruskal–Wallis test was performed to compare the median values of the parameters age, albumin, β2M, calcium, eGFR, and hemoglobin across the three risk groups for both the MMIn and MMRF datasets. There was a significant increase (p < 0.05) in the values of age and β2M, while there was a significant decrease (p < 0.05) in the values of albumin, eGFR, and hemoglobin as the risk of disease increased (**Figures S5, S6**) for both the MMIn and MMRF datasets. Wilcoxon rank-sum test was performed to compare the median values of the parameters between two successive risk groups and showed significant variation of parameters for both datasets.

Model Interpretation

To ascertain the impact of individual parameters on risk stage predictions by CRSS, decision tree models built using the MMIn and MMRF datasets were analyzed using SHAP (**Figures 3**, **7**). Key

TABLE 3 | Prediction of progression-free survival and overall survival (in %) for CRSS and R-ISS at 1, 2, 3, 4, and 5 years in the MMIn (n = 384) and MMRF datasets (n = 800).

		MMIn data					
		R-ISS (n = 355)			CRSS (n = 384)		
	Year	1	2	3	1	2	3
PFS	1	0.9318	0.8305	0.6967	0.8966	0.7812	0.7196
	2	0.8606	0.6601	0.5223	0.7709	0.6265	0.4472
	3	0.6404	0.5124	0.3632	0.6449	0.4729	0.2515
	4	0.3422	0.4179	0.2810	0.5251	0.3624	0.0587
	5	0.2738	0.2856	0.2342	0.4014	0.2679	0.0587
OS	1	0.9773	0.9387	0.7784	0.9630	0.8938	0.7976
	2	0.9540	0.8415	0.6393	0.9466	0.7679	0.6155
	3	0.9282	0.7764	0.5342	0.9098	0.6702	0.5831
	4	0.8895	0.6790	0.4953	0.8979	0.5691	0.4574
	5	0.8895	0.6422	0.3698	0.8979	0.4791	0.3136
		MMRF data					
		R-ISS (n = 658)			CRSS (n = 800)		
	Year	1	2	3	1	2	3
PFS	1	0.9033	0.8132	0.6358	0.9325	0.8367	0.6611
	2	0.7957	0.6261	0.4040	0.8162	0.6734	0.4423
	3	0.6295	0.4862	0.3059	0.7008	0.5084	0.3129
	4	0.4641	0.3414	0.2781	0.5151	0.3711	0.2249
	5	0.2769	0.2450	0.2781	0.4121	0.2637	0.1799
OS	1	0.9807	0.9092	0.8559	0.9869	0.9379	0.8231
	2	0.9612	0.8372	0.6460	0.9689	0.8772	0.6780
	3	0.9286	0.7799	0.5211	0.9478	0.8217	0.5814
	4	0.8833	0.7461	0.4904	0.9478	0.7844	0.5293
	5	0.5748	0.7108	0.3678	0.9478	0.6569	0.4691

contributors of high-risk predictions in the MMIn dataset were the presence of HRCA, elevated levels of β2M, higher age, and lower levels of albumin (**Figure 3**). Furthermore, lower levels of eGFR and hemoglobin along with elevated levels of calcium also contributed to high-risk prediction in the patients. It was observed from the waterfall plots (**Figures 4**–**6**) of the randomly chosen patients in different risk stages that the order of the impact of the parameters varied in different patients within the same risk category. For the high-risk category (**Figure 6**), HRCA had the highest impact on one of the randomly chosen patients; in another patient, β2M had the highest impact in contributing to high risk, while in the third patient, age and albumin had the highest prognostic impact. This suggests that the risk assessment in MM is a cumulative function of multiple factors. An individual parameter cannot adequately capture the risk associated with MM given that other prognostic parameters could influence the outcome. Furthermore, the complex association among different parameters that encapsulates the disease risk varies according to the patients, thereby leading to a varying order of impact of parameters in the patients. Hence, the AI-based decision tree algorithms can handle such an integrated analysis. This analysis reveals that each patient is unique and multiple factors interact and impact the outcome differently in individual patients.

DISCUSSION

The influence of ethnicities on clinical characteristics in patients belonging to distinct ethnic groups is well known, and therefore, it is of paramount interest to integrate the ethnic group-specific information in risk-staging models as it can affect the risk score prediction. The R-ISS (3) is the current standard of care for staging myeloma patients which includes a few HRCA, but molecular aberrations such as 1q gain and chromothripsis associated with adverse outcome have been overlooked (24). In fact, it includes t (4;14), which has lost significance in patients treated with triplet regimens (25). Besides, the R-ISS does not include any ethnic-specific information and, therefore, is not robust considering the large heterogeneous population of MM patients globally. An ideal risk-staging system would be based on all the known adverse prognostic factors including clinical, ethnic, and molecular aberrations. There is a tremendous heterogeneity in global healthcare systems that limit the availability of high-end molecular testing for all patients, and yet, the internet/electronic connectivity allows patients to receive medical advice from global leaders in medicine. Recently, an AI-supported risk-staging model, MRS (26), has been developed for NDMM; however, it does not include HRCA and ethnicity information. Considering the present world scenario, it is, thus, desirable to develop a simple risk-staging model that integrates ethnic-specific characteristics of the prognostic parameters that are easy to acquire in the healthcare settings worldwide.

Risk-Staging Models and Their Performance as Compared With the R-ISS

In contrast to the R-ISS which utilizes four parameters, seven parameters were taken into consideration for designing the

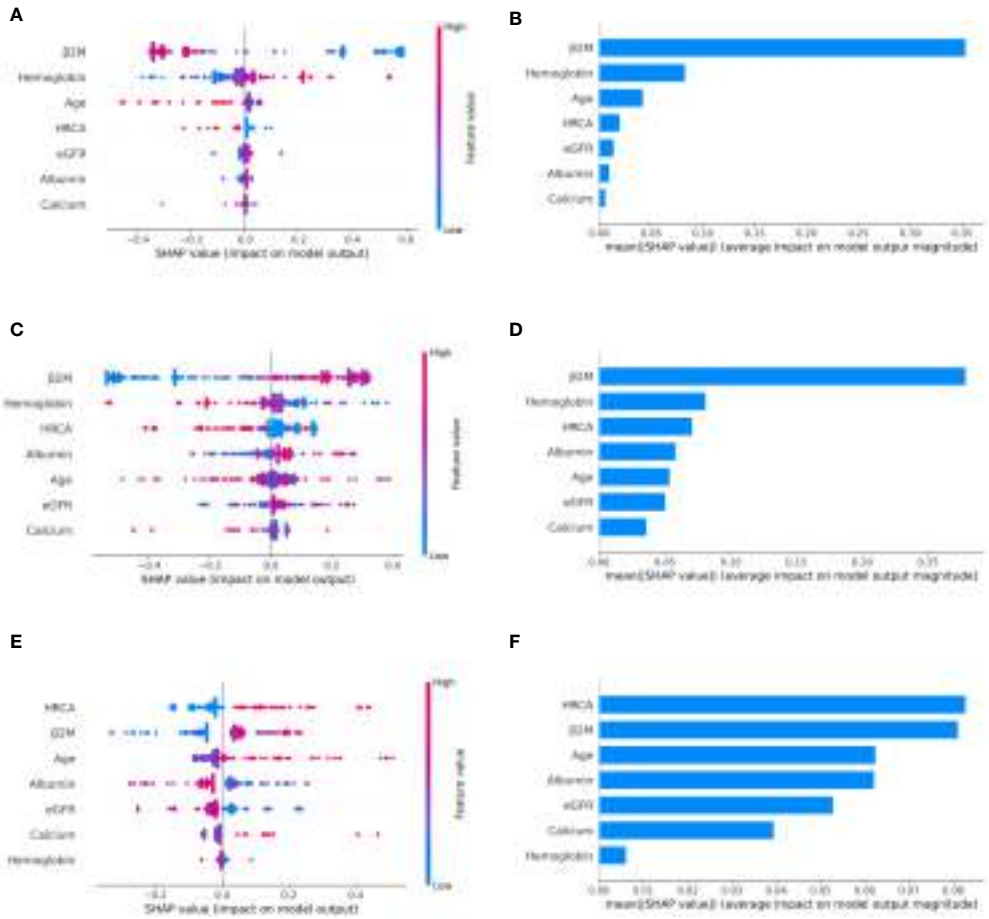


FIGURE 3 | Model interpretation using SHAP (SHapley Additive exPlanations). SHAP summary plots for different risk stages inferred from MMIn data showing the relative impact of different parameters (top to bottom) contributing to a particular risk stage prediction. **(A, B)** CRSS-1: Normal levels of β2M and hemoglobin are the key contributors to the low-risk stage prediction. Furthermore, high values of age on the left side of the summary plot are pushing the model away from the low-risk prediction and are indicative of either intermediate or high risk. Overall, β2M has the highest impact and calcium has the lowest impact on the low-risk stage prediction. **(C, D)** CRSS-2: β2M and hemoglobin are the key contributors to the intermediate-risk stage. Elevated levels of β2M with lower levels of hemoglobin are indicative of intermediate risk. **(E, F)** CRSS-3: Presence of HRCA is contributing the most to the high-risk stage. Elevated values of β2M and calcium and lower levels of albumin, hemoglobin, and eGFR are contributing toward the high-risk stage prediction.

CRSS. It was observed that the cutoff values for these parameters derived using KAP vary in the two cohorts, one of which belongs to Indian and the other belongs to the American population. For the Indian data, there was a change in the cutoff values for β2M, age, eGFR, and hemoglobin, while there was no change in the cutoff value for calcium and albumin as shown in **Table 1**. For the MMRF data, there was a change in cutoff values for calcium, eGFR, hemoglobin, and age, while the cutoff values for albumin and β2M remain unchanged. The median age of onset of MM in the Indian population is almost a decade early as compared with the population in the USA (27, 28). This supported our assertion of choosing different cutoffs of age for MMIn from the MMRF dataset.

Various models were built on the different combinations of the parameters using both the established and proposed cutoffs for the two datasets. The best staging model for both datasets was obtained when the proposed cutoffs for the respective cohorts were used. When the ML-derived cutoffs were used for the parameters age, eGFR, hemoglobin, and β2M in the A3 model, performance was enhanced significantly in terms of high C-index and hazard ratios as compared with the R-ISS. A similar observation was noticed in the M4 model which utilized ML-derived cutoffs obtained for the MMRF dataset and achieved the best performance among all the models with a significant improvement in the C-index as well as hazard ratios as compared with the R-ISS. Overall, A3 and M4 were the best staging models for the MMIn and MMRF data, respectively. The improvement in the performance of the model verified our hypothesis that the cutoffs of the different parameters vary with different ethnicities.

The plausibility of the proposed model was further substantiated by performing significance testing. The Kruskal–Wallis test showed statistically significant variations ($p < 0.05$) in

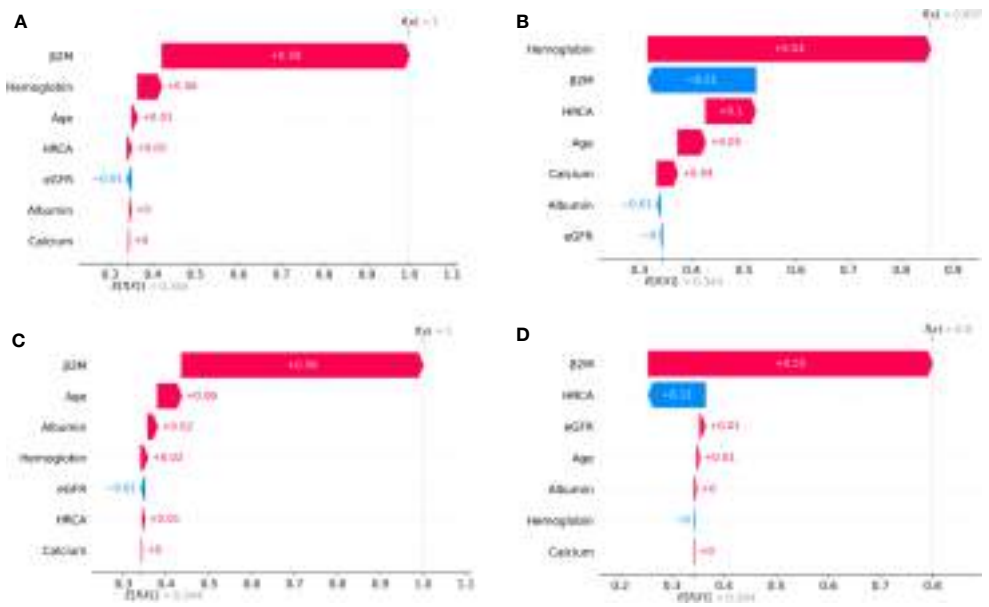


FIGURE 4 | SHAP waterfall plots for the randomly chosen four patients in low-risk stage (CRSS-1) from the MMIn dataset. The pink color shows the positive impact of the feature, while the blue color shows the negative impact of the feature. Features with a positive impact contributed to the class of low-risk stage prediction, while features with a negative impact contributed to class opposite to low risk. β2M, hemoglobin, age, and HRCA have the highest overall impact on low-risk stage prediction in the MMIn dataset. However, this ranking itself differs from patient to patient as can be seen in **(A–D)**. **(A)** β2M has the highest impact followed by hemoglobin, age, and HRCA. **(B)** Hemoglobin has the highest impact followed by β2M and age. **(C, D)** β2M has the highest impact followed by age and HRCA.

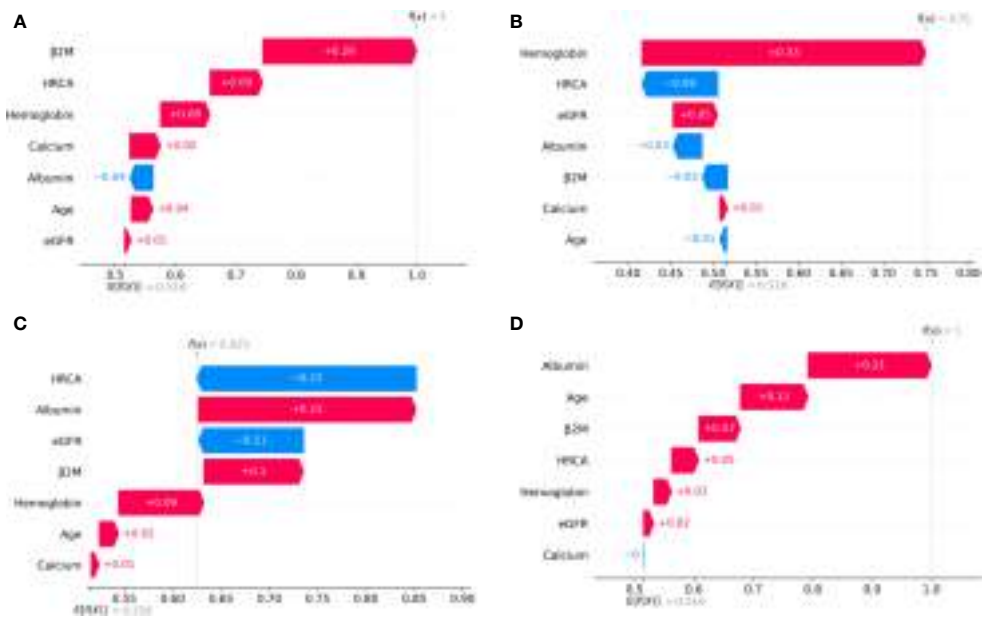


FIGURE 5 | SHAP waterfall plots for the randomly chosen four patients in the intermediate-risk stage (CRSS-2) from the MMIn dataset. The pink color shows the positive impact of the feature, while the blue color shows the negative impact of the feature. Features with a positive impact contributed to the class of intermediate-risk stage prediction, while features with a negative impact contributed to the class opposite to intermediate risk. β2M, hemoglobin, HRCA, and albumin have the highest overall impact on the intermediate-risk stage prediction in the MMIn dataset. However, the ranking of the features itself differs from patient to patient as can be seen in **(A–D)**. **(A)** β2M has the highest impact followed by HRCA. **(B)** Hemoglobin has the highest impact followed by HRCA. **(C)** HRCA has the highest impact followed by albumin. **(D)** Albumin has the highest impact followed by age.

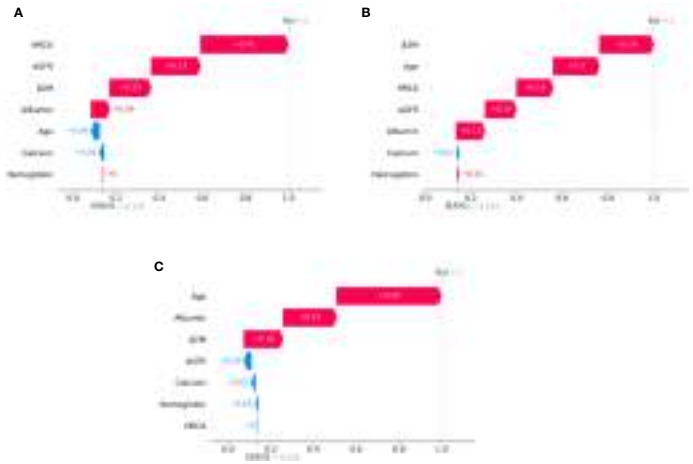


FIGURE 6 | SHAP waterfall plots for randomly chosen patients in high-risk stage (CRSS-3) from the MMIn dataset. The pink color shows the positive impact of the feature, while the blue color shows the negative impact of the feature. Features with a positive impact contributed to the class of high-risk stage prediction, while features with a negative impact contributed to class opposite to highest risk. HRCA, β2M, age, and albumin have the highest overall impact on high-risk stage prediction. However, this ranking differs from patient to patient as can be seen in **(A–C)**. **(A)** HRCA has the highest impact. **(B)** β2M has the highest impact. **(C, D)** Age and albumin have the highest impact.

the median values of the parameters age, albumin, β2M, eGFR, and hemoglobin across the three risk groups (**Figures S4, S5**) for both datasets. Furthermore, the Wilcoxon rank-sum test revealed statistically significant variations ($p < 0.05$) in the median values of the parameters between two successive risk groups (CRSS-1 and CRSS-2; CRSS-2 and CRSS-3). Furthermore, CRSS for the MMIn and MMRF datasets were interpreted using SHAP (13) to establish the clinical relevance of the risk stages predicted by the CRSS. For the MMIn data, elevated levels of β2M and calcium with lower levels of eGFR and hemoglobin contributed to high risk, whereas in the MMRF data, elevated levels of β2M and lower levels of hemoglobin, eGFR, and albumin contributed to high risk in myeloma patients. These findings are in accordance with the observations mostly identified in high-risk MM patients. Additionally, it was observed that the order of impact of hemoglobin was higher in low-risk stage prediction in the MMIn dataset as compared with the MMRF dataset, while the order of impact of hemoglobin was higher in high-risk stage prediction in the MMRF dataset as compared with the MMIn dataset (**Figures 3**, **7**). The difference in the rankings can be attributed to the varying ethnicities and further confirmed our claim of using ethnicity-aware risk-staging models for MM. In the present study, we have used the MMIn and MMRF cohorts belonging to Indian and American ethnicities, respectively, for building CRSS models. Results on both cohorts have strengthened our claim that the robustness of the staging model is amplified by inclusion of ethnicity-specific cutoffs of the prognostic factors as well as by utilizing AI techniques.

The classification rules were obtained using a decision tree classifier on the classification output of the best performing models in both MMIn and MMRF data. Overall classification accuracy was 94.79% and 98% for the MMIn and MMRF data, respectively. Final risk stages were evaluated using the classification rules in both datasets. Furthermore, it is evident from the UMAP plots that both the MMIn and MMRF data were not visible as three separate risk groups initially in the absence of CRSS risk labels (**Figures S3A, C, E**). With the addition of these risk labels with every patient sample, the subjects could be seen to be grouped separately (where a group corresponds to one risk label) in the UMAP plot (**Figures S3B, D**). This demonstrates the ability of the CRSS model in identifying the risk groups correctly from the non-separable data. To further validate our model, we found risk stages in 123 prospective subjects of MMIn data that were not used to build the CRSS model. UMAP plots (**Figure S3F**) suggest that the prospective subjects got correctly aligned to their respective risk stages inferred *via* CRSS.

For the MMIn data, β2M was in the highest level of hierarchy in the classification rules followed by hemoglobin and HRCA (**Figure S2A**). For the MMRF data, the prognostic factor in the highest level of hierarchy was β2M followed by albumin and Hb (**Figure S2B**). The cutoff values for β2M, albumin, and Hb were 5.2, 3.55, and 9.64. The cutoffs for β2M and albumin were not changed, but the cutoff value proposed for Hb was 9.59, which was close to the observed value in the classification rules. This observation further justified our choice of using new cutoffs for the risk-staging model.

Conclusion

In this work, we examined the impact of ethnicity-based cutoffs of laboratory parameters derived using the ML algorithm on risk prediction in Indian and American patients with MM. We trained different risk-staging models for both the MMRF and MMIn datasets. The best predictor model was obtained when ethnicity-specific cutoffs of the clinical parameters were utilized. Furthermore, we presented a new reliable and robust AI-enabled risk-staging system, namely, CRSS, which utilizes easily acquirable

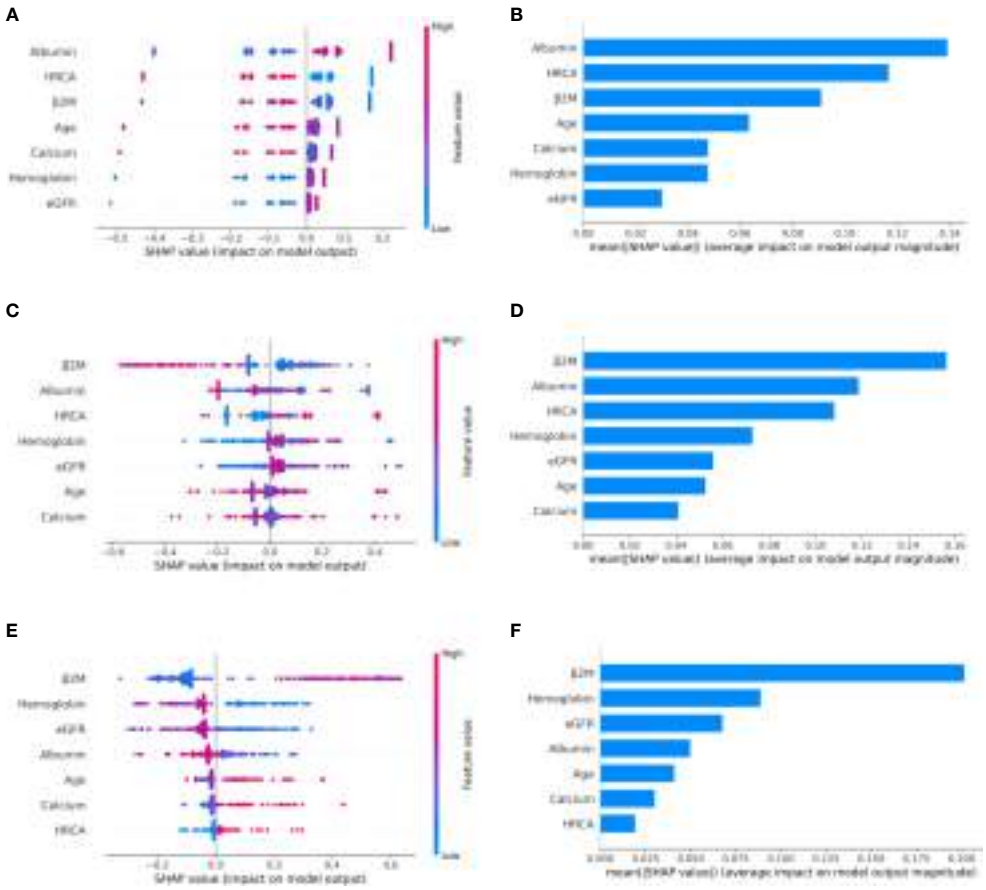


FIGURE 7 | Model interpretation using SHAP. SHAP summary plots for different risk stages inferred in MMRF data showing the impact of different parameters used in the model. **(A, B)** CRSS-1: albumin, HRCA, and β2M have the highest impact on the low-risk stage. Normal levels of albumin, absence of HRCA, and lower values of β2M are contributing to low risk (CRSS-1) in myeloma patients. **(C, D)** CRSS-2: β2M, albumin, and HRCA are the key contributors to the intermediate-risk stage. **(E, F)** CRSS-3: β2M and hemoglobin have the highest impact on the high-risk stage. Elevated levels of β2M and lower values of hemoglobin are contributing toward the high-risk stage in the patient. Lower values of albumin and eGFR are further promoting high-risk stage prediction.

laboratory and clinical parameters, i.e., age, albumin, β2M, calcium, eGFR, and hemoglobin along with HRCA (**Table S5**). Risk stratification achieved by AI-assisted CRSS is able to better separate the patients into different risk groups as compared with the R-ISS. High concordance-index and hazard ratios reveal the superior performance of the CRSS as compared with the R-ISS.

Furthermore, the clinical and biological significance of the decision tree classifier rules for risk stage prediction in MM patients was deduced *via* SHAP analysis on both datasets. The successful evaluation of our proposed staging system on both datasets establishes the utility of the proposed ethnicity-aware staging system for NDMM patients, treated largely with novel agents or a combination thereof, in a real-world scenario. Our study also highlights the importance of application of AI in building CRSS, thereby enhancing the prediction of survival outcome and separability of risk stages in NDMM patients. We have also developed a web platform-based AI-assisted ethnicity-aware MM risk-staging calculator.

Limitations and Future Work

The CRSS has been built on a smaller set of NDMM patients as compared with the R-ISS (3) study. In the future, the CRSS model may be tested on larger datasets with varying ethnic groups as the cohort size of the present study is 25% of the cohort used in the R-ISS reported in 2015. As the CRSS calculator becomes available online, data could be generated by independent groups for further validation in real-world scenarios.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding authors. CRSS calculator can be found at: http://sbilab.iiitd.edu.in/pub_files/CRRScalculator_edit.html.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by IEC, AIIMS. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AF: methodology, software, formal analysis, investigation, validation, and writing—original draft preparation. AG: methodology, investigation, validation, writing—original draft preparation, resources, project management, and supervision. KS: formal analysis, validation, and supervision. LK: resources. AS: resources. RG: conceptualization, investigation, validation, resources, writing—original draft preparation, project management, and supervision. All the authors had full access to the final version of the report. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the Department of Biotechnology, Govt. of India (Grant: BT/MED/30/SP11006/2015), and the Department of Science and Technology, Govt. of India (Grant: DST/ICPS/CPS-Individual/2018/279(G)). The funding bodies had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

ACKNOWLEDGMENTS

AF would like to thank the University Grants Commission, Govt. of India, for the UGC-Senior Research Fellowship. The authors acknowledge the MMRF and dbGaP (Project #18964) for providing the dataset. These data were generated as part of the Multiple Myeloma Research Foundation Personalized Medicine Initiative. The authors would also like to thank the Centre of Excellence in Healthcare, IIIT-Delhi for the support in their research.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2021.720932/full#supplementary-material

REFERENCES

1. Durie BG, Salmon SE. A Clinical Staging System for Multiple Myeloma Correlation of Measured Myeloma Cell Mass With Presenting Clinical Features, Response to Treatment, and Survival. *Cancer* (1975) 36(3):842–54. doi: 10.1002/1097-0142(197509)36:3<842::AID-CNCR2820360303>3.0.CO;2-U
2. Greipp PR, SanMiguel J, Durie BG, Crowley JJ, Barlogie B, Boccadoro M, et al. International Staging System for Multiple Myeloma. *J Clin Oncol* (2005) 23:3412–20. doi: 10.1200/JCO.2005.04.242
3. Palumbo A, Avet-Loiseau H, Oliva S, Lokhorst HM, Goldschmidt H, Rosinol L, et al. Revised International Staging System for Multiple Myeloma: A Report From International Myeloma Working Group. *J Clin Oncol* (2015) 33 (26):2863. doi: 10.1200/JCO.2015.61.2267
4. Rago A, Grammatico S, Za T, Levi A, Mecarocci S, Siniscalchi A, et al. Prognostic Factors Associated With Progression of Smoldering Multiple Myeloma to Symptomatic Form. *Cancer* (2012) 118(22):5544–9. doi: 10.1002/cncr.27657
5. Schinke M, Ihorst G, Duyster J, Wäsch R, Schumacher M, Engelhardt M. Risk of Disease Recurrence and Survival in Patients With Multiple Myeloma: A German Study Group Analysis Using a Conditional Survival Approach With Long-Term Follow-Up of 815 Patients. *Cancer* (2020) 126(15):3504–15. doi: 10.1002/cncr.32978
6. Howlader N, Noone AM, Krapcho M, Miller D, Bishop K, Kosary CL, et al. *SEER Cancer Statistics Review, 1975–2014*. (2017), based on November 2016 SEER data submission.
7. Ailawadhi S, Aldoss IT, Yang D, Razavi P, Cozen W, Sher T, et al. Outcome Disparities in Multiple Myeloma: A SEER-Based Comparative Analysis of Ethnic Subgroups. *Br J Haematol* (2012) 158(1):91–8. doi: 10.1111/j.1365-2141.2012.09124.x
8. Waxman AJ, Mink PJ, Devesa SS, Anderson WF, Weiss BM, Kristinsson SY, et al. Racial Disparities in Incidence and Outcome in Multiple Myeloma: A Population-Based Study. *Blood* (2010) 116(25):5501–6. doi: 10.1182/blood-2010-07-298760
9. Costa LJ, Brill IK, Omel J, Godby K, Kumar SK, Brown EE. Recent Trends in Multiple Myeloma Incidence and Survival by Age, Race, and Ethnicity in the United States. *Blood Adv* (2017) 1(4):282–7. doi: 10.1182/bloodadvances.2016002493
10. Derman BA, Jasielec J, Langerman SS, Zhang W, Jakubowiak AJ, Chiu BC. Racial Differences in Treatment and Outcomes in Multiple Myeloma: A Multiple Myeloma Research Foundation Analysis. *Blood Cancer J* (2020) 10 (8):1–7. doi: 10.1038/s41408-020-00347-6
11. Yellapragada SV, Fillmore NR, Frolov A, Zhou Y, Dev P, Yameen H, et al. Vitamin D Deficiency Predicts for Poor Overall Survival in White But Not African American Patients With Multiple Myeloma. *Blood Adv* (2020) 4 (8):1643. doi: 10.1182/bloodadvances.2019001411
12. Alexander DD, Mink PJ, Adami HO, Cole P, Mandel JS, Oken MM, et al. Multiple Myeloma: A Review of the Epidemiologic Literature. *Int J Cancer* (2007) 120(S12):40–61. doi: 10.1002/ijc.22718
13. Lundberg SM, Lee S-I. A Unified Approach to Interpreting Model Predictions. In: *Proceedings of the 31st International Conference on Neural Information Processing Systems (Nips'17)*. Red Hook, NY, USA: Curran Associates Inc. (2017). p. 4768–77.
14. Farswan A, Gupta A. TV-DCT: Method to Impute Gene Expression Data Using DCT Based Sparsity and Total Variation Denoising, in: *InICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019 May 12. pp. 1244–8, IEEE.
15. Gehlot S, Farswan A, Gupta A, Gupta R. CT-NNBI: Method to Impute Gene Expression Data Using DCT Based Sparsity and Nuclear Norm Constraint With Split Bregman Iteration, in: *In2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, 2019 Apr 8. pp. 1315–8, IEEE.
16. Farswan A, Gupta A, Gupta R, Kaur G. Imputation of Gene Expression Data in Blood Cancer and Its Significance in Inferring Biological Pathways. *Front Oncol* (2020) 9:1442. doi: 10.3389/fonc.2019.01442
17. Montealegre JR, Zhou R, Amirian ES, Scheurer ME. Uncovering Nativity Disparities in Cancer Patterns: Multiple Imputation Strategy to Handle Missing Nativity Data in the Surveillance, Epidemiology, and End Results Data File. *Cancer* (2014) 120(8):1203–11. doi: 10.1002/cncr.28533
18. Gupta R, Kaur G, Kumar L, Rani L, Mathur N, Sharma A, et al. Nucleic Acid Based Risk Assessment and Staging for Clinical Practice in Multiple Myeloma. *Ann Hematol* (2018) 97(12):2447–54. doi: 10.1007/s00277-018-3457-8

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) 17:e328–46. doi: 10.1016/S1470-2045(16)30206-6
20. Florkowski CM, Chew-Harris JS. Methods of Estimating GFR–different Equations Including CKD-EPI. *Clin Biochem Rev* (2011) 32(2):75.
21. Rajkumar SV. Multiple Myeloma: 2016 Update on Diagnosis, Risk-Stratification, and Management. *Am J Hematol* (2016) 91(7):719–34. doi: 10.1002/ajh.24402
22. Eo SH, Kang HJ, Hong SM, Cho H. K-Adaptive Partitioning for Survival Data, With an Application to Cancer Staging. (2013). arXiv preprint arXiv:1306.4615.
23. Monti S, Tamayo P, Mesirov J, Golub T. Consensus Clustering: A Resampling-Based Method for Class Discovery and Visualization of Gene Expression Microarray Data. *Mach Learn* (2003) 52(1-2):91–118. doi: 10.1023/A:1023949509487
24. Kaur G, Gupta R, Mathur N, Rani L, Kumar L, Sharma A, et al. Clinical Impact of Chromothriptic Complex Chromosomal Rearrangements in Newly Diagnosed Multiple Myeloma. *Leukemia Res* (2019) 76:58–64. doi: 10.1016/j.leukres.2018.12.005
25. Avet-Loiseau H, Leleu X, Roussel M, Moreau P, Guerin-Charbonnel C, Caillot D, et al. Bortezomib Plus Dexamethasone Induction Improves Outcome of Patients With T (4; 14) Myeloma But Not Outcome of Patients With Del (17p). *J Clin Oncol* (2010) 28(30):4630–4. doi: 10.1200/JCO.2010.28.3945
26. Farswan A, Gupta A, Gupta R, Hazra S, Khan S, Kumar L, et al. AI-Supported Modified Risk Staging for Multiple Myeloma Cancer Useful in Real-World Scenario. *Trans Oncol* (2021) 14(9):101157. doi: 10.1016/j.tranon.2021.101157
27. Unnikrishnan A, Khan AM, Narayan P, Norkin M. Striking Age Differences of Multiple Myeloma (MM) Diagnosis in Patients of Indian and Pakistani Descent in the United States Compared to Native Countries. *J Clin Oncol* (2017) 35:e13070. doi: 10.1200/JCO.2017.35.15_suppl.e13070
28. Konatam AM, Sadashivudu G. Age of Onset of Multiple Myeloma: A Paradigm Shift in Indian Patients. *Indian J Appl Res* (2016) 6:3.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Graded Depth of Response and Neoplastic Plasma Cell Index as Indicators of Survival Outcomes in Patients With Multiple Myeloma Following Autologous Stem Cell Transplant

Nupur Das, MD,[1] Meetu Dahiya, MSc,[1] Ritu Gupta, MD,[1] Lalit Kumar, DM,[2] Lata Rani, PhD,[1] Anubha Gupta, PhD,[3] Akanksha Farswan, PhD,[3] Atul Sharma, DM,[2] and OmDutt Sharma, MSc[1]

From the [1]Laboratory Oncology Unit, Dr. B.R.A. IRCH, All India Institute of Medical Sciences (AIIMS), New Delhi, India; [2]Department of Medical Oncology, Dr. B.R.A. IRCH, AIIMS, New Delhi, India; and [3]SBI Lab, Department of ECE, Indraprastha Institute of Information Technology, New Delhi, India.

ABSTRACT

Objectives: With a substantial number of patients with multiple myeloma (MM) experiencing disease relapse, the quest for more sensitive methods to assess deeper responses indicative of cure continues.

Methods: In this prospective analysis of 170 patients with MM at day 100 after autologous stem cell transplant, we evaluated the predictive value of conventional response, measurable residual disease (MRD_{TOTAL}: the aberrant percentage of plasma cells [PC%] among total bone marrow cells), and neoplastic plasma cell index scores (NPCI: the aberrant PC% of total PCs).

Results: Significantly better progression-free survival (PFS) and overall survival (OS) were observed with deepening conventional response. Conventional response–based stratification within the MRD-positive and MRD-negative subgroups showed a significantly higher PFS (hazard ratio [HR], 3.11; $P < .005$) and OS (HR, 3.08; $P = .01$) in the conventional response-positive/MRD-positive group compared with the conventional response-negative/MRD-positive group. Using K-adaptive partitioning to find the optimum threshold for MRD, patients achieving less than 0.001% MRD_{TOTAL} had superior PFS (MRD_{TOTAL} 0.001% to <0.1%: HR, 6.66, $P < .005$; MRD_{TOTAL} ≥0.1%: HR, 11.52, $P < .005$) and OS (MRD_{TOTAL} 0.001% to <0.1%: HR, 5.3, $P < .05$; MRD_{TOTAL} ≥0.1%: HR = 9.21, $P < .005$). The C index and Akaike information criterion metrics demonstrated the superior performance of the NPCI compared with MRD_{TOTAL} in predicting treatment outcome.

Conclusions: Progressive deepening of response, conventional as well as MRD, correlates with superior survival outcomes. The NPCI proved to be a superior determinant of survival and can be explored as a better statistic than MRD.

KEY POINTS

- Increasing depth of conventional response (from less than very good partial response to stringent conventional response) and a gradual fall in measurable residual disease (MRD) predicts superior survival.
- Patients achieving maximal conventional response have survival benefits in both the MRD-positive and MRD-negative subgroups, highlighting the essence of conventional response in the era of MRD in response assessment.
- Survival analysis and comparative performance metrics show that the neoplastic plasma cell index is a better predictor of disease progression and mortality than MRD.

KEY WORDS

Multiparametric flow cytometry; Measurable residual disease; Autologous stem cell transplant; K-adaptive partitioning; Neoplastic plasma cell index

Am J Clin Pathol XXXX 2022;XX:1-0
HTTPS://DOI.ORG/10.1093/AJCP/AQAC129

Received: June 2, 2022
Accepted: September 13, 2022

Corresponding author: Ritu Gupta, MD; drritugupta@gmail.com, drritu.laboncology@aiims.edu.

Funding: This work was supported by a grant from the Department of Biotechnology, Government of India (Grant: BT/MED/30/SP11006/2015).

INTRODUCTION

The current standard of care in multiple myeloma (MM) includes induction therapy with a combination of proteasome inhibitor/immunomodulatory drug/monoclonal antibody conjugate followed by autologous stem cell transplant (ASCT) in transplant-eligible patients.[1-3] Achievement of complete response (CR) is associated with improved survival in both transplant-eligible and non–transplant-eligible patients, as various studies have shown.[4,5] A significant proportion of patients achieving CR eventually progress, however, calling into question the utility of CR as a predictor of prognostic relevance. Disease progression after achievement of CR highlights the persistence of residual tumor cells that are not detectable by conventional techniques. Measurable residual disease (MRD) assessment by multiparametric flow cytometry (MFC) and next-generation sequencing (NGS) or quantitative polymerase chain reaction (qPCR) molecular methods has emerged as an independent prognostic indicator and thus has been incorporated into the International Myeloma Working Group (IMWG) as a response criterion.[6,7]

The MFC scores better than the molecular methods in clinical laboratories primarily because of the fast turnaround time of MFC-MRD, the availability of flow cytometers in clinical laboratories, and the relatively resource-effective nature of the MFC assay. The process of MRD assessment has been standardized over the years, and currently the EuroFlow bulk lysis protocol is recommended for sample preparation; a 2-tube, 8-color panel and a single-tube, 10-color panel are preferred across Europe and the United States, respectively.[8,9] The use of the next-generation flow cytometry has resulted in accurate and reproducible MRD results, with a sensitivity of 10^{-6} for MRD detection.[8] The inherent challenge associated with MFC-MRD assays, however, is the use of bone marrow aspirate (BMA), which is susceptible to inadequate representation of the myeloma cells because of either procedural hemodilution or patchy involvement, both of which result in underreporting of plasma cell (PC) fractions and false-negative results. To offset the challenge of hemodilution in MRD assessments, several solutions have been put into use: (1) first-pull BMA, (2) assessing the adequacy of hematopoietic elements, (3) analyzing a large number of events, and (4) using denominators other than total bone marrow cells.[10] The European Myeloma Network has thus recommended the use of first-pull BMA for MRD assessment in MM.[11] Determination of sample adequacy based on the presence of bone marrow–specific cell populations, such as erythroid precursors, PCs (normal or abnormal), B-cell progenitors (hematogones), and mast cells, have also been found to be associated with better PC yield in adequate specimens.[12] In this context, we introduced the concept of the neoplastic plasma cell index (NPCI)—that is, the percentage of clonal aberrant PCs in the total PCs in all the samples for MRD evaluation—as a parameter to document MRD.

In this study, we evaluated the role of depth of conventional response and MRD assessment in predicting survival outcomes in patients with MM following ASCT. We used K-adaptive partitioning (KAP) to find a clinically optimum threshold for MRD levels to assess their role in response evaluation. We further assessed the clinical relevance of the NPCI compared with total bone marrow MRD levels in patients with MM who had received an ASCT.

MATERIAL AND METHODS

Participants

We conducted a prospective analysis of patients with MM who underwent ASCT at our center from 2014 until December 2019 and for whom the MFC-based MRD assessments were done at day 100 following ASCT. The demographic, clinical, and laboratory data were retrieved from patient medical records. We conducted a response assessment of the patients per the IMWG response criteria.[7] All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Declaration of Helsinki and its later amendments. All patients provided their informed consent after proper counseling before participation in the study.

MRD Evaluation by Flow Cytometry

We assessed MRD based on BMAs collected at day 100 following ASCT. We processed BMAs collected in EDTA per the consensus recommendations for MM by the EuroFlow Consortium.[13,14] Two-tube, 8-color antibody combination (CD138, CD56, CD38, CD45, CD117, CD81, CD19, CD27, CD138, CD56, CD38, CD45, Cy-κ, Cy-λ, CD19, or CD27) or a single-tube, 10-color combination (CD138, CD56, CD19, CD38, CD45, CD117, Cy-κ, Cy-λ, CD81, CD27) were used to assess the PCs.[15] The samples were acquired using a flow cytometer (Gallios; Beckman Coulter), and the data were analyzed using Kaluza, version 2.0, analysis software (Beckman Coulter). Plasma cells were gated based on the expression pattern of CD38, CD138, and CD45, as described previously.[15,16] An aberrant PC cluster was defined by presence of more than 30 PCs with an aberrant immunophenotype and monoclonal for involved light chain. Samples with less than 3×10^6 events with an absence of neoplastic PCs may be false-negative for MRD and thus were excluded. We conducted the MRD assessment using 2 different analytic approaches: (1) MRD_{TOTAL}, defined as the percentage of aberrant neoplastic PCs (APCs) in the total nucleated bone marrow cells (MRD_{TOTAL} = APC/total nucleated bone marrow cells), or (2) NPCI, calculated as the percentage of APCs in the total number of PCs (NPCI = APC/total PC).

Clinical End Points and Statistical Analysis

Progression free survival (PFS) was measured from the date of ASCT to the date of progression or relapse. Overall survival (OS) was calculated from the date of ASCT to the date of death or last follow-up visit. Patients who had not progressed or relapsed were censored on the last date they were known to be alive. Survival analysis was done using Kaplan-Meier survival curves, with the log-rank test for comparison between the groups; the hazard ratio (HR) was calculated using the Cox proportional hazards model. Statistical analyses were performed using SigmaPlot, version 13.0, software (Systat Software).

K-Adaptive Partitioning

The KAP algorithm is generally used to divide the data into K heterogeneous subgroups based on the information from a prognostic factor.[17] We used KAP to segregate patients with varied levels of MRD_{TOTAL} and NPCI based on PFS.

The C Index and the Akaike Information Criterion

We used both the C index and the Akaike information criterion (AIC) to compare MRD_{TOTAL} and NPCI as predictors of survival. The C index is an evaluation metric used to compare the goodness of fit of risk prediction models.[18] It quantifies the proportion of the concordant pairs divided by all the possible pairs of participants. In survival analysis, a pair of patients (i, j) is called concordant if patient "i," who experienced an event, had a higher risk score than patient "j," who either experienced the event at a later stage (survival time is higher for patient "j" compared with patient "i") or has not yet experienced the event at all. Its value ranges from 0 to 1, where a value below 0.5 represents a very poor model and values near to 1 represent a good model.

The AIC is a mathematical model used to compare multiple models built from the same data set.[19] Hence, the AIC assists in the selection of the best-fit model for the data. Note that AIC scores across different models are meaningful only if these models are built on the same data set. The model with a lower AIC score best fits the data.

RESULTS

Patient Characteristics

Our sample consisted of 194 patients with MM who had undergone ASCT between January 2014 and December 2019 and for whom medical records were available. The total number of events acquired on MFC were less than 3 million in 24 patients, so those patients were excluded from subsequent analysis. The clinical and laboratory characteristics of the patient cohort are provided in TABLE 1.

The median time from diagnosis to ASCT was 12 months (range, 4-48 months). The median PFS after ASCT was 36 months (range, 5-88 months), and OS was 41.5 months (range, 5-93 months). In all, 65 (38.2%) patients progressed and 38 (22.4%) died before the analysis date.

None of the baseline clinical and laboratory parameters were predictive of PFS and OS following ASCT TABLE 1. Patients achieving stringent CR (sCR) before transplant had superior PFS and OS compared with patients achieving very good partial response (VGPR) or less (PFS: HR, 3.29 [95% confidence interval (CI), 1.6-6.78], $P < .005$; OS: HR, 2.9 [95% CI, 1.11-7.59], $P = .03$) FIGURES 1 and 2. Depth of response after initial induction therapy was not associated with PFS and OS in this cohort, however TABLE 1.

Conventional Response Following ASCT

Forty-three patients were in sCR before transplant. Of the remaining 127 patients, 72 (57%) improved their response at day 100 following ASCT. Compared with the patients who had an improved response, the risk of progression and mortality was higher in patients with no improvement in response at day 100 after ASCT (PFS: HR, 2.61 [95% CI, 1.52-4.48], $P < .005$; OS: HR, 2.46 [95% CI, 1.22-4.97], $P = .01$) FIGURES 1 and 2.

At day 100 after ASCT, sCR, CR, and VGPR and less were documented in 95 (55.9%), 38 (22.4%), and 37 (21.7%) patients, respectively. Compared with the patients with sCR, patients with a response less than sCR had progressively higher risk of progression (CR: HR, 4.31 [95% CI, 2.33-8.00], $P < .005$; ≤VGPR: HR, 6.01 [95% CI, 3.17-11.4], $P < .005$) FIGURE 1 as well as mortality (CR: HR, 3.01 [95% CI, 1.32-6.84], $P = .01$; ≤VGPR: HR, 4.86 [95% CI, 2.21-10.69], $P < .005$) FIGURE 2.

Effect of MRD_{TOTAL} at Day 100 After ASCT on Survival

Of 170 patients, 50% (n = 85) were MRD positive. MRD-positive patients had a higher risk of progression (HR, 7.81 [95% CI, 3.85-15.84]; $P < .005$) FIGURE 3A and mortality (HR, 6.4 [95% CI, 2.49-16.49]; $P < .005$) FIGURE 3B compared with the MRD-negative patients. Patients were then segregated into 3 groups based on MRD_{TOTAL} levels as MRD_{TOTAL}–1 (<0.001%), MRD_{TOTAL}–2 (0.001% to <0.1%), and MRD_{TOTAL}–3 (≥0.1%). Patients achieving MRD_{TOTAL}–1 levels had superior PFS (MRD_{TOTAL}–2: HR, 6.66 [95% CI, 3.2-13.88], $P < .005$; MRD_{TOTAL}–3: HR, 11.52 [95% CI, 5.19-25.57], $P < .005$) FIGURE 3C and OS (MRD_{TOTAL}–2: HR, 5.3 [95% CI, 1.97-14.24], $P < .05$; MRD_{TOTAL}–3: HR, 9.21 [95% CI, 3.27-25.93], $P < .005$) FIGURE 3D compared with the other 2 groups, suggesting that the risk of progression and mortality increases with an increase in MRD levels.

Effect of the NPCI on Survival

The NPCI determination is based on the percentage of APCs within the total number of PCs and is likely to offset the influence of hemodilution, which is a major concern for falsely low MRD levels. Therefore, the patients were segregated into 3 groups based on NPCI cutoffs derived by KAP. For the NPCI, the 3 groups were NPCI_1 ($0 \le NPCI \le 0.98$; n = 86), NPCI_2 ($0.98 < NPCI \le 23.71$; n = 51), and NPCI_3 ($NPCI > 23.71$; n = 33). Risk of progression and mortality significantly increased with an increase in the value of the NPCI. Compared with NPCI_1, risk of progression and mortality was significantly higher in NPCI_2 (PFS: HR, 5 [95% CI, 2.33-10.71], $P < .005$; OS: HR, 3.65 [95% CI, 1.28-10.41], $P = .02$) FIGURE 3E and NPCI_3 (PFS: HR, 17.49 [95% CI, 8.23-37.162], $P < .005$; OS: HR, 12.4 [95% CI, 4.64-33.11], $P < .005$) FIGURE 3F.

Effect of Conventional Response and MFC-MRD–Based Stratification on Survival

Based on conventional response and MRD_{TOTAL} status, the patients were distributed into four groups: Conv–MRD– (n = 65), Conv–MRD+ (n = 30), Conv+MRD– (n = 20), and Conv+MRD+ (n = 55). Conv– is equivalent to sCR, whereas Conv+ is equivalent to CR and less. As compared with the Conv–MRD– subgroup, the risk of progression was significantly higher in the rest of the 3 subgroups (Conv–MRD+: HR, 6.36 [95% CI, 2.09-19.37], $P < .005$; Conv+MRD–: HR, 4.06 [95% CI, 1.09-15.14], $P = .04$; Conv+MRD+: HR, 21.36 [95% CI, 7.7-60.76], $P < .005$) FIGURE 4A. Compared with the Conv–MRD– subgroup, the risk of mortality was significantly higher in the Conv+MRD+ subgroup (HR, 8.48 [95% CI, 2.96-24.32]; $P < .005$) FIGURE 4B. Within the MRD-positive subgroup,

TABLE 1 Baseline Demographic, Clinical, and Laboratory Characteristics of Patients With Multiple Myeloma Evaluated for Measurable Residual Disease

Parameter	No. of Patients (n = 170)	PFS, mo (*P* Value)	OS, mo (*P* Value)
Age, y			
Median (range)	52 (28-72)		
<65/≥65	158/12	59/NR (.291)	NR (.202)
Sex			
M/F	115/55	59/NR (.240)	NR/82 (.656)
Hemoglobin, g/dL			
<10/≥10	91/79	59/65 (.233)	NR/NR (.207)
Albumin, g/L			
<3.5/≥3.5/NA	63/105/02	79/59 (.256)	NR/NR (.354)
eGFR, mL/min/1.73m^2			
≤48.2/>48.2	46/124	59/65 (.958)	NR/NR (.553)
Serum β2 microglobulin, μg/mL			
<5.5/≥5.5/NA	99/64/07	59/NR (.923)	NR/NR (.735)
M band subtype			
IIMM/LCMM/mixed	37/43/90	67/65/52 (.494)	37/43/90 (.339)
ISS staging			
ISS1/2/3/NA	39/59/65/7	65/59/NR (.589)	77/NR/NR (.086)
R-ISS staging			
R-ISS1/2/3/NA	15/87/16/52	NR/59/5) (.665)	NR (.293)
MRS staging[20]			
MRS1/2/3/NA	55/80/25/10	53/67/NR (.855)	82/NR/NR (.647)
CRSS staging[21]			
CRSS1/2/3/NA	45/46/01/78	NR/59/NR (.534)	82/NR/NR (.03)
Induction/first-line therapy			
Doublet (DT/RD/VD) Triplet (VCD/VRD/VTD)	50 (11/25/14) 120 (46/55/19)	59 69 (.928)	NR NR (.207)
Molecular abnormalities on FISH (n = 92)			
del13q/del17p/t(4;14)/t(11;14)/t(14;16)/Negative	14/5/8/6/NA/59	–	–
Poor risk/not poor risk	13/79	59/NR (.182)	NR/NR (.501)
Lines of therapies before transplant			
01/02	129/41	79/43 (.02)	NR/NR (.435)
Response after first-line therapy			
sCR/CR/≤VGPR	33/43/94	59/52/65 (.633)	NR/82/NR (.942)
Pre-ASCT response			
sCR/CR/≤VGPR	43/41/86	NR/53/44 (.002)	NR (.059)

ASCT, autologous stem cell transplant; CR, complete response; CRSS, Consensus-Based Risk Staging System[23]; DT, dexamethasone-thalidomide; eGFR, estimated glomerular filtration rate; FISH, fluorescence in situ hybridization; IIMM, intact immunoglobulin multiple myeloma; ISS, International Staging System; LCMM, light-chain-only multiple myeloma; MRS, modified risk staging[22]; NA, not applicable; NR, not reached; OS, overall survival; PFS, progression-free survival; R-ISS, Revised International Staging System; RD, lenalidomide-dexamethasone; sCR, stringent conventional response; VCD, bortezomib-cyclophosphamide-dexamethasone; VD, bortezomib-dexamethasone; VGPR, very good partial response; VRD, bortezomib-lenalidomide-dexamethasone; VTD, bortezomib-thalidomide-dexamethasone.

the conventional response further stratified patients with different survival outcomes. In the MRD-positive subgroup, Conv+MRD+ patients had a significantly higher risk of progression (HR, 3.11 [95% CI, 1.63-5.92]; $P < .005$) FIGURE 4C and mortality (HR, 3.08 [95% CI, 1.3-7.28]; $P = .01$) FIGURE 4D compared with Conv−MRD+ patients. In the MRD-negative subgroup, Conv+MRD− patients had a significantly higher risk of progression (HR, 4.06 [95% CI, 1.09-15.14]; $P = .04$) FIGURE 4E compared with Conv−MRD− patients; however, no significant difference was found in terms of mortality (HR, 0.77 [95% CI, 0.09-6.86]; $P = .81$) FIGURE 4F.

Effect of Conventional Response and NPCI-Based Stratification on Survival

Based on conventional response and NPCI status, patients were distributed into 6 groups: Conv−NPCI_1 (n = 65), Conv+NPCI_1 (n = 20), Conv−NPCI_2 (n = 21), Conv+NPCI_2 (n = 29), Conv−NPCI_3 (n = 9), and Conv+NPCI_1 (n = 26). In the NPCI_1 subgroup, conventional response-positive patients had a significantly higher risk of progression (HR, 4.007 [95% CI, 0.86-18.59]; $P = .024$) FIGURE 5A but no significant impact on mortality (HR, 1.357 [95% CI, 0.83-10.06]; $P = .76$) FIGURE 5B compared with conventional response-negative

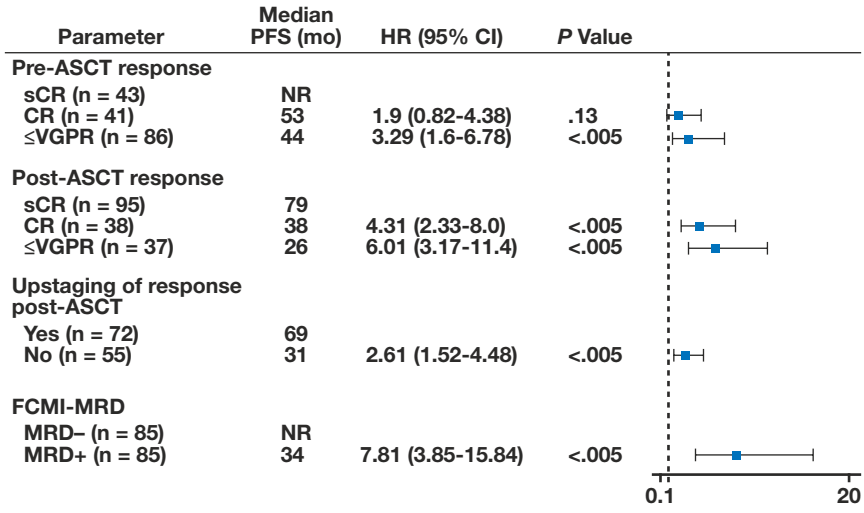


FIGURE 1 Forest plot showing the influence of various parameters on progression-free survival (PFS) in post–autologous stem cell transplant (ASCT) recipients with multiple myeloma (n = 170). CI, confidence interval; CR, complete response; FCMI, flow cytometric immunophenotyping; HR, hazard ratio; MRD, measurable residual disease; NR, not reached; sCR, stringent conventional response; VGPR, very good partial response.

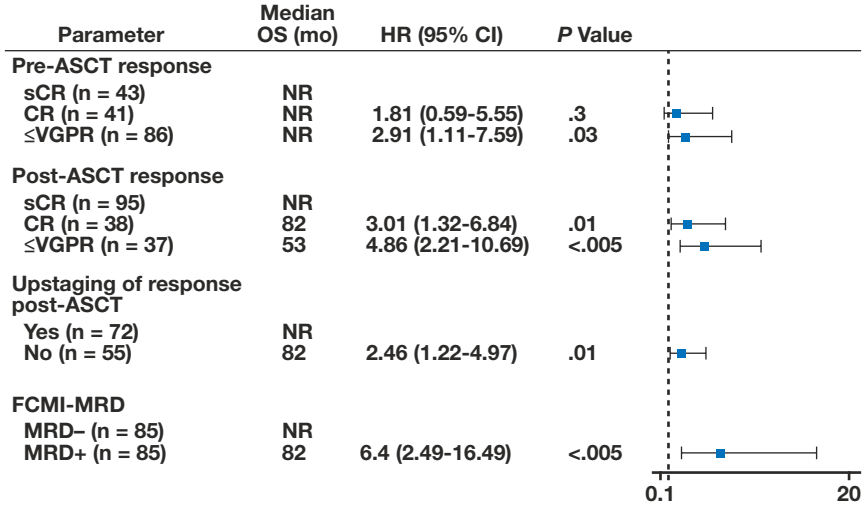


FIGURE 2 Forest plot showing the influence of various parameters on overall survival (OS) in post–autologous stem cell transplant (ASCT) recipients with multiple myeloma (n = 170). CI, confidence interval; CR, complete response; FCMI, flow cytometric immunophenotyping; HR, hazard ratio; MRD, measurable residual disease; NR: not reached; sCR, stringent conventional response; VGPR, very good partial response.

patients. In the NPCI_2 subgroup, conventional response-positive patients had a significantly higher risk of progression and mortality (PFS: HR, 2.27 [95% CI, 1.01-5.05], $P = .026$; OS: HR, 3.34 [95% CI, 1.01-10.98], $P = .047$) FIGURE 5C and 5D compared with Conv–NPCI_2 patients. In the NPCI_3 subgroup, however, conventional response could not stratify the patients in terms of progression (HR, 2.14 [95% CI, 1.01-4.50]; $P = .06$) FIGURE 5E or mortality (HR, 1.63 [95% CI, 0.63-4.19]; $P = .36$) FIGURE 5F.

Comparison of NPCI With MRD_{TOTAL} Levels for Prediction of Survival Outcomes

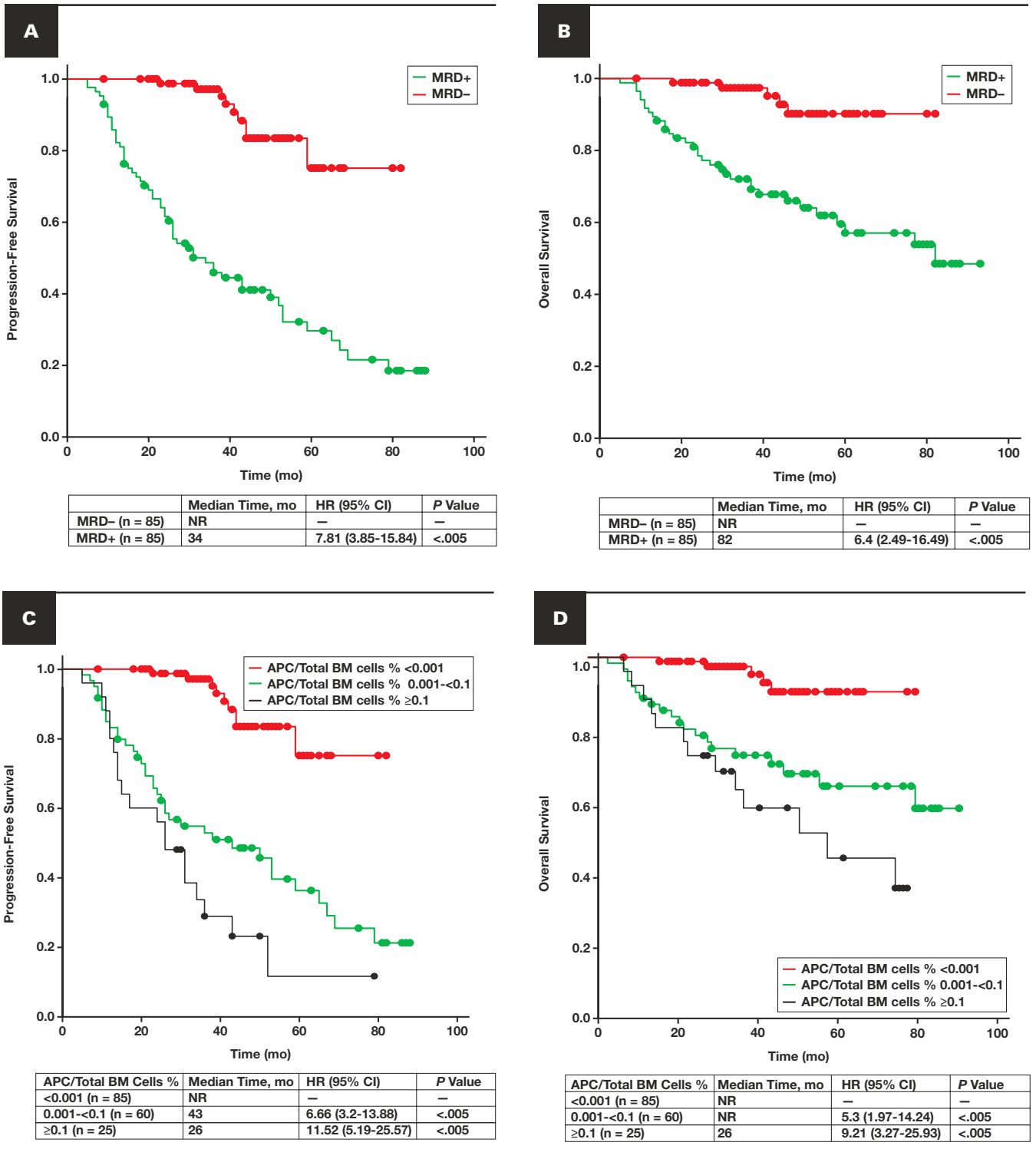
The C index and AIC metrics demonstrated the slightly superior performance of the NPCI compared with MRD_{TOTAL} levels in predicting PFS (NPCI: C index, 0.8; AIC, 520.86; MRD_{TOTAL}: C index, 0.74; AIC, 542.45) and OS (NPCI: C index, 0.77; AIC, 321.52; MRD_{TOTAL}: C index, 0.70; AIC, 334.23).

DISCUSSION

Since the commencement of stem cell transplantation, ASCT remains the standard therapy of choice for treating young as well as fit elderly patients (>65 years) with MM.[24,25] As the depth of response directly correlates with superior PFS and OS in MM, the recent focus is on assessment of residual disease by sensitive methods such as MFC, qPCR, and NGS. Although the increasing depth of response following ASCT has been shown to translate into superior PFS, the same is not true for OS for patients with MM treated before the era of novel agents.[26-29]

In a large retrospective study (n = 865), achievement of CR after ASCT was shown to be associated with significantly prolonged PFS but not OS, which was, however, significantly superior in the subgroup of patients who received novel agent–based induction therapy.[30] Recently, Nishimura et al[31] reported that patients with

HTTPS://DOI.ORG/10.1093/AJCP/AQAC129



	Median Time, mo	HR (95% CI)	*P* Value
MRD– (n = 85)	NR	—	—
MRD+ (n = 85)	34	7.81 (3.85-15.84)	<.005

	Median Time, mo	HR (95% CI)	*P* Value
MRD– (n = 85)	NR	—	—
MRD+ (n = 85)	82	6.4 (2.49-16.49)	<.005

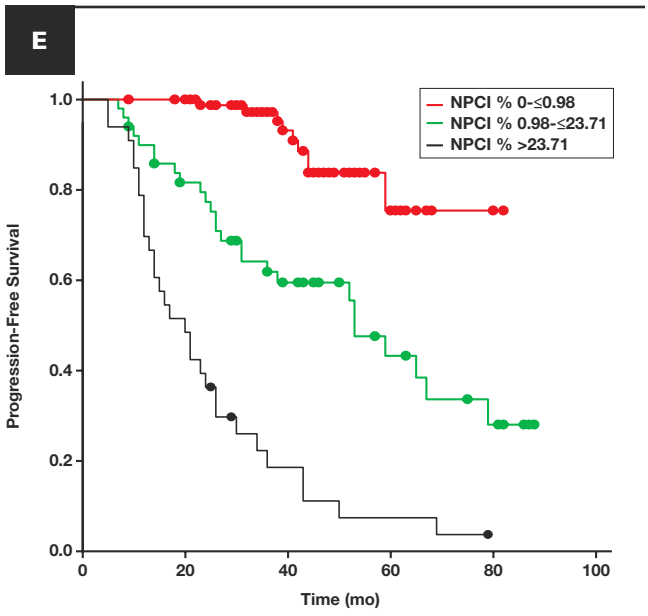
APC/Total BM Cells %	Median Time, mo	HR (95% CI)	*P* Value
<0.001 (n = 85)	NR	—	—
0.001-<0.1 (n = 60)	43	6.66 (3.2-13.88)	<.005
≥0.1 (n = 25)	26	11.52 (5.19-25.57)	<.005

APC/Total BM Cells %	Median Time, mo	HR (95% CI)	*P* Value
<0.001 (n = 85)	NR	—	—
0.001-<0.1 (n = 60)	NR	5.3 (1.97-14.24)	<.005
≥0.1 (n = 25)	26	9.21 (3.27-25.93)	<.005

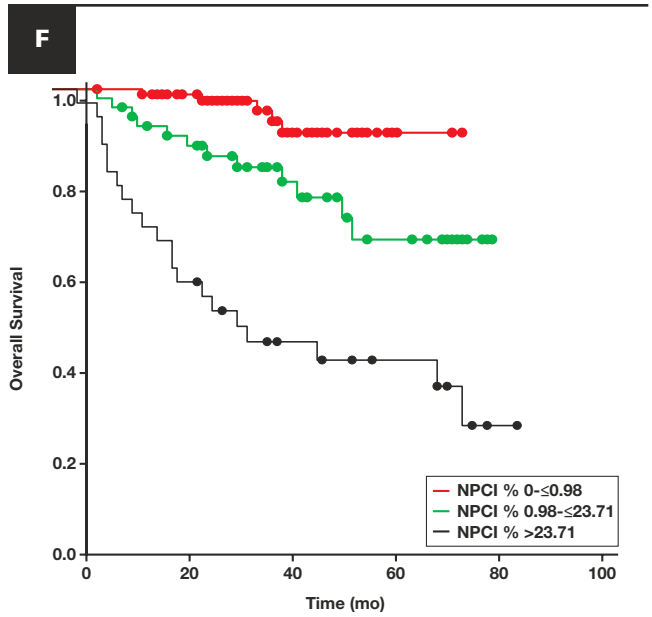
FIGURE 3 The Kaplan-Meier curves for progression-free survival and overall survival in post–autologous stem cell transplant recipients with multiple myeloma (n = 170) based on measurable residual disease (MRD) status (**A, B**); percentage of aberrant neoplastic plasma cell (APC) per total bone marrow (BM) cells (**C, D**); neoplastic plasma cell index (NPCI) (**E, F**). CI, confidence interval; HR, hazard ratio; NR, not reached.

MM treated in 2014 or later had superior OS compared with patients treated in 1997 or earlier and intermediate survival in patients treated during intervening time periods. These studies indicate that the CR attained by conventional therapy is of insufficient depth to positively influence long-term survival, and the addition of novel agents to the treatment regimen enhances the depth of response to a level where patients achieve OS benefits.

The current study includes patients who received at least 1 of the novel agents before ASCT. It is evident that as response deepens from VGPR and less to attainment of sCR, a dramatic improvement



NPCI %	Median Time, mo	HR (95% CI)	*P* Value
0-≤0.98 (n = 87)	NR	—	—
>0.98-≤23.71 (n = 50)	53	5 (2.33-10.71)	<.005
>23.71 (n = 33)	20	17.49 (8.23-37.16)	<.005



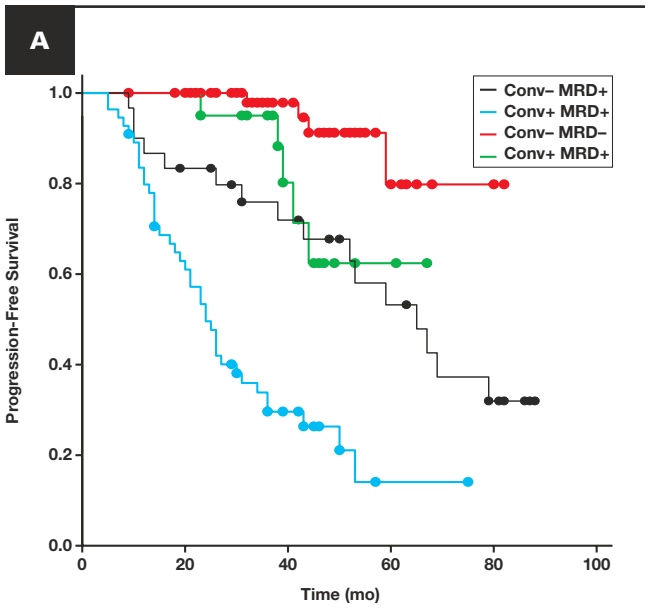
NPCI %	Median Time, mo	HR (95% CI)	*P* Value
0-≤0.98 (n = 87)	NR	—	—
>0.98-≤ 23.71 (n = 50)	NR	3.65 (1.28-10.41)	.02
>23.71 (n = 33)	37	9.21 (4.64-33.11)	<.005

FIGURE 3 (cont)

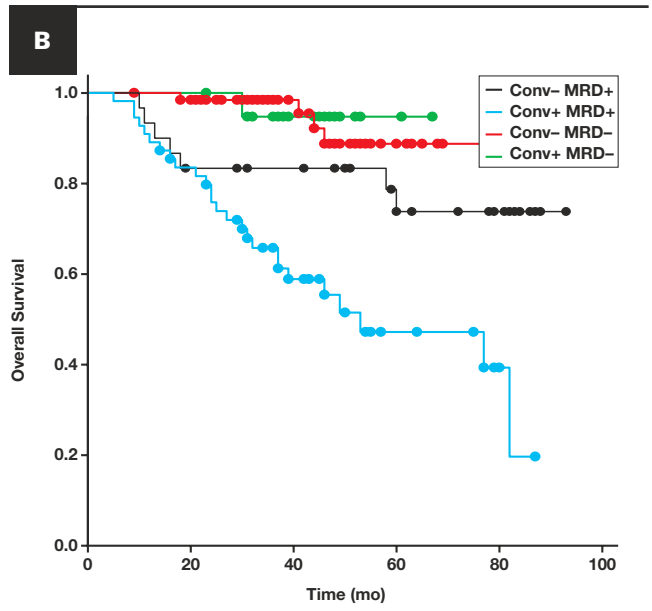
in long-term survival is observed, signifying that the increasing depth of response following ASCT is predictive of superior survival outcomes. This study further demonstrates an upgradation in depth of response after ASCT in 57% of the patients, which translates into superior PFS and OS in this subset of patients. The results of our study emphasize that ASCT plays a crucial role in bringing a significant reduction in disease level that, if attained, facilitates achievement of superior PFS and OS, as reported previously.[32,33]

The presence of high-risk molecular aberrations or any other clinical or laboratory parameter assessed at the diagnostic time point did not influence long-term outcomes in the transplant recipients evaluated in the current study (TABLE 1). A pooled analysis of patients with MM enrolled in the GIMEMA-MM-03-05 and GEM05MAS65 studies also showed that the depth of response is a stronger predictor of survival than cytogenetic profile and disease stage in patients with MM.[34] In this study, Gay et al[34] showed that although the 3-year OS rate was significantly different for risk categories based on cytogenetics and disease stage, the rates did not differ in patients achieving CR. These findings indicate that the category of conventional response achieved is a better predictor of outcome compared with any of the individual host- or disease-specific parameters, including high-risk cytogenetics. Further, because ASCT as a treatment strategy improves the depth of response, it is an apt treatment for patients with high-risk cytogenetics. Another point of clinical relevance in this context is the concept of dynamic risk monitoring, which is nascent in MM response evaluation but worthwhile to explore. Recently, Fan et al[35] evaluated the prognostic impact of dynamic changes in cytogenetic abnormalities in a cohort of 80 paired patients at diagnosis and recurrence. In this study, it was demonstrated that the patients showing new high-risk cytogenetic abnormalities and escalated genetic risk stratification at relapse had poor outcomes compared with those with standard risk. Thus, it may be relevant to re-stratify all the posttransplant patients at the time of progression based on molecular and cytogenetic abnormalities to see their impact on overall clinical outcome.

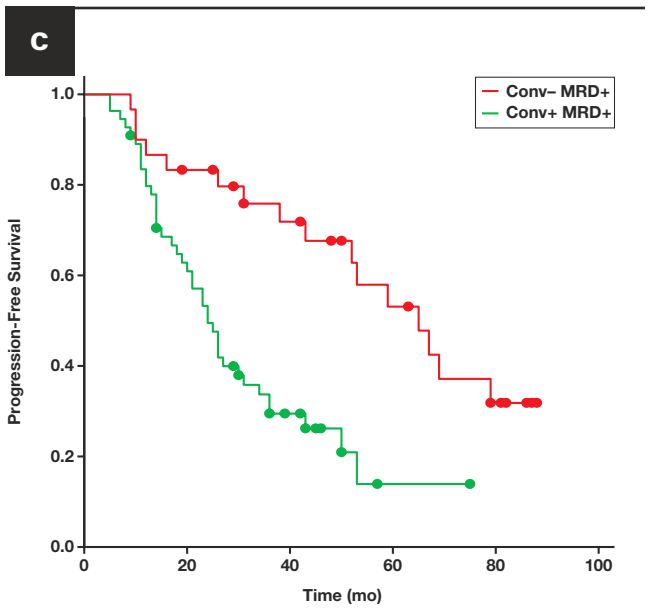
MRD level after therapy has emerged as one of the strongest prognostic indicators in patients with MM with persistent MRD negativity as a predictor of superior PFS and OS.[36,37] In our cohort of patients with MM, MRD-positive status was associated with poor outcome compared with MRD-negative status, with progressive worsening of both PFS and OS with an increase in MRD levels from less than 0.001 to higher levels. Because previous studies have shown that MRD surpasses CR and is a better indicator of treatment outcomes, we sought to evaluate whether conventional response currently has any role. We observed that within the MRD-positive group, conventional response further stratified patients with poor outcome (ie, patients with detectable disease on protein electrophoresis or deranged serum free light-chain ratio [sFLCR] had poor survival outcomes compared with those who attained sCR) FIGURE 4C and 4D. Similarly, within the MRD-negative group, the PFS was superior in patients attaining sCR compared with those with CR less than sCR FIGURE 4E. From this information, it can be deduced that achievement of maximal conventional response is an important independent predictor of survival outcomes in MM. This finding jibes with the observation made by Lauherta et al,[38] who demonstrated that MRD-negative status surpasses CR. On further stratifying the patients based on conventional response, however, patients who achieved maximal conventional response have survival benefits compared with less than maximal response in both the MRD-positive and MRD-negative subgroups. Similarly, patients with NPCI_1 and NPCI_2 achieving maximal response using conventional technology had better survival outcomes. The NPCI_3 group



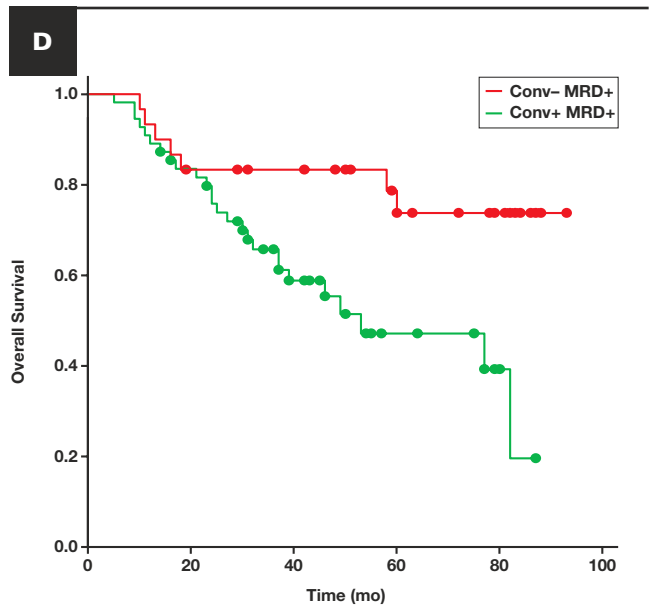
Conv + MRD Status	Median Time, mo	HR (95% CI)	*P* Value
Conv– MRD– (n = 65)	NR	—	—
Conv+ MRD– (n = 20)	NR	4.06 (1.09-15.14)	.04
Conv– MRD+ (n = 30)	65	6.36 (2.09-19.37)	<.005
Conv+ MRD– (n = 55)	24	21.36 (7.7-60.76)	<.005



Conv + MRD Status	Median Time, mo	HR (95% CI)	*P* Value
Conv– MRD– (n = 65)	NR	—	—
Conv+ MRD– (n = 20)	NR	0.77 (0.09-6.86)	.81
Conv– MRD+ (n = 30)	NR	2.67 (0.77-9.33)	.12
Conv+ MRD+ (n = 55)	53	8.48 (2.96-24.32)	<.005



Conv + MRD Status	Median Time, mo	HR (95% CI)	*P* Value
Conv– MRD+ (n = 30)	65	—	—
Conv+ MRD+ (n = 55)	24	3.11 (1.63-5.92)	<.005

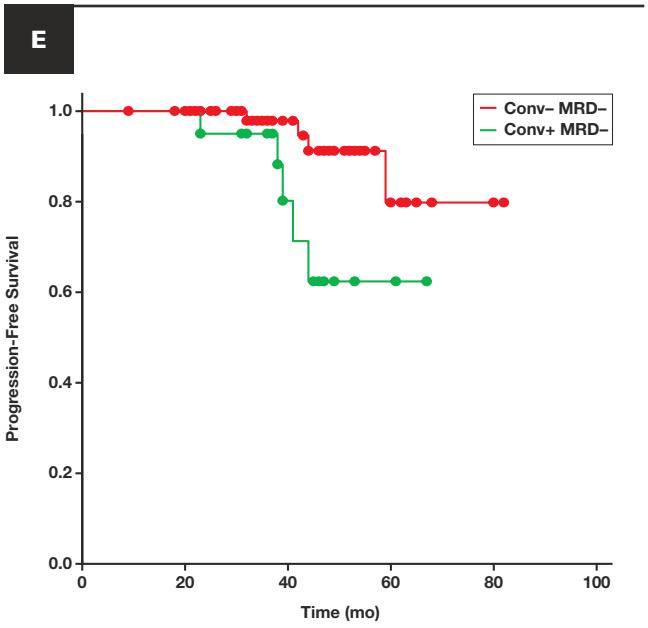


Conv + MRD Status	Median Time, mo	HR (95% CI)	*P* Value
Conv– MRD+ (n = 30)	NR	—	—
Conv+ MRD+ (n = 55)	53	3.08 (1.3-7.28)	.01

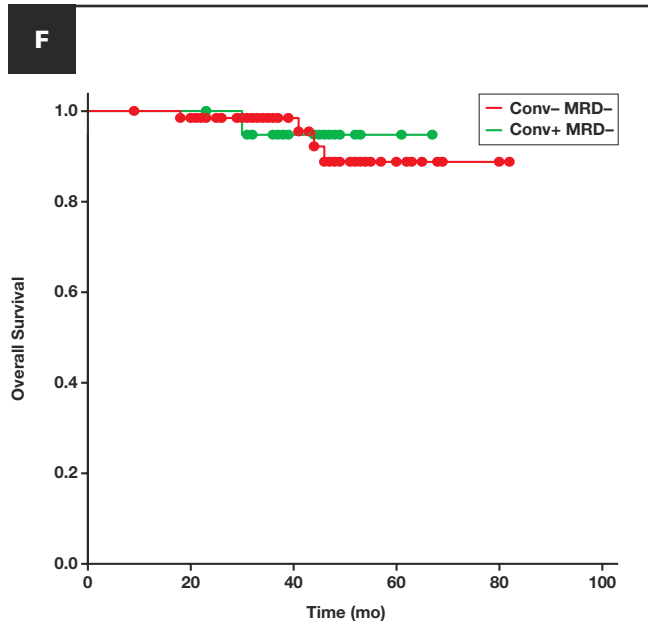
FIGURE 4 The Kaplan-Meier curves for progression-free survival and overall survival in post–autologous stem cell transplant recipients with multiple myeloma (n = 170) based on conventional response and measurable residual disease (MRD) status in all subgroups (**A, B**), stratified based on conventional response (Conv) within the MRD-positive (**C, D**) and MRD-negative (**E, F**) groups. CI, confidence interval; HR, hazard ratio; NR, not reached.

had high tumor load; hence, conventional response was not useful in further risk stratification of this group of patients. Because we also considered normalization of sFLCR and undetectable disease on protein electrophoresis in conventional response assessment, this may likely explain the superior survival of MRD-positive patients in the sCR group.

As the conventional response undoubtedly helps stratify patients with distinct survival outcomes within the MRD-positive and MRD-negative groups, it is important to give due diligence to monitoring of M protein, which is a true biomarker and helps assess the secretory component of the disease, whether intra- or extramedullary. The MRD assay, in contrast, helps assess only the



Conv + MRD Status	Median Time, mo	HR (95% CI)	*P* Value
Conv– MRD– (n = 65)	NR	—	—
Conv+ MRD– (n = 20)	NR	4.06 (1.09-15.14)	.04



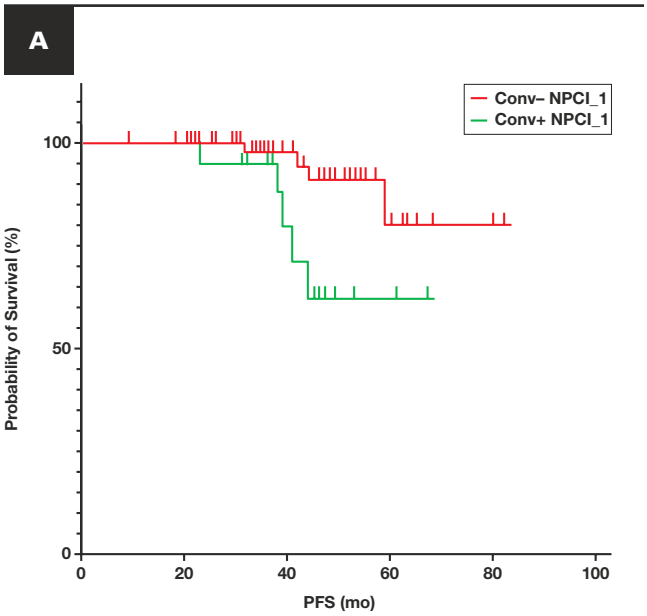
Conv + MRD Status	Median Time, mo	HR (95% CI)	*P* Value
Conv– MRD– (n = 65)	NR	—	—
Conv+ MRD– (n = 20)	NR	0.77 (0.09-6.86)	.81

FIGURE 4 (cont)

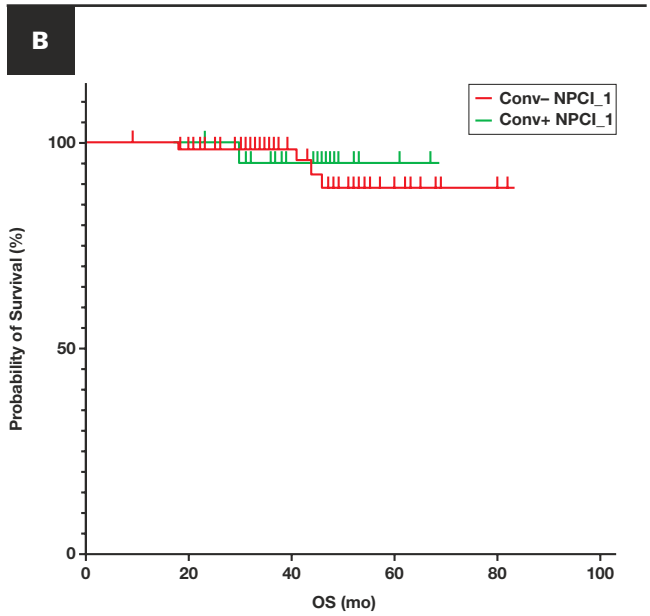
intramedullary component of the disease and is limited by patient discomfort associated with serial sampling of the bone marrow, overall expense, and the likelihood of missing the residual cells because of the patchy nature of the disease and hemodilution. Thus, although blood-based disease monitoring using M protein estimation plays a pivotal role in response assessment, the current methods of detecting monoclonal protein (ie, protein electrophoresis and sFLC) are not sensitive enough to evaluate the deep responses observed with the institution of novel agents in MM. To take advantage of the benefits of M protein detection in predicting long-term treatment outcome, an ultrasensitive method for monitoring M protein in blood is the prerequisite. Mass spectrometry (MS)–based technologies have recently demonstrated significant potential in this area and are emerging as a potential strategy to monitor residual disease in the near future.[39-41] Two new methods for MS-based M protein detection include the heavy-chain variable-region clonotypic approach and the light-chain MS-based approach.[42,43] Recently, Mills et al[44] demonstrated the role of ultrasensitive M protein detection in patients with MM achieving sCR and observed that although monoclonal immunoglobulin rapid accurate mass measurement (miRAMM) positivity at single time points after ASCT did not correlate with PFS, sCR patients with persistently elevated miRAMM M protein intensities in serial measurements had shorter PFS than those with reduced miRAMM intensities (median, 17.9 months vs 51.6 months; $P <$ .0017). Based on the existing data on the role of MS in M protein detection in MM, the IMWG Mass Spectrometry Committee has acknowledged lower rates of CR in clinical trials by MS-based M protein estimation than protein-immunofixation studies and recommended the M protein estimation by MS as a useful alternative to immunofixation in clinical practice.[20] Hence, a blood-based ultrasensitive M protein assay can serve as an essential tool that extends the current value of monitoring M proteins from response assessment as a predictor of survival outcomes.

Hemodilution is one of the most challenging factors influencing MRD assessments not only in MM but across the entire spectrum of hematolymphoid malignancies. As both neoplastic and normal PCs are proportionately reduced in a hemodiluted BMA,[21] the NPCI was calculated in this study as a percentage of neoplastic PCs within the total number of PCs in contrast to MRD_{TOTAL}, which is a percentage of neoplastic or aberrant PCs within the total number of bone marrow cells. Both NPCI and MRD_{TOTAL} detection performed equally well, with superior survival in the NPCI/MRD_{TOTAL}–low subgroups (NPCI_1 and MRD_{TOTAL}–1) compared with the NPCI/MRD_{TOTAL}–high subgroups (NPCI_2; NPCI_3 and MRD_{TOTAL}–2; MRD_{TOTAL}–3) FIGURE 3C-3F. The NPCI performed better, however, than MRD_{TOTAL} on comparative performance metrics of C index and HR, with higher C index and HR values observed for NPCI both for PFS and OS. Thus, it may be worthwhile to explore the relevance of the NPCI as an indicator of residual disease monitoring in patients with MM.

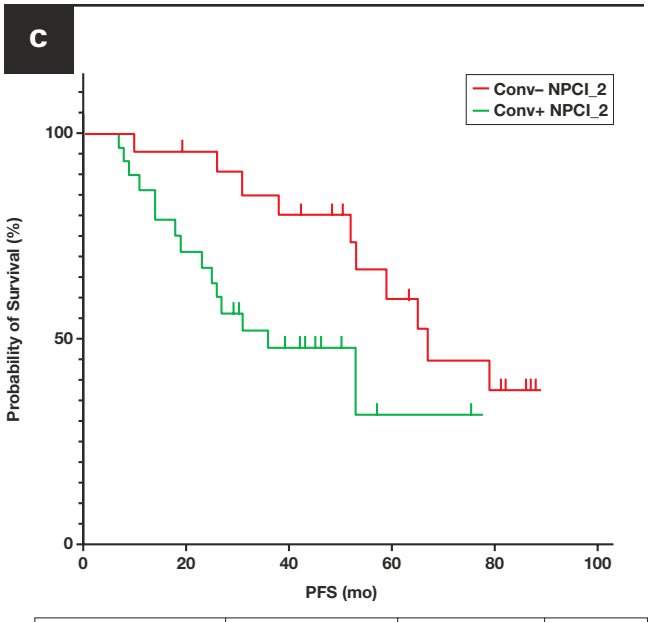
We have demonstrated that a progressive increase in the depth of response, conventional as well as MFC-MRD, correlates with graded survival outcomes, with the best outcomes in patients achieving sCR and with MRD-negative status. Furthermore, the NPCI has proved to be a better indicator of survival outcomes compared with MRD_{TOTAL} and thus should be evaluated in larger cohorts for validation. This study shows that conventional response helps stratify patients within the MRD-positive and MRD-negative subgroups. Thus, more sensitive methods to detect the monoclonal protein may play a major role in residual disease monitoring in secretory MM in the near future.



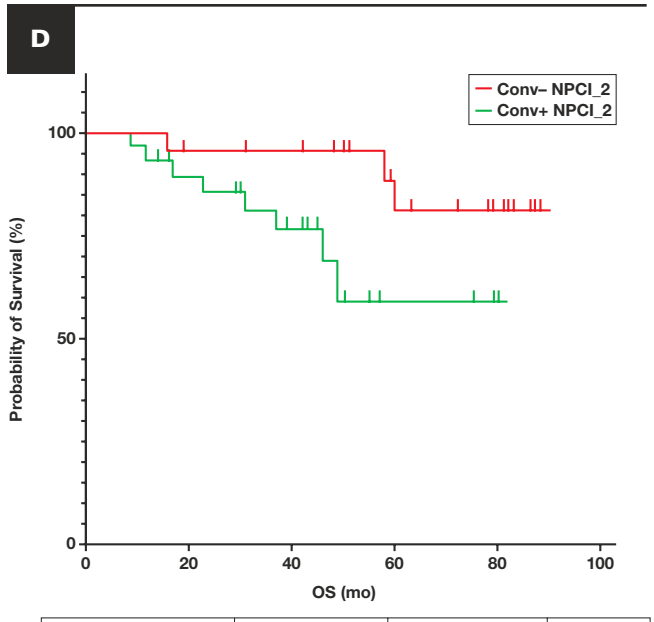
Conv + NPCI_1 Status	Median Time, mo	HR (95% CI)	*P* Value
Conv– NPCI_1 (n = 65)	NR	—	—
Conv+ NPCI_1 (n = 20)	NR	4.007 (0.86-18.59)	.024



Conv + NPCI_1 Status	Median Time, mo	HR (95% CI)	*P* Value
Conv– NPCI_1 (n = 65)	NR	—	—
Conv+ NPCI_1 (n = 20)	NR	1.357 (0.83-10.06)	.764

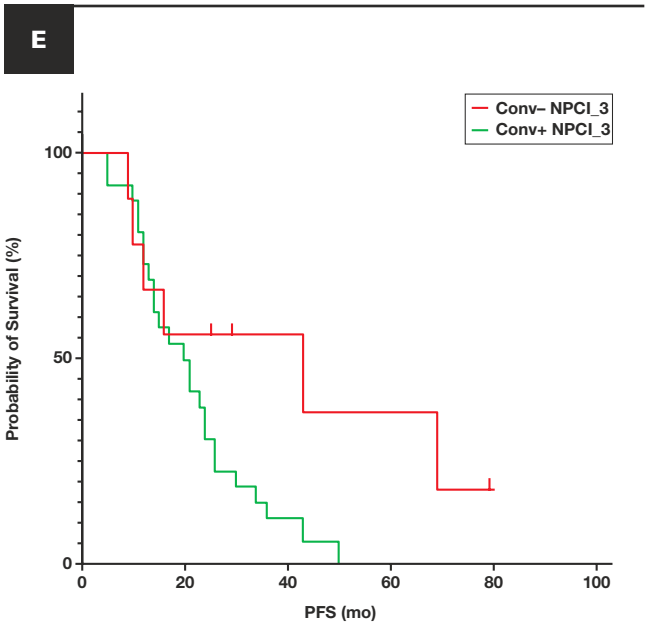


Conv + NPCI_2 Status	Median Time, mo	HR (95% CI)	*P* Value
Conv– NPCI_2 (n = 21)	67	—	—
Conv+ NPCI_2 (n = 29)	36	2.270 (1.01-5.05)	.026

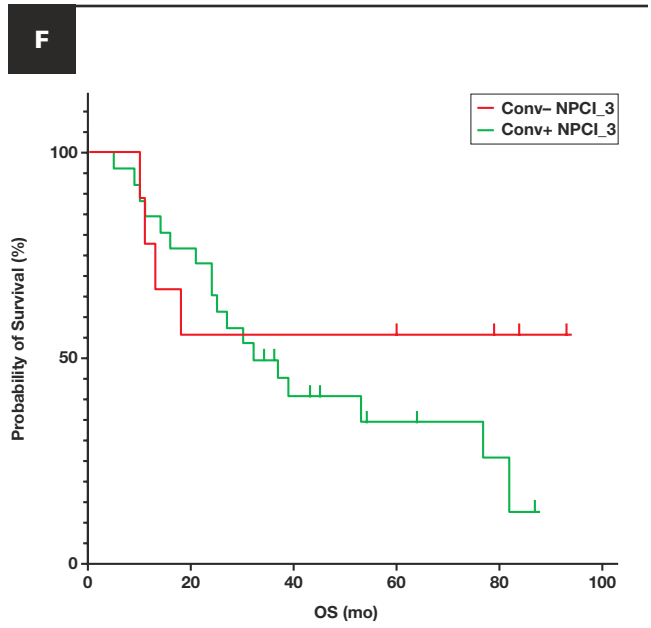


Conv + NPCI_2 Status	Median Time, mo	HR (95% CI)	*P* Value
Conv– NPCI_2 (n = 21)	NR	—	—
Conv+ NPCI_2 (n = 29)	NR	3.341 (1.01-10.98)	.047

FIGURE 5 The Kaplan-Meier curves for progression-free survival (PFS) and overall survival (OS) in post–autologous stem cell transplant recipients with multiple myeloma (n = 170) based on conventional response (Conv) within neoplastic plasma cell index (NPCI)_1 (**A, B**), NPCI_2 (**C, D**), and NPCI_3 (**E, F**) groups. CI, confidence interval; HR, hazard ratio; NR, not reached.



Conv + NPCI_3 Status	Median Time, mo	HR (95% CI)	*P* Value
Conv– NPCI_3 (n = 9)	43	—	—
Conv+ NPCI_3 (n = 26)	21	2.140 (1.018-4.502)	.006



Conv + NPCI_3 Status	Median Time, mo	HR (95% CI)	*P* Value
Conv– NPCI_3 (n = 9)	NR	—	—
Conv+ NPCI_3 (n = 26)	35	1.631 (0.63-4.198)	.36

FIGURE 5 (cont)

REFERENCES

1. Moreau P, Attal M, Facon T. Frontline therapy of multiple myeloma. *Blood*. 2015;125:3076-3084. doi:10.1182/blood-2014-09-568915.
2. Kumar SK, Rajkumar SV, Dispenzieri A, et al. Improved survival in multiple myeloma and the impact of novel therapies. *Blood*. 2008;111:2516-2520. doi:10.1182/blood-2007-10-116129.
3. Richardson PG, Weller E, Lonial SN, et al. Lenalidomide, bortezomib, and dexamethasone combination therapy in patients with newly diagnosed multiple myeloma. *Blood*. 2010;116:679-686. doi:10.1182/blood-2010-02-268862.
4. Martinez-Lopez J, Blade J, Mateos MV, et al; Grupo Español de MM; Programa para el Estudio de la Terapé Utica en Hemopatía Maligna. Long-term prognostic significance of response in multiple myeloma after stem cell transplantation. *Blood*. 2011;118:529-534. doi:10.1182/blood-2011-01-332320.
5. Lahuerta JJ, Mateos MV, Martínez-López J, et al. Influence of pre- and post-transplantation responses on outcome of patients with multiple myeloma: sequential improvement of response and achievement of complete response are associated with longer survival. *J Clin Oncol*. 2008;26:5775-5782. doi:10.1200/JCO.2008.17.9721.
6. 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-e346. doi:10.1016/S1470-2045(16)30206-6.
7. Dispenzieri A, Kyle R, Merlini G, et al; International Myeloma Working Group. International Myeloma Working Group guidelines for serum-free light chain analysis in multiple myeloma and related disorders. *Leukemia*. 2009;23:215-224. doi:10.1038/leu.2008.307.
8. 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. doi:10.1038/leu.2017.29.
9. Roshal M, Flores-Montero JA, Gao Q, et al. MRD detection in multiple myeloma: comparison between MSKCC 10-color single-tube and EuroFlow 8-color 2-tube methods. *Blood Adv*. 2017;1:728-732. doi:10.1182/bloodadvances.2016003715.
10. Jain G, Das N, Gajendra S et al. Effect of the sequence of pull of bone marrow aspirates on plasma cell quantification in plasma cell proliferative disorders. *Int J Lab Hematol*. 2022;44:837-845. doi:10.1111/ijlh.13887
11. Rawstron AC, Orfao A, Beksac M, et al; European Myeloma Network. Report of the European Myeloma Network on multiparametric flow cytometry in multiple myeloma and related disorders. *Haematologica*. 2008;93:431-438. doi:10.3324/haematol.11080.
12. Kalina T, Flores-Montero J, van der Velden VH, et al; EuroFlow Consortium (EU-FP6, LSHB-CT-2006-018708). EuroFlow standardization of flow cytometer instrument settings and immunophenotyping protocols. *Leukemia*. 2012;26:1986-2010. doi:10.1038/leu.2012.122.
13. Van Dongen JJM, Lhermitte L, Böttcher S, et al; EuroFlow Consortium (EU-FP6, LSHB-CT-2006-018708). EuroFlow antibody panels for standardized n-dimensional flow cytometric immunophenotyping of normal, reactive and malignant leukocytes. *Leukemia*. 2012;26:1908-1975. doi:10.1038/leu.2012.120.
14. 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. doi:10.1200/JCO.2012.46.2119.
15. Das N, Dahiya M, Gupta R, et al. Flow cytometric immunophenotyping of plasma cells across the spectrum of plasma cell proliferative disorders: a fresh insight with pattern-based recognition. *Cytometry B Clin Cytom*. 2022;102:292-302. doi:10.1002/cyto.b.22062.
16. Das N, Dahiya M, Gupta R, et al. Relevance of polyclonal plasma cells and post-therapy immunomodulation in MRD assessment in multiple myeloma. *Cytometry B Clin Cytom*. 2022;102:209-219. doi:10.1002/cyto.b.22068.

17. Eo SH, Kang HJ, Hong SM, et al. Adaptive partitioning for survival data, with an application to cancer staging. *arXiv*. https://arxiv.org/pdf/1306.4615.pdf. Published November 4, 2014. Accessed September 22, 2022.
18. Longato E, Vettoretti M, Di Camillo B. A practical perspective on the concordance index for the evaluation and selection of prognostic time-to-event models. *J Biomed Inform*. 2020;108:103496. doi:10.1016/j.jbi.2020.103496.
19. Cavanaugh JE, Neath AA. The Akaike information criterion: background, derivation, properties, application, interpretation, and refinements. *WIRES Comput Stat*. 2019;11:e1460. doi:10.1002/wics.1460.
20. Murray DL, Puig N, Kristinsson S, et al. Mass spectrometry for the evaluation of monoclonal proteins in multiple myeloma and related disorders: an International Myeloma Working Group Mass Spectrometry Committee Report. *Blood Cancer J*. 2021;11:24. doi:10.1038/s41408-021-00408-4.
21. Gupta R, Bhaskar A, Kumar L, et al. Flow cytometric immunophenotyping and minimal residual disease analysis in multiple myeloma. *Am J Clin Pathol*. 2009;132:728-732. doi:10.1309/AJCP1GYI7EHQYUYK.
22. Farswan A, Gupta A, Gupta R, et al. AI-supported modified risk staging for multiple myeloma cancer useful in real-world scenario. *Transl Oncol*. 2021;14:101157. doi:10.1016/j.tranon.2021.101157.
23. Farswan A, Gupta A, Sriram K, et al. Does ethnicity matter in multiple myeloma risk prediction in the era of genomics and novel agents? Evidence from real-world data. *Front Oncol*. 2021;11:720932. doi:10.3389/fonc.2021.720932.
24. Barlogie B, Hall R, Zander A, et al. High-dose melphalan with autologous bone marrow transplantation for multiple myeloma. *Blood*. 1986;67:1298-1301.
25. Harousseau JL, Moreau P. Autologous hematopoietic stem-cell transplantation for multiple myeloma. *N Engl J Med*. 2009;360:2645-2654. doi:10.1056/NEJMct0805626.
26. Koreth J, Cutler CS, Djulbegovic B, et al. High-dose therapy with single autologous transplantation versus chemotherapy for newly diagnosed multiple myeloma: a systematic review and meta-analysis of randomized controlled trials. *Biol Blood Marrow Transplant*. 2007;13:183-196. doi:10.1016/j.bbmt.2006.09.010.
27. Lenhoff S, Hjorth M, Turesson I, et al; Nordic Myeloma Study Group. Intensive therapy for multiple myeloma in patients younger than 60 years: long-term results focusing on the effect of the degree of response on survival and relapse patterns after transplantation. *Haematologica*. 2006;91:1228-1233.
28. Rajkumar SV, Fonseca R, Dispenzieri A, et al. Effect of complete response on outcome following autologous stem cell transplantation for myeloma. *Bone Marrow Transplant*. 2000;26:979-983. doi:10.1038/sj.bmt.1702640.
29. Gupta R, Kumar L, Dahiya M, et al. Minimal residual disease evaluation in ASCT recipients with multiple myeloma. *Leuk Lymphoma*. 2017;58:1234-1237. doi:10.1080/10428194.2016.1228930.
30. Lehners N, Becker N, Benner A, et al. Analysis of long-term survival in multiple myeloma after first-line autologous stem cell transplantation: impact of clinical risk factors and sustained response. *Cancer Med*. 2018;7:307-316. doi:10.1002/cam4.1283.
31. Nishimura KK, Barlogie B, van Rhee F, et al. Long-term outcomes after autologous stem cell transplantation for multiple myeloma. *Blood Adv*. 2020;4:422-431. doi:10.1182/bloodadvances.2019000524.
32. Kumar L, Iqbal N, Mookerjee A, et al. Complete response after autologous stem cell transplant in multiple myeloma. *Cancer Med*. 2014;3:939-946. doi:10.1002/cam4.257.
33. van de Velde HJ, Liu X, Chen G, et al. Complete response correlates with long-term survival and progression-free survival in high-dose therapy in multiple myeloma. *Haematologica*. 2007;92:1399-1406. doi:10.3324/haematol.11534.
34. Gay F, Larocca A, Mateos M, et al. Complete response is a stronger predictor of survival than cytogenetic profile and International Staging System stage in elderly myeloma patients treated with bortezomib: analysis of 771 patients EHA [abstract]. *Haematologica*. 2012;97:242.
35. Fan H, Liu J, Yuan C, et al. The prognostic impact of dynamic changes of genetic risk stratification in multiple myeloma. *Blood*. 2020;136:1-3. doi:10.1182/blood-2020-139345.
36. 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. doi:10.1182/blood-2014-01-550020.
37. Rawstron AC, Gregory WM, de Tute RM, et al. Minimal residual disease in myeloma by flow cytometry: independent prediction of survival benefit per log reduction. *Blood*. 2015;125:1932-1935. doi:10.1182/blood-2014-07-590166.
38. 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. doi:10.1200/JCO.2016.69.2517.
39. Eveillard M, Rustad E, Roshal M, et al. Comparison of MALDI-TOF mass spectrometry analysis of peripheral blood and bone marrow-based flow cytometry for tracking measurable residual disease in patients with multiple myeloma. *Br J Haematol*. 2020;189:904-907. doi:10.1111/bjh.16443.
40. Zajec M, Langerhorst P, VanDuijn MM, et al. Mass spectrometry for identification, monitoring, and minimal residual disease detection of M-proteins. *Clin Chem*. 2020;66:421-433. doi:10.1093/clinchem/hvz041.
41. Bergen HR 3rd, Dasari S, Dispenzieri A, et al. Clonotypic light chain peptides identified for monitoring minimal residual disease in multiple myeloma without bone marrow aspiration. *Clin Chem*. 2016;62:243-251. doi:10.1373/clinchem.2015.242651.
42. Barnidge DR, Tschumper RC, Theis JD, et al. Monitoring M-proteins in patients with multiple myeloma using heavy-chain variable region clonotypic peptides and LC-MS/MS. *J Proteome Res*. 2014;13:1905-1910. doi:10.1021/pr5000544.
43. Remily-Wood ER, Benson K, Baz RC, et al. Quantification of peptides from immunoglobulin constant and variable regions by LC-MRM MS for assessment of multiple myeloma patients. *Proteomics Clin Appl*. 2014;8:783-795. doi:10.1002/prca.201300077.
44. Mills JR, Barnidge DR, Dispenzieri A, et al. High sensitivity blood-based M-protein detection in sCR patients with multiple myeloma. *Blood Cancer J*. 2017;7:e590. doi:10.1038/bcj.2017.75.