

List of ten best papers of Dr. Ritu Gupta, highlighting the important discoveries/contributions described in them briefly (*Corresponding author)

1. Farswan A, Gupta A, Jena L, Kaur G, **Gupta R***. Characterizing the mutational landscape of MM and its precursor MGUS. **Am J Cancer Res** **2022**;12(4):1919-1933. PMID: 35530275

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 from three different global regions were analyzed and single base substitutions, mutational signatures and TMB were inferred from the variants identified in MGUS and MM patients. This study finds a change in the mutational spectrum with a statistically significant increase from MGUS to MM in the frequency of all the three categories of variants, non-synonymous (NS), synonymous (SYN), and others (OTH) ($P < 0.05$). However, there was a statistically significant rise in the TMB values for TMB_NS and TMB_SYN only. We also reported 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. 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.

2. Kaur G, Jena L, **Gupta R***, Farswan A, Gupta A, Sriram K. Correlation of changes in subclonal architecture with progression in the MMRF CoMMpass study. **Transl Oncol.** **2022** Jun 28; 23:101472. doi: 10.1016/j.tranon.2022.101472. PMID: 35777247.

Multiple myeloma 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. To address this, we evaluated clonal evolution at multiple time points in MM patients. 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 synchronize molecular events of drug response, myelomatogenesis and hold potential for early prognostication and development of risk stratified personalized therapies in MM.

3. Farswan A, Jena L, Kaur G, Gupta A, **Gupta R***, Rani L, Sharma A, Kumar L. Branching clonal evolution patterns predominate mutational landscape in Multiple Myeloma. **Am J Cancer Res** 2021; 11(11):5659-5679. PMID: 34873486

Multiple Myeloma (MM) arises from malignant transformation and deregulated proliferation of clonal plasma cells (PCs) harbouring heterogeneous molecular anomalies. The effect of evolving mutations on clone fitness and their cellular prevalence shapes the progressing myeloma genome and impacts clinical outcomes. Although clonal heterogeneity in MM is well established, which subclonal mutations emerge/persist/perish with progression in MM and which of these can be targeted therapeutically remains an open question. To address this, we sequenced pairwise whole exomes of 62 MM patients collected at diagnosis and on progression and demonstrated marked intraclonal heterogeneity and identified actionable/druggable gene targets that varied at different time points within the same patient. Branching evolution was predominant pattern of clonal evolution and associated with had low TMBs (<10) and 2 or more founder clones. A distinct temporal fall in subclonal driver mutations was identified from diagnosis to progression e.g., in *PABPC1*, *BRAF*, *KRAS*, *CRI*, *DIS3* and *ATM* genes suggesting such patients could be treated early with target specific drugs like Vemurafenib/Cobimetinib. An analogous rise in driver mutations was observed in *KMT2C*, *FOXD4L1*, *SP140*, *NRAS* whereas a few drivers such as *FAT4*,

IGLL5 and *CDKN1A* retained consistent distribution patterns at two time points. These findings are clinically relevant and provide evidence for multi time point evaluation of subclonal mutational landscapes for time to time risk stratification and tailoring time-adapted combination therapies in MM.

4. Ruhela V, Jena L, Kaur G, **Gupta R***, Gupta A. BDL-SP: A Bio-inspired DL model for the identification of altered Signaling Pathways in Multiple Myeloma using WES data. *Am J Cancer Res.* 2023 Apr 15;13(4):1155-1187. PMID: 37168334.

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. 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.

5. Katiyar A, Kaur G, Rani L, Jena L, Singh H, Kumar L, Sharma A, Kaur P, **Gupta R***. Genome-wide identification of potential biomarkers in multiple myeloma using meta-analysis of mRNA and miRNA expression data. *Sci Rep.* 2021; 11(1):10957. PMID: 34040057

Differentially expressed genes (DEGs) and miRNAs (DEMs) in MM may influence disease pathogenesis, clinical presentation / drug sensitivities. But these signatures overlap meagrely plausibly due to complexity of myeloma genome, diversity in primary cells studied, molecular technologies/ analytical tools utilized. This warrants further investigations since DEGs/DEMs can impact clinical outcomes and guide personalized therapy. We have conducted genome-wide meta-analysis of DEGs/DEMs in MM versus Normal Plasma Cells (NPCs) and derived unified putative signatures for MM. 100 DEMs and 1,362 DEGs were found deranged between MM and NPCs. Signatures of 37 DEMs ('Union 37') and 154 DEGs ('Union 154') were deduced that shared 17 DEMs and 22 DEGs with published prognostic signatures, respectively. Two miRs (miR-16-2-3p, 30d-2-3p) correlated with survival outcomes. PPI analysis identified 5 topmost functionally connected hub genes (UBC, ITGA4, HSP90AB1, VCAM1, VCP). Transcription factor regulatory networks were determined for five seed DEGs with ≥ 4 biomarker applications (CDKN1A, CDKN2A, MMP9, IGF1, MKI67) and three topmost up/ down regulated DEMs (miR-23b, 195, let7b/ miR-20a, 155, 92a). This study identified the functionally connected hub genes and transcription factor regulatory networks with miR-16–2-3p and 30d-2-3p as prognostically relevant miRs that correlated with survival outcomes in MM.

6. Kaur G, **Gupta R***, Mathur N, Rani L, Kumar L, Sharma A, Singh V, Gupta A, Sharma OD. Clinical impact of chromothriptic complex chromosomal rearrangements in newly diagnosed multiple myeloma. **Leuk Res.** 2019; 76:58-64. PMID: 30576858

Complex Chromosomal Rearrangements (CCRs) are increasingly being reported as genetic risk factors of clinical significance in cancer owing to their identification using high resolution whole genome profiling technologies. This study employed high resolution CGH + SNP microarrays for whole genome copy number variations (CNV) profiling and identified CCRs in 10% of newly diagnosed MM patients. Multivariable Cox regression model demonstrated a significant association of CTH with poor PFS (HR = 3.09, p = 0.010) and OS (HR = 3.31, p = 0.024) which suggests that CTH is an additional independent prognostic marker in multiple myeloma. Addition of CTH in risk stratification models in clinical setting in multiple myeloma is helpful in upfront identification of high risk patients for suitable customized therapy.

7. Farswan A, Gupta A, **Gupta R***, Hazra S, Khan S, Kumar L, Sharma A. AI-supported modified risk staging for multiple myeloma cancer useful in real-world scenario. **Transl Oncol.** 2021 Jul 8;14(9):101157. doi: 10.1016/j.tranon.2021.101157. PMID: 34247136.

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, β 2-microglobulin (β 2M), calcium, estimated glomerular filtration rate (eGFR) and hemoglobin. MRS was designed using a training set of 716 NDMM patients of our inhouse MM Indian (MMIn) cohort and validated on MMIn (n=354) cohort and MMRF (n=900) cohort. Risk staging rules, obtained via training a J48 classifier, were used to build MRS. New thresholds were identified for albumin (3.6 g/dL), β 2M (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. Our 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 and can be implemented in India with limited resources for genomic testing.

8. Farswan A, Gupta A, Sriram K, Sharma A, Kumar L, **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. doi: 10.3389/fonc.2021.720932. PMID: 34858811.

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. 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. 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. Our methodology discovers changes in the cutoffs with ethnicities from the established cutoffs of prognostic features and the best predictor model for both cohorts was obtained with the new ethnicity-specific cutoffs of clinical parameters. Our study revealed the efficacy of AI in building a deployable risk prediction system for Indian patients with MM which is now available to Indian medical community.

9. Sagar D, Aggarwal P, Farswan A, **Gupta R***, Gupta A. GCRS: A hybrid graph convolutional network for risk stratification in multiple myeloma cancer patients. *Comput Biol Med.* 2022; 149:106048. doi: 10.1016/j.compbiomed.2022.106048. PMID: 36113255.

In this study, we present an efficient Graph Convolutional Network based Risk Stratification system (GCRS) for cancer risk-stage prediction of newly diagnosed multiple myeloma (NDMM) patients. GCRS is a hybrid graph convolutional network consisting of a fusion of multiple connectivity graphs that are used to learn the latent representation of topological structures among patients. This proposed risk stratification system integrates these connectivity graphs prepared from the clinical and laboratory characteristics of NDMM cancer patients for partitioning them into three cancer risk groups: low, intermediate, and high. Extensive experiments demonstrate that GCRS outperforms the existing state-of-the-art methods in terms of C-index and hazard ratio on two publicly available datasets of NDMM patients. We have statistically validated our results using the Cox Proportional-Hazards model, Kaplan-Meier analysis, and log-rank test on progression-free survival (PFS) and overall survival (OS). We have also evaluated the contribution of various clinical parameters as utilized by the GCRS risk stratification system using the SHapley Additive exPlanations (SHAP) analysis, an interpretability algorithm for validating AI methods. Our study reveals the utility of the deep learning approach in building a robust system for cancer risk stage prediction.

10. Das N, Dahiya M, **Gupta R***, Kumar L, Rani L, Gupta A, Farswan A, Sharma A, Sharma OD. Graded Depth of Response and Neoplastic Plasma Cell Index as Indicators of Survival Outcomes in Patients with Multiple Myeloma Following Autologous Stem Cell Transplant. *Am J Clin Pathol.* 2023 Jan 4;159(1):69-80. doi: 10.1093/ajcp/aqac129. PMID: 36317501.

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. In this prospective analysis of patients with MM after autologous stem cell transplant, we evaluated the predictive value of conventional response, measurable residual disease, and neoplastic plasma cell index. 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 had superior PFS (MRD 0.001% to <0.1%: HR, 6.66, $P < .005$; MRD $\geq 0.1\%$: HR, 11.52, $P < .005$) and OS (MRD 0.001% to <0.1%: HR, 5.3, $P < .05$; MRDTOTAL $\geq 0.1\%$: HR = 9.21, $P < .005$). The C index and Akaike information criterion metrics demonstrated the superior performance of the NPCI compared with MRD in predicting treatment outcome. 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, especially in hemodiluted aspirates.