An APOBEC/Inflammation-based classifier improves the stratification of multiple myeloma patients and identifies novel risk subgroups

Sarah Grasedieck  
University of British Columbia

Afsaneh Panahi  
British Columbia Cancer

Matthew C. Jarvis  
University of Minnesota

Faezeh Borzooee  
Simon Fraser University

Reuben S. Harris  
University of Minnesota

Mani Larijani  
Simon Fraser University

Hervé Avet-Loiseau  
IUCT-Oncopole Toulouse

Mehmet Samur  
Harvard Medical School

Nikhil Munshi  
Harvard Medical School

Kevin Song  
Vancouver General Hospital, BC Cancer

Arefeh Rouhi  
British Columbia Cancer

Florian Kuchenbauer  
British Columbia Cancer

Research Article

Keywords: MM, nucleotide editing, prognostic, risk score, t(4;14), del(17)

Posted Date: October 24th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2184702/v1
Abstract

Background: Recent insights into the pathogenesis of multiple myeloma (MM) have highlighted inflammation and genome editing, e.g. by APOBEC enzymes, as major drivers of disease onset and progression. We hypothesized that inclusion of molecular features corresponding to these two mechanisms can be utilized to define novel MM risk groups at initial diagnosis.

Methods: Using two independent patient cohorts (MMRF and IFM/DFCI 2009), we developed and validated an easy-to-calculate novel risk-score that is based on mRNA expression levels of APOBEC2 and APOBEC3B, as well as inflammatory cytokines (IL11, TGFB1 and TGFB3) and serum levels of β2-microglobulin and LDH.

Results: Performance of the Editor- and Inflammation-based score (EI-score) was superior to current cytogenetics-based risk classifiers. Moreover, the EI-score was able to identify previously unrecognized MM patients who experience favourable outcomes despite carrying adverse risk cytogenetics.

Conclusions: Through accurate risk stratification we can identify patients who are currently over-or undertreated. The EI-score is a contemporary and superior prognostic score, calculated based on transcript levels at diagnosis, allowing the identification of unrecognized MM risk subgroups potentially leading to adjustment of clinical treatment and improvement of patient outcomes.

Introduction

Multiple myeloma (MM) is a hematological malignancy characterized by the clonal proliferation of transformed plasma cells [1]. Acquired genetic aberrations drive the initiation and subsequent progression of MM and play a critical role in patient outcome and prognosis [2]. Recent genomic studies have identified mutational genome signatures induced by Apolipoprotein B mRNA Editing Catalytic Polypeptide-like (APOBEC) as independent adverse prognostic factors in MM [3]. The APOBEC family consists of 10 enzymes in humans: APOBEC1, APOBEC2, seven APOBEC3 proteins (APOBEC3A, -3B, -3C, -3D, -3F, -3G, and −3H) and APOBEC4 [4]. They belong to an endogenous family of deaminase enzymes that enhance immune responses through inhibition of retrovirus and retrotransposon replication [5–7]. APOBEC transcription is frequently dysregulated in cancer [8–14] and stimulated by a complex network of innate immune responses, involving factors such as interferons and interleukins, highlighting the tight connection between the immune system, inflammation, and genome editors in cancer cells [5–7, 15].

Although these recent molecular insights provide a better understanding of MM pathogenesis and progression, they have not yet been translated into clinical applications. The current gold standards for MM patient risk classification are the International Staging System (ISS) and the Revised International Staging System (R-ISS), which combine elements of tumor burden (serum β2 microglobulin (β2M), serum albumin, and lactate dehydrogenase (LDH)) with high-risk cytogenetics (del(17p), t(4;14) and t(14;16)) [16]. The Mayo Stratification for Myeloma And Risk-adapted Therapy (mSMART) score is even more inclusive of cytogenetic aberrations, allocating MM patients with t(4;14), t(14;16), t(14;20), gain(1q) or
del(17p) into a high-risk category [16]. Considering that chromosomal translocations often reflect early genetic events in MM cells, caused by Activation-induced cytidine deaminase (AID) [17], we argue that cytogenetic classifiers do not reflect disease progression and could therefore misclassify patients with high-risk cytogenetics but good outcomes.

Early microarray-based gene expression profiling studies that yielded transcriptional classifiers such as UAMS-70 or EMC-92 [18] paved the way for genome/transcriptome sequencing approaches in MM diagnostics, which have already become standard of care in many blood cancers [19]. Recent studies have demonstrated a pivotal role of inflammatory processes, AID and APOBEC3 enzymes in the onset and progression of MM [20–23]. Based on these findings, we hypothesized that RNA-seq data is able to capture disease activity more accurately than cytogenetics and can provide improved prognostic information to predict outcomes, treatment benefits, and potentially identify new MM patient subgroups. Utilizing two independent patient cohorts with available RNA-seq information, we built and validated a polynucleotide Editor/Inflammation based scoring system (EI-score) which is based on mRNA gene expression levels of APOBECs, inflammatory cytokines, and serum levels of the known prognostic proteins β2M and LDH. This simple score allowed us to identify MM risk groups more accurately and to predict treatment outcomes independent of cytogenetics.

**Materials And Methods**

**Patient cohorts and available data**

Data from 1,143 patients with newly diagnosed MM (NDMM) and available survival information was obtained through the CoMMpass database version IA14, which was generated as part of the Multiple Myeloma Research Foundation Personalized Medicine Initiatives (www.themmrf.org). Fluorescence *in situ* hybridization (FISH) test results, ISS and R-ISS staging information was available for 711, 1114 and 690 patients, respectively. For 599 patients, information on both blood parameters and RNA-seq from CD138+ sorted bone marrow (BM) cells was available. As an independent validation cohort, we utilized data collected from 263 NDMM patients in the IFM/DFCI 2009 clinical trial (ClinicalTrials.gov identifier: NCT01191060) with available clinical data, FISH result (del (17p), t(4;14), t(14;16)) and RNA-seq data from CD138+ BM cells. IFM/DFCI MM patients were treated with lenalidomide, bortezomib, and dexamethasone (RVD) alone or with RVD+ autologous stem cell transplant (ASCT). All patient baseline characteristics (CoMMpass and IFM/DFCI) are summarized in *Appendix Table 1 (online only)*.

**Feature selection, model building, performance evaluation, and validation**

A stepwise workflow for the evaluation and selection of individual features and multivariate models in the CoMMpass data is shown in *Figure 1*. Transcript per million (TPM) normalized gene expression and canonical genomic mutational variants per patient were obtained through the CoMMpass database MMRF [24]. First, we performed univariate Cox Proportional Hazard (CoxPH) analyses using the *R survival* package [25] on 163 pre-selected variables (*Appendix Table 2-4 (online only)*) comprised of demographic features (n=6), clinical blood parameters (n=19), cytogenetic abnormalities (n=5), cell
surface receptors assessed by flow cytometry \((n=9)\), mRNA transcript levels of \textit{APOBEC}, pro-/anti-inflammatory cytokines and cytokine receptors \((n=119)\), total mutational burden, as well as computed single base substitution (SBS) genomic mutational signatures linked to ageing and APOBEC activity \((n=5)\). SBS mutational signatures as listed in COSMIC version 3 \((\text{SBS1/5} = \text{age-associated}, \text{SBS2/13} = \text{APOBEC-associated})\) as well as an improved C-to-T/C-to-G in 5’TC(A/T motifs) APOBEC enrichment score were calculated as described by Jarvis \textit{et al.} \cite{26}. All features were assessed in univariate as well as in multivariate CoxPH analyses in order to select only for variables that impacted on the association between either patient age or first-line treatment with patient survival, based on either log likelihood or Chi-square tests. All continuous parameters \(\text{(including RNA transcript levels and blood measurements)}\) were transformed to categorical variables applying maximally selected rank statistics \(\text{\textit{R survival package}}\) to define the cut-off between \textit{high} and \textit{low} patient groups for each individual feature \cite{27}. Although maximally selected rank statistics calculated cut-offs were used to calculate the final cut-offs in our presented models and scores, we have also computed median- as well as 1st and 3rd quantile cut-offs for all candidate variables and only selected variables to move on to multivariate testing that were significantly associated with outcome in at least two of the four cut-off categories. All remaining age- and treatment-independent prognostic variables with expression >5 fragments per kilobase per million were then combined into multivariate CoxPH models. Parameters that showed the greatest effect size \((z\text{-score})\) were combined and weighted based on their rounded integer multivariate \(z\text{-score}\), yielding an easy to calculate score formula \(\text{\textit{see Results section}}\).

Prediction accuracy metrics of this combination of score variables were determined through multivariate CoxPH analyses and by training three individual machine-learning models after randomly splitting the data into training \((75\%)\) and testing \((25\%)\) cohorts applying the Monte Carlo cross-validation approach, including a random forest \((\text{rf})\), gradient boosting \((\text{gbm})\), and negative binomial \((\text{nb})\) model, both for overall survival \((\text{OS})\) and progression free survival \((\text{PFS})\) prediction \(\text{\textit{R caret, tidymodels, tidyverse, ranger, MLeval} [28-32]. Model performance on the testing data was evaluated based on receiver operating characteristic \text{\textit{(ROC) as well as precision recall \text{\textit{(PR) efficiency and recorded in the form of ROC area under the curve \text{\textit{(AUC) values. Reproducibility was assessed using a leave-one-out approach for multivariate Cox regression, and 10-fold cross validation for all machine learning approaches. An independent validation in the IFM/DFCI cohort was performed by applying the score formula. Stratification of patients into EI-score \textit{high, intermediate, and low} groups was determined based on natural valleys in the score distribution.}}}}}}

\textbf{Results}

\textbf{Construction of a prognostic Editor/Inflammation-based score}

An overview of the workflow is shown in \textbf{Figure 1}. In brief, we applied univariate and multivariate Cox regression analysis testing 163 selected features \((\text{Appendix Table 1-4 (online only)})\) to identify covariates with significant time-to-event outcomes that were independent from patient age and first-line treatment \(\text{(hazard ratios and individual test statistics are shown in \text{Figure 2 and Appendix Table 5 (online only)})}\). To
identify potentially interesting relationships, we computed pairwise Pearson's correlation between all survival-associated parameters and observed a positive correlation between multiple APOBEC and cytokine transcript levels, which was most significant among APOBEC members 3C, -3D, -3F, and -3G ($R=0.81-0.9$, $p<5\times10^{-18}$, Appendix Figure 1A (online only)). Furthermore, APOBEC3A levels correlated significantly with a subset of interleukin receptors comprising the IL-8 receptors, CXCR1 and CXCR2 ($R=0.87$, $p<2\times10^{-24}$), CSF3R ($R=0.92$, $p=4\times10^{-11}$), as well as the IL10 receptor paralog IFNAR1 ($R>0.89$, $p=2\times10^{-21}$), further supporting the connection between APOBECs and inflammation in MM (Appendix Figure 1B, (online only)).

A fourteen-parameter model that reproducibly showed the best performance in predicting PFS and OS included β2M, hemoglobin, LDH as well as RNA transcript levels of three APOBEC family members (APOBEC2, APOBEC3B, APOBEC3C), six pro/anti-inflammatory cytokines (IL10, IL11, IL27, IFNG, TGFB1, TGFB3) and all tested genome mutational signatures (SBS1, SBS2, SBS5, SBS13, SBS40) [33]. Interestingly, prediction of PFS in CoMMpass patients strongly relied on low APOBEC2 expression levels (PFS $z=-4.34$, $p=0.00001$ vs. OS $z=-3.56$, $p=0.0004$, (Appendix Table 6, (online only))) and was improved by inclusion of the SBS1 ageing- and SBS2 APOBEC-associated mutational signatures (PFS: $p=0.004$ and $p=0.005$, respectively vs OS: not significant). In contrast, OS prediction in this patient cohort was improved by the inclusion of creatinine blood levels and RNA levels of cytokines such as TGFB3 and IL11 (Appendix Table 6, (online only)). Notably, only one out of five cytogenetic features, being gain(1q), passed all of our selection criteria.

It should be noted that the aim of this study was not to establish a de novo transcription-based classifier for NDMM patients but rather to test whether incorporation of disease-associated and therefore pre-selected features would be able to predict patient outcome with comparable or greater accuracy than existing classifiers that heavily rely on cytogenetics. As a proof of principle, we therefore decided to further reduce our editor/inflammation-associated feature set from 16 to 7 variables, only retaining those that were most significant for both PFS and OS prediction. Mutational signatures were excluded as they are difficult to compute in clinical practice. Lastly, we applied a simplified weighting based on the rounded integer multivariate z-score of each parameter:

**EI-score[OS]:** $(APOBEC2: \leq 0.20=0, >0.20=3) + (APOBEC3B: \leq 3.79=0, >3.79=3) + (IL11: \leq 0.48=0, >0.48=2.5) + (TGFB1: \leq 0.11=0, >0.11=1) + (TGFB3: \leq 0.10=0, >0.10=2) + (\beta2M: \leq 4.22=0, >4.22=4) + (LDH: \leq 3.18=0, >3.18=2)$

**EI-score[PFS]:** $(APOBEC2: \leq 0.19=0, >0.19=4) + (APOBEC3B: \leq 6.91=0, >6.91=3) + (IL11: \leq 0.56=0, >0.56=2) + (TGFB1: \leq 0.43=0, >0.43=1) + (TGFB3: \leq 0.03=0, >0.03=1) + (\beta2M: \leq 3.33=0, >3.33=4) + (LDH: \leq 3.08=0, >3.08=2)$

A total of 599 CoMMpass patients had sufficient data to compute both EI-scores. Computed EI-score[OS] values applying the above formula in these 599 patients ranged from 0 to 17.5 and computed EI-score[PFS] values from 0 to 15 (Appendix Figure 2A and B, (online only)). The median PFS and OS was
688 (95% CI; 605 to 760) and 1052 (95% CI; 996 to 1094) days, respectively. Dichotomization of this patient subset into EI-score[OS] standard risk and high-risk patients groups allowed the classification of a standard-risk cohort with a 5-year OS rate of more than 50% and median OS of 2207 days (95% CI; 2207 to 2207) as well as a high-risk group with a median OS of 1500 days (95% CI; 934 to 1500; Figure 3A). Similarly, dichotomization into EI-score[PFS] standard- and high-risk patient groups allowed the classification of a standard-risk cohort with a median PFS of 1302 days (95% CI; 1176 to 1472) and a high-risk cohort with a median PFS of 485 days (95% CI; 390 to 604, Figure 3B). Overall, the EI-score accurately predicted outcome in the CoMMpass patient cohort and we observed a positive correlation between EI-score[PFS] and EI-score[OS] (R=0.78; P<0.0001, Appendix Figure 2C, (online only)). Applying the EI-score formulas to the independent IFM/DFCI patient cohort (n=263), we could confirm the above findings by dichotomization into two EI-score[OS] risk groups (47% and 53%, respectively) with a 5-year OS rate of 87% and 73% (p=0.0062; Figure 3C), respectively. Similarly, PFS was dichotomized into two EI-score[PFS] risk groups (45% and 55%, respectively) with 37% and 25% 5-year OS rates, respectively (p=0.0007; Figure 3D). Of note, the reported 5-year OS rate of 78% in the IFM/DFCI 2009 study was superior to only 68% in the CoMMpass cohort.

The EI-score outperforms clinically established MM risk classifiers

To compare the prognostic accuracy of the EI-score to ISS, R-ISS, and mSMARTcyto (a reduced version of the mSMART score based exclusively on presence of t(4;14), t(14;16), gain 1q and/or del 17p), we computed EI-score performance metrics including the multivariate CoxPH Concordance index (Cᵢ) and ROC-AUC in MMRF patients applying a standardized machine learning pipeline (Table 1). The EI-score achieved the best performance ratings for OS and PFS prediction (n=599; Cᵢ 0.7 and 0.69, respectively; Table 1) followed by ISS (n=1113; Cᵢ 0.66 and 0.6), R-ISS (n=690; Cᵢ 0.64 and 0.6), and mSMARTcyto (n=817; Cᵢ 0.58 and 0.54). Even after exclusion of patients allocated to the ISS, R-ISS, and mSMARTcyto intermediate risk classes, the EI-score showed superior classification accuracy, especially in predicting unfavourable outcomes (Table 1).

We then assessed which “score classes” (blood markers, editors, or cytokines) had the most weight in clinical risk prediction. Although the best stratification was achieved when applying a combination of all three score classes, both APOBEC genes and blood parameters showed the most significant differential presence in high compared to low clinically defined patient risk groups (Appendix Figure 3A-B, online only). Next, we asked whether addition of gene expression information can improve the performance of established risk classifiers. Therefore, we added the individual EI-score components of each score class to the risk classifiers ISS, R-ISS and mSMARTcyto, which significantly improved their performance of PFS and OS prediction. The best results were achieved when transcript levels of both editors and inflammatory cytokines were added simultaneously to the established scores (Table 1). Overall, classifiers that included routinely assessed blood parameters, which are already known to be predictive of MM disease outcome (β2M and LDH), had the highest classification accuracy and scores based on cytogenetics performed poorest among all tested combinations (Appendix Table 6 online only). In line
with our observation that adding gene expression parameters of the EI-score to ISS, R-ISS, and mSMARTcyto significantly improved their accuracy, we observed further improvement of these classifiers when cytogenetic information was removed completely (Table 1, see R-ISSnocyto), underscoring the potential of gene expression classifiers such as the EI-score for MM outcome prediction.

Applying the EI-score, we next aimed to identify MM patients with very high and very low risk of progression or death, similar to the three-risk-group stratifiers ISS and R-ISS. Based on natural valleys in the score distribution (Appendix Figure 2A,B, online only), we selected two cut-offs (EI-score 9 and 3), yielding EI-score low-, intermediate- and high-risk groups (comprising 31%, 56% and 13% of the investigated MMRF patient cohort (n=599); Figure 3E-F). The 5-year OS rate for patients classified into the individual groups were 85% (EI-score[OS] low-risk), 65% (intermediate-risk) and 35% (high-risk). Overall, hazard ratios for patients in the EI-score[OS] high category were four times higher than those of patients categorized as EI-score[OS] low (16.9 versus 69.9, 95% CI). A detailed comparison between the EI-score, ISS and R-ISS risk group survival probabilities are shown in Appendix Table 7 and Appendix Figure 4 (online only). Eventually, we applied the EI-score to already classified ISS and R-ISS risk groups which allowed for a further sub-stratification of these patients, revealing previously unrecognized risk subgroups (Figure 4A-B).

To adjust for the heterogeneous treatment protocols of patients included in the CoMMpass dataset, we performed a sub analysis for MM patients receiving CyBorD or VRD ± autologous stem cell transplantation. Also in these treatment groups, the EI-score outperformed ISS and R-ISS in predicting patient outcome (Figures 5A and B).

The EI-score identifies novel MM risk subgroups

Recent publications point towards the necessity of a more refined classification of high-risk MM patients. To test the suitability of the EI-score for sub-stratifying adverse risk MM patients, we applied the EI-score to patient subgroups that carried either del(17p), gain(1q) or t(4;14). With this approach, we were able to identify patients with poor and favorable outcomes in the CoMMpass (Figure 4C) and the IFM/DFCI cohorts (Figure 4D). In the CoMMpass dataset, the EI-score was able to subclassify del(17p) MM patients into three main risk subgroups: a very good prognosis group (0% with additional TP53mut) with 5-year OS of 100%, an intermediate group (30% with additional TP53mut) with 5-year OS rate of 75%, and a very poor prognosis group (40% with additional TP53mut) with 5-year OS of 0% (2-year OS: 40%) (CoMMpass data, Figure 4C). These finding could be reproduced in the IFM/DCFI cohort, where application of the EI-score allowed the subclassification of del(17p) MM patients into a standard (5-year OS rate: 73%), and a high-risk group (5-year OS of 35%) (Figure 4D). Similarly, the EI-score allowed us to identify t(4;14) MM patients with a very poor prognosis (median OS: 20 months) (Figure 4D). In line with the reported pathophysiological relevance of nucleotide editors for MM, we found that CoMMpass study patients who carried either del(17p), gain(1q) or t(4;14) and also achieved a high EI-score, displayed an enrichment of APOBEC-induced genomic mutations compared to intermediate and low EI-score patients (Appendix Figure 5, online only). Taken together, our data demonstrates that a simple gene expression-based score,
composed to reflect disease biology, is able to identify previously unrecognized subgroups of MM patients who display adverse risk cytogenetics but experience favorable outcomes.

**Discussion**

High *APOBEC3B* expression has been shown to induce mutations that contribute to disease evolution in pro-inflammatory environments, which in turn induces *APOBEC* expression and thus creates a feedback loop that ultimately causes cell damage [7, 34, 35]. Based on this recently discovered pathological framework in MM, we have developed a prognostic score that incorporates mRNA levels of *APOBEC* family- and inflammatory genes as well as established blood parameters. In contrast, probe-based RNA-based classifiers such as EMC-92 and UAMS-70 rather reflect optimal performance of a gene signature, rather than biological processes driving the pathogenesis of MM and disease burden. A potential explanation why *APOBEC* members have not been part of past, probe-based classifiers is the probes’ propensity to cross-hybridize to sequences with high similarity, as is the case among members of the AID/APOBEC families. Cross-hybridization can result in decreased specificity as has been shown for *APOBEC* mRNAs in the Affymetrix Gene Chip Human Genome U133 Plus 2.0 Array, the platform that was used to build the UAMS-70 and EMC-92 gene classifiers [36].

Notably, the majority of cytogenetic abnormalities did not pass our stringent score building criteria due to their - in relative terms - weak association with patient outcomes. Among all APOBEC/AID family members, transcript levels of *APOBEC3B* achieved the highest predictive importance metrics in our analyses and exhibited greater prognostic value for NDMM patients compared to APOBEC-associated genomic mutational signatures. As of yet, the function of *APOBEC2* in MM still needs to be defined. *APOBEC2* was reported to induce mitophagy in muscle cells and is thought to affect liver cancer cell proliferation [8, 37]. MM patients typically exhibit a dysregulation of pro/anti-inflammatory cytokines which was shown to promote MM proliferation and tissue invasion [20, 38, 39], a possible explanation of the observed *APOBEC2* transcriptional upregulation. IL-11 is part of a cytokine group including IL-6 [40], leading to the activation of the JAK-STAT pathway and has been considered as an osteoclastogenic interleukin supporting the growth of osteoclasts [20]. Notably, IL-11 levels were found to be elevated in MM patients [41]. Levels of TGF-b, another cytokine considered in the EI-score, have also been implicated in MM bone lesions through blocking the differentiation of osteoblasts [42] as well as the proliferation of MM cells [43]. These findings support the relevance of IL-11 and TGF-b as potential molecular markers of osteolytic activity.

Currently, the R-ISS cytogenetics-based risk score is the most widely used classifier in clinical application. Within this study, we could show that the prognostic value of cytogenetics-based risk scores is limited as defining high-risk groups based on cytogenetic abnormalities potentially masks patient subsets experiencing better outcomes. This finding is especially relevant as this unexpectedly favorable risk group could benefit from reduced treatment intensity, thus minimizing side effects and cost impacts on the health care system. We further observed a significant correlation between *APOBEC3B* transcript levels and the APOBEC genomic mutation enrichment score in MM patients carrying del(17p) and/or TP53mut.
These findings support the idea of a synergy between del(17p) and simultaneously elevated APOBEC mRNA levels to serve as an optimal environment for APOBEC-induced mutagenesis and tumor progression [44].

To implement novel risk classifiers such as the EI-score, prospective validation - ideally under real world conditions rather than in highly selected MM patient cohorts - will be necessary, preferentially including up-to-date treatment regimens such as daratumumab. In addition, the EI-score provides a rationale for further preclinical research to target APOBEC3B high MM cells as a novel therapeutic approach [45].

In conclusion, this study describes the development of the EI-score, a simple seven-parameter diagnostic tool that functions independent of patient cytogenetics in classifying patients into high-risk and standard-risk MM. Besides the documented clinical relevance, our data further supports the implementation of molecular biomarkers to stratify the risk of NDMM patients, which may also help to shed light on the intricate molecular mechanisms that drive MM disease progression.

**Abbreviations**

Multiple myeloma (MM); Newly diagnosed multiple myeloma (NDMM); Progression free survival (PFS); Overall survival (OS); Apolipoprotein B mRNA Editing Catalytic Polypeptide-like (APOBEC); Editor- and Inflammation-based score (EI-score); International Staging System (ISS), Revised International Staging System (R-ISS); ß2 microglobulin (ß2M); Lactate dehydrogenase (LDH)); The Mayo Stratification for Myeloma And Risk-adapted Therapy (mSMART); Activation-induced cytidine deaminase (AID); Fluorescence in situ hybridization (FISH), Lenalidomide; bortezomib, and dexamethasone (RVD); Autologous stem cell transplant (ASCT); Bone marrow (BM); Transcript per million (TPM); Univariate Cox Proportional Hazard (CoxPH); Single base substitution (SBS); Receiver operating characteristic (ROC); Precision recall (PR); Area under the curve (AUC); Random forest (rf); Gradient boosting (gbm); Negative binomial (nb); Concordance index (C).

**Declarations**

**Ethics approval and consent to participate:** The MM Research Foundation (MMRF) CoMMpass (ClinicalTrials.gov identifier: NCT01454297) data set was used to construct and validate the EI-score. Samples from patients with MM were collected from the IFM/DFCI 2009 study (ClinicalTrials.gov identifier: NCT01191060) after written informed consent, and clinical and genomic data were de-identified in accordance with the Declaration of Helsinki and used to validate the EI-score.

**Consent for publication:** not applicable.

**Availability of data and materials:** The MMRF CoMMpass dataset was generated as part of the Multiple Myeloma Research Foundation Personalized Medicine Initiatives and access can be requested at: https://research.themmrf.org and www.themmrf.org). Data from the IFM/DFCI 2009 study was provided by Dr. Nikhil Munshi and will be made available from the corresponding author upon reasonable request.
Competing interests: The authors declare no potential conflicts of interest.

Funding: This work was supported by the BC Cancer Foundation and the Park, Armstrong, McKeen, Hungerford, Carruthers and DeLasalle families as well as through the International Myeloma Society (IMS) and a Paula and Rodger Riney Foundation Translational Research Grant and the LLSC. SG was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation – project 446251518), the Michael Smith Foundation for Health Research (MSFHR) and the Lotte & John Hecht Memorial Foundation (project RT-2020-0578). AP was supported by a four-year UBC doctoral fellowship (#6569) and an Echoridge Educational Foundation Scholarship (VGH & UBC Hospital Foundation). ML received support through LLSC and CIHR grants (MOP111132). FB was supported by Weyerhaeuser Grad Schol (MBB) scholarship from Simon Fraser University.

Authors' contributions: This study was designed by FK, AR, SG, AP, and supervised by FK, AR, KS and ML. Data analysis was carried out by SG, AP, and MJ. NM, MS, and HA provided (IFM/DFCI 2009) patient RNA sequencing and clinical data. The manuscript was written by FK, AR, SG, AP, FB and revised by FK, AR, KS, ML, RH, NM, and MS. All authors approved the manuscript.

Acknowledgements: We thank the MMRF for kindly providing access to the data from the CoMMpass study.

References


41. Urbańska-Ryś H, Wiersbowska A, Stepień H, Robak T. Relationship between circulating interleukin-10 (IL-10) with interleukin-6 (IL-6) type cytokines (IL-6, interleukin-11 (IL-11), oncostatin M (OSM)) and soluble interleukin-6 (IL-6) receptor (sIL-6R) in patients with multiple myeloma. *Eur Cytokine Netw*. 2000;11(3):443-451.


**Table**

Table 1 is available in the Supplementary Files section

**Figures**
Figure 1

Feature selection and multivariate model construction workflow.

1. **Feature selection**
   - CoMMpass Trial, n=1143 patients
   - Sum of tested features = 163
   - [demographics (6), cytogenetics (5), cell surface markers (9), genes (119), clinical and blood parameters (19), genomic mutational signatures (5)]

2. **Feature evaluation**
   - Univariate CoxPH and Kaplan-Meier analysis
   - Features sign. associated with survival, HR>1 (3rd Quartile gene expression >0.5 TPM): Meet criteria: OS(25), PFS(21)

3. **El-score building**
   - CoxPH/Kaplan-Meier analysis
   - Find optimal cut-offs and feature weightings to build an easily applicable score that has a high predictive accuracy.

4. **El-score evaluation**
   - Multivariate CoxPH + machine learning (ROC-AUC)
   - Evaluate El-score sensitivity and specificity in comparison to established diagnostic scores.

5. **El-score validation**
   - IFM/DFCI trial, n=263 patients
   - Apply El-score to an independent patient cohort and evaluate performance compared to established diagnostic scores.

**Figure 1.**
### Figure 2

**Univariate Hazard ratios for survival-associated variables in MMRF datasets.** Cox proportional hazard was calculated for 163 selected variables at 95% confidence, shown are all variables that were significantly associated with OS and PFS. Cut-offs for continuous variables were set either based on maximally selected rank statistics (for RNA classifiers), to 3rd quartiles (for genome mutational signatures), or to a fixed cut-off, which was applied for patient age (>75).
Graphical representation of EI-score Kaplan-Meier estimates in MMRF and IFM/DFCI 2009 dataset. A. EI-score[OS] and B. EI-score[PFS] calculated for each patient in the MMRF dataset. Dichotomization into two risk groups was performed based on maximally selected rank statistics. C. EI-score[OS] and D. EI-score[PFS] calculated for each patient in the IFM/DFCI 2009 dataset. E-F. Graphical representation of Kaplan-Meier estimates based on EI-score[OS] and EI-score[PFS] calculated for each patient in the MMRF
dataset. Stratification into high-, intermediate- and low-risk groups was performed based on naturally occurring valleys in the EI-score[OS] and EI-score[PFS] data distribution.

Figure 4

Applying the EI-score to pre-risk classified MM patients allows identification of favourable/unfavourable survival outcomes. Shown are graphical representations of Kaplan-Meier estimates based on
reapplication of the EI-score[PFS] to MMRF patients who were classified into **A.** ISS stage I, II and III risk groups; **B.** R-ISS stage I, II and III risk groups; **C.** t(4;14), gain1(q) and del 17(p) adverse risk groups. **D.** to IFM/DFCI patients who were classified into t(4;14) and del 17(p) adverse risk groups.

**Figure 5.**
EI-score allowed an accurate prediction of patient outcomes in MMRF dataset. A. KM curve analysis of patients who received CyBorD as the induction regimen with or without SCT. B. KM curve analysis of patients who received VRD as the induction regimen with or without SCT.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Table.pdf
- AppendixTable1.docx
- AppendixTable2.docx
- AppendixTable3.docx
- AppendixTable4.docx
- AppendixTable5.docx
- AppendixTable6.xlsx
- AppendixTable7.docx
- AppendixFigure1.docx
- AppendixFigure2.docx
- AppendixFigure3.pdf
- AppendixFigure4.docx
- AppendixFigure5.docx