Development and Validation of a Novel Prognostic Model for Overall Survival in Newly Diagnosed Multiple Myeloma Integrating Tumor Burden and Comorbidities

Shuangshuang Jia†, Lei Bi†, Yuping Chu, Xiao Liu, Juan Feng, Li Xu, Tao Zhang, Hongtao Gu, Lan Yang, Qingxian Bai, Rong Liang, Biao Tian, Yaya Gao, Hailong Tang* and Guangxun Gao*

Department of Hematology, Xijing Hospital, Air Force Medical University, Xi’an, China

Background: Multiple myeloma (MM) is a highly heterogeneous disease with enormously variable outcomes. It remains to be a major challenge to conduct a more precise estimation of the survival of MM patients. The existing stratifications attached less importance to the prognostic significance of comorbidities. In the present study, we aimed to develop and validate a novel and simple prognostic stratification integrating tumor burden and comorbidities measured by HCT-CI.

Method: We retrospectively enrolled 385 consecutive newly diagnosed multiple myeloma (NDMM) patients in Xijing Hospital from January 2013 to December 2020. The cohort between January 2016 and December 2020 was selected as development cohort (N = 233), and the cohort between January 2013 and December 2015 was determined as validation cohort (N = 152). By using LASSO analysis and univariate and multivariable Cox regression analyses, we developed the MM-BHAP model in the way of nomogram composed of β2-MG, HCT-CI, ALB, and PBPC. We internally and externally validated the MM-BHAP model and compared it with ISS stage and R-ISS stage.

Results: The MM-BHAP model was superior to the ISS stage and partially better than the R-ISS stage according to time-dependent AUC, time-dependent C-index, DCA, IDI, and continuous NRI analyses. In predicting OS, only the MM-BHAP stratification clearly divided patients into three groups while both the ISS stage and R-ISS stage had poor classifications in patients with stage I and stage II. Moreover, the MM-BHAP stratification and the R-ISS stage performed well in predicting PFS, but not for the ISS stage.
Besides, the MM-BHAP model was also applied to the patients with age ≤65 or age >65 and with or without HRCA and could enhance R-ISS or ISS classifications.

**Conclusions:** Our study offered a novel simple MM-BHAP stratification containing tumor burden and comorbidities to predict outcomes in the real-world unselected NDMM population.

**Keywords:** prognostic model, risk stratification, HCT-CI, comorbidity, multiple myeloma

**INTRODUCTION**

Multiple myeloma (MM) is the second most common hematologic malignancy characterized by hyperproliferation of clonal plasma cells within the bone marrow (1). Despite its incurability, the median overall survival (OS) in MM patients has obtained a significant improvement over the years due to the introduction of novel therapies, such as proteasome inhibitors (PIs), immunomodulatory drugs (IMiDs), monoclonal antibodies, and immunotherapy. However, even among the patients with the same genetic background currently known, the high heterogeneity of the outcome still exists, with the survival varying from a few months to more than 10 years (2–4). Therefore, it remains to be a major challenge to conduct precise estimation of the survival in MM patients.

The current International Staging System (ISS) (5) and Revised International Staging System (R-ISS) (6) were widely used in clinical practices. The R-ISS stage combined the ISS stage, lactate dehydrogenase (LDH), and high-risk chromosomal abnormalities (HRCA), which achieved an improvement in estimating prognosis compared to the ISS stage. However, the R-ISS was composed of disease-related factors, which mainly reflected the inherent biological characteristics of myeloma but reckoned without the patient-related factors like comorbidities and performance status. Moreover, there was a large portion of patients distributed into R-ISS stage II (6), indicating that it was reckoned without the patient-related factors like comorbidities and performance status. Therefore, it remains to be a major challenge to conduct a more precise estimation of the survival in these patients. Besides, the R-ISS stage was based on the data derived from clinical trials and might result in some limitations for the application in the real-world unselected patients.

Based on the abovementioned reasons, efforts were always underway to explore new risk stratifications for the prediction of survival in the patients with newly diagnosed multiple myeloma (NDMM). Some studies established prognostic tools with biological parameters, such as gene expression (7–14) and lncRNA (15), followed by the limitations of the complexity and non-standardization. A few of new prognostic models predicting survival were based on new integration of chromosomal abnormalities and clinical indicators, whereas these only focused on disease-related factors (2, 16–18).

Therefore, a more comprehensive assessment, orchestrating the genetic landscape of myeloma and host characteristics, may be more appropriate to distinguish the benefits from existing treatments in an unselected community setting. For example, the UK Myeloma Research Alliance Risk Profile (MRP) taking account of WHO performance status and age, together with ISS stage and C-reactive protein, could help to predict the prognosis and therapy delivery in patients who are not candidates for transplantation (19). Furthermore, a survival matrix was created using the factors of age, del (17p), triplet therapy use, EQ-5D mobility, ISS stage, solitary plasmacytoma, history of diabetes, platelet count, Eastern Cooperative Oncology Group performance status, and serum creatinine for predicting the outcomes of NDMM patients (20). Moreover, a pleural effusion-based nomogram including factors of pleural effusion, plasma cell proportion in the bone marrow, ISS stage, Charlson Comorbidity Index (CCI), 1q21 gain, and autologous hematopoietic stem cell transplantation (HSCT) (21) was also performed to evaluate the prognosis in unselected MM population.

It is obvious that there is still a lack of comprehensive studies for the prognostic significance of comorbidities which are usually excluded from the clinical trials but maybe partially determine the treatment options, treatment intensity, and time to next treatment (3, 4). Both the CCI and hematopoietic cell transplantation–comorbidity index (HCT-CI) were ways of evaluating the impact of comorbidities. However, the HCT-CI could have clearer definitions of comorbidities than the CCI (22) and also considered recent infections which usually led to less intensive therapies and associated with worse OS (23). Moreover, HCT-CI was widely used to predict the survival probabilities of patients after HSCT (24–27), while little attention was paid to its prognostic value in the outcome of NDMM patients.

Hence, we explored the possibility of predicting the survival of MM patients with the HCT-CI evaluated at the time of diagnosis. Using routinely available clinical factors that integrated comorbidities and tumor burden, we developed and validated a new prognostic model and risk stratification for predicting the probability of 6-month, 1-year, 2-year, and 4-year OS of patients with NDMM. Our study offered a novel simple tool to predict outcomes in NDMM patients, as a supplement for a better stratification on the basis of current risk classifications.

**METHODS**

**Cohort Selection**

We retrospectively enrolled 385 consecutive newly diagnosed MM patients in our institution from January 2013 to December 2020. All patients were aged at least 18 years, diagnosed...
according to International Myeloma Working Group (IMWG) criteria (28) and followed up for the available treatment and survival information until June 1, 2021. The cohort between January 2016 and December 2020 was selected as development cohort \((N = 233)\), and the cohort between January 2013 and December 2015 was determined as validation cohort \((N = 152)\).

**Data Collection**

The baseline characteristics, such as patient/disease-specific data at diagnosis and treatment information, were collected. The patient-specific data included age, sex, body mass index (BMI, kg/m\(^2\)), history of hypertension, history of thrombosis, and HCT-CI. The disease-specific data contained white blood cell (WBC, \(10^9/L\)), neutrophil granulocyte (NEU, \(10^9/L\)), lymphocyte (LYM, \(10^9/L\)), monocyte (MONO, \(10^9/L\)), platelet (PLT, \(10^9/L\)), \(\beta\)-2-microglobulin (\(\beta\)-2-MG, mg/L), albumin (ALB, g/L), serum calcium (Ca\(^{2+}\), mmol/L), lactate dehydrogenase (LDH, IU/L), high-risk chromosomal abnormalities (HRCA), bone marrow plasma cells (BMPC), peripheral blood plasma cells (PBPC), DS stage, ISS stage, and R-ISS stage. The treatment information involved novel therapies and autologous stem-cell transplantation (ASCT).

**Definition of Some Variables and Survival Outcomes**

The condition of infection in HCT-CI was assessed at the time before induction therapy and other comorbidities that were measured at diagnosis in our study. The HRCA was defined as the presence of t (4;14) and/or t (14;16) and/or del(17p), detected by fluorescence in situ hybridization (FISH) (6). The thresholds were set at 10% for t (4;14) and t (14;16) and at 20% for del(17p), recommended by the European Myeloma Network (29). The BMPC was referred to as the proportion of clonal plasma cells on a bone marrow smear. The PBPC meant the proportion of clonal plasma cells on a peripheral blood smear.

The time of last follow-up was on June 1, 2021. The primary end point was OS defined as the time from the start of diagnosis until all-cause death or until the last follow-up time the patient was known to be alive. The secondary end point was progression-free survival (PFS) defined as the time from the start of diagnosis until progression or all-cause death or until the last follow-up time the patient was known to be progression-free.

**Development of the New Prognostic Model**

Before variable selection, we examined the non-linear association between continuous variables with OS via restricted cubic splines based on Cox regression (Supplemental Figure S1), then transformed the variable into categorical variable according to the cutoff points when the \(p\) value for non-linearity <0.05. We handled missing data on candidate prognostic variables using multivariate imputation by chained equation (MICE) and created five imputed datasets. Then, we evaluated the potential prognostic value of candidate variables by using univariate Cox regression analysis in all the five imputed datasets and the coefficients were combined with R package “mice.” The variables with \(p\) value < 0.10 were subjected to the least absolute shrinkage and selection operator (LASSO) analysis to prevent overfitting (30), and we finally picked up four predictors of them taking account of clinical importance and prediction of the model (Supplemental Methods).

**Validation of the New Prognostic Model**

We internally and externally validated the predictive power of the model respectively in the development cohort and validation cohort via the following analyses: (1) discrimination: assessed by time-dependent area under the curve (AUC) of the receiver operator characteristic (ROC) and time-dependent Harrell’s concordance index (C-index) analyses; (2) calibration: examined by calibration curve with 1,000 bootstrap resamples, which indicated a good fit of the predicted probabilities with actual outcome frequencies when the curve had a good agreement with the 45° diagonal line; (3) clinical usefulness: estimated by decision curve analysis (DCA), which showed that the model was the best choice for all patients that had the highest net benefits among all the range of risk thresholds (31); and (4) improvement in prediction: tested by integrated discrimination improvement (IDI) and continuous net reclassification index (NRI), which suggested that the new model had an improvement in predictive capacity compared with the old model when they were greater than zero. In particular, we would like to compare the performance of the new prognostic model with that of the R-ISS stage which showed missing data in the development cohort. To reduce the influence of data deficiency, we carried out the abovementioned analyses in each imputed dataset.

**Statistical Methods**

All the statistical analyses were carried out using R version 4.1.0 and SPSS version 26.0, and a two-sided \(p < 0.05\) suggested a statistical significance. The chi-square test or Fisher’s exact test was used to analyze qualitative variables, and the Mann–Whitney U test was used to analyze quantitative variables. Survival was analyzed by Kaplan–Meier curves and log-rank tests. Uni- and multivariable Cox proportional hazard models were used to assess the prognostic factors and calculated hazard ratios (HR) with 95% confidence intervals (CI). The R packages used in the abovementioned analyses were shown in Supplemental Methods.

**RESULTS**

**Patients’ Characteristics**

We presented the baseline characteristics of the development cohort \((N = 233)\) and validation cohort \((N = 152)\) in the Supplemental Table S1. Moreover, the median age was respectively 59 (35–88) and 58.5 (18–89) years. In the development cohort, the vast majority (99.6%) of the patients received anti-myeloma treatment and most (98.3%) patients underwent novel therapy in their induction treatment. Similarly, 98.7% of the patients were treated with anti-myeloma therapy and 98.0% of the patients received novel therapy in the validation cohort. As of the end of June 1, 2021,
the median time of follow-up was 24.0 (range from 0.1 to 61.8) and 28.8 (range from 0.4 to 101.9) months in the development cohort and validation cohort, respectively.

Development and Evaluation of the MM-BHAP Model
In the development cohort, we evaluated the correlation of candidate variables with OS using univariate Cox regression analysis (Table 1), and the variables with p value < 0.10 were subjected to the LASSO analysis. Finally, we picked up four variables to construct a new prognostic model in the way of nomogram (Figure 1 and Supplemental Table S2) by using the multivariate Cox proportional hazard model based on the β coefficients of predictive factors (Table 2). The nomogram consisted of β2-MG, HCT-CI, ALB, and PBPC, which comprised a new prognostic model called MM-BHAP model for predicting the probability of 6-month, 1-year, 2-year, and 4-year OS in NDMM patients. Also, we constructed two additional models individually including only ISS stage or R-ISS stage, comparing the predictivity of the MM-BHAP model with them in each imputed dataset (we just showed the results in one of the five imputed datasets in the text due to their similarity).

In the development cohort, the 50-sample bootstrapped calibration curve, with 1,000 bootstrap resamples, was used to examine the calibration of the MM-BHAP model. Table 3 summarizes other evaluations of models. The predictive OS probabilities were basically in accordance with those observed in 2-year OS (Figure 2A). As for discrimination, both time-dependent AUC and C-index of MM-BHAP model were globally higher than that of ISS stage and R-ISS stage (Figures 2B, C and Table 3). We also assessed clinical effect by DCA for 2-year OS (Figure 2D). The MM-BHAP model could achieve positive net benefit over a wider range of risk threshold, with higher area under the decision curve analysis (AUDC) than ISS stage and R-ISS stage in most time-points (Table 3). Moreover, the results of calibration curve and DCA for 6-month, 1-year, and 4-year OS are shown in Supplemental Figure S3.

Besides, we evaluated the IDI and continuous NRI to test the improvement in the prediction efficiency of the MM-BHAP model (Table 3). Compared to ISS stage, the MM-BHAP model showed the statistical improvement of predicting 2-year OS (6.3%, p = 0.046, Figure 2E) and had a tendency to perform better for the prediction of 6-month and 1-year OS (p = 0.088 and 0.080, respectively, Supplemental Figures S4A, C), but it was not statistically different for predicting 4-year OS (Supplemental Figure S4E) according to the IDI values. Moreover, there was no statistical difference in the prediction of OS between the MM-BHAP model and ISS stage according to the continuous NRI values. Moreover, the prediction efficiency of the MM-BHAP model was comparable to R-ISS stage, with no statistical difference according to the values of IDI and continuous NRI (Figure 2F and Supplemental Figures S4B, D, F).

In addition, the MM-BHAP model had great calibration for predicting PFS (Supplemental Figure S5A). Moreover, it also had globally higher AUC and C-index (Supplemental Figures S5B, C), as well as clinical usefulness (Supplemental Figure S5D) than that of ISS stage and R-ISS stage in each imputed dataset. Also, there was no statistical difference between the MM-BHAP model and ISS stage/R-ISS stage according to IDI and continuous NRI analyses (Supplemental Figures S5E, F).

Construction of MM-BHAP Stratification
We calculated the total point for each patient according to the nomogram (Figure 1 and Table S2) and divided the patients into low-, medium-, and high-risk subgroups according to the optimal cutoff points calculated by the X-Tile program (32). MM-BHAP stratification stage I was defined as the point ≤110; stage II was defined as the point from 110 to 248; stage III was defined as the point ≥248.

In the development cohort (N = 233), the patients were distributed across the three stages of the MM-BHAP stratification as follows (Table 4 and Figure 3A): stage I (38.6%), stage II (45.5%), and stage III (15.9%), with median overall survivals that were not reached (NR), 50.1 months and 26.2 months, respectively.
contrast, both ISS stage (Figure 3B) and R-ISS stage (Figure 3C) were not satisfactory in stratifying the patients between stages I and II in all five imputed datasets. Previous studies also indicated the prognostic value of ISS stage II (33) and R-ISS stage II (34) required further improvement.

The median OS of stage III of the MM-BHAP stratification was shorter than that of ISS stage and R-ISS stage (26.2, 46.1, and 33.4 months, respectively), suggesting that the MM-BHAP stratification performed better in identifying a specific group of high-risk patients. We also assessed the stratification for PFS (Table 4), indicating that both the MM-BHAP stratification and R-ISS stage had good prognostic stratification (Figures 3D, F), while ISS stage was not satisfactory in stratifying patients between stage I and stage II (Figure 3E) in all the five imputed datasets.

Subgroup Analyses

Next, we explored the performance of MM-BHAP stratification in specific groups of patients with age ≤65 or age >65 years and patients with or without HRCA in predicting OS. Both age and HRCA were important prognostic factors, but the MM-BHAP stratification still applied to the four different subgroups (Figures 4A-D). Then we further analyzed the applicability of our model in the subgroup of patients with at least two HRCAs recommended by mSMART 3.0, and the result indicated the applicability of MM-BHAP stratification in the double-hit or triple-hit myeloma patients (Figure 4E).

We also examined the distribution and co-occurrence of the MM-BHAP stratification, R-ISS stage, and ISS stage in the development cohort (N = 233) (Figure 5A). There were 37 patients among MM-BHAP stage III, all of whom were distributed in ISS stage III simultaneously. Among them, 22 patients existed in R-ISS stage III and 15 patients had R-ISS stage II.

Notably, there was a substantial portion of the patients with R-ISS stage II or ISS stage III, indicating that the two groups of patients needed an accurate stratification. The patients with ISS stage III (Figure 5B) could be further divided into three groups by our MM-BHAP stratification. Moreover, in the patients with R-ISS stage II, our model identified a group of patients with favorable outcomes (Figure 5C).

External Validation of the MM-BHAP Model

We calculated the total point for each patient of the validation cohort according to the abovementioned nomogram, as a factor for subsequent analyses (35). In the validation cohort, the calibration curve indicated an optimal agreement between the prediction and actual observation for the probability of OS (Figure 6A and Supplemental Figure S6).

Table 5 presents other comprehensive evaluations of the MM-BHAP model and ISS stage. Both time-dependent AUC and C-index of the MM-BHAP model were higher than those of the ISS stage, showing a greater prediction performance compared to the ISS stage (Figures 6B, C). In the analysis of DCA, the MM-BHAP model had higher net benefits among wider risk thresholds than that of the ISS stage at all time-points (Figure 6D and Supplemental Figure S7). Surprisingly enough, the MM-BHAP model had a remarkable
TABLE 3 | Comprehensive evaluations of the different models in the development cohort.

|                          | OS             | 6 months | 12 months | 24 months | 48 months |
|--------------------------|----------------|----------|-----------|-----------|-----------|
|                          | AUC, n (95% CI)|          |           |           |           |
| MM-BHAP                  | 0.793 (0.691–0.895) | 0.781 (0.684–0.878) | 0.789 (0.702–0.875) | 0.721 (0.666–0.875) |
| ISS stage                | 0.720 (0.626–0.815) | 0.749 (0.666–0.813) | 0.724 (0.644–0.803) | 0.771 (0.647–0.896) |
| R-ISS stage              | 0.790 (0.590–0.871) | 0.755 (0.655–0.856) | 0.728 (0.696–0.886) | 0.744 (0.527–0.746) |
| C-index, n               |               |          |           |           |           |
| MM-BHAP                  | 0.788          | 0.760    | 0.724     | 0.667     |
| ISS stage                | 0.720          | 0.732    | 0.680     | 0.670     |
| R-ISS stage              | 0.714          | 0.736    | 0.722     | 0.600     |
| Range*, n (%)            |               |          |           |           |           |
| MM-BHAP                  | 1.25%–26.64%   | 2.28%–38.16% | 3.78%–61.25% | 11.51%–95.07% |
| ISS stage                | 2.06%–8.99%    | 3.69%–15.74% | 5.88%–24.12% | 16.42%–55.84% |
| R-ISS stage              | 2.68%–13.29%   | 4.84%–22.97% | 8.01%–35.53% | 25.23%–78.31% |
| AUDC, n                  |               |          |           |           |           |
| MM-BHAP                  | 0.0027         | 0.0094   | 0.0195    | 0.0757    |
| ISS stage                | 0.0015         | 0.0052   | 0.0106    | 0.0472    |
| R-ISS stage              | 0.0018         | 0.0058   | 0.0144    | 0.0922    |
| IDI, n (95% CI), p value |               |          |           |           |           |
| Vs. ISS stage            | 2.4% (0.3%–14.8%) | 3.9% (-0.4%–19.5%) | 6.3% (1.0%–22.1%) | -1.8% (-23.1%–24.7%) |
| Vs. R-ISS stage          | 0.6% (-3.7%–11.3%) | 1.1% (-5.4%–16.4%) | 3.1% (-13.1%–15.3%) | 0% (-26.9%–34.5%) |
| Continuous NRI, n (95% CI), p value |               |          |           |           |           |
| Vs. ISS stage            | 1.2% (-14.9%–48.6%) | -2.1% (-16.1%–41.9%) | 12.4% (-4.4%–50.2%) | -21.1% (-82.1%–70.9%) |
| Vs. R-ISS stage          | 5.9% (-29.7%–43.9%) | 12.1% (-23.7%–44.2%) | 3.3% (-34.8%–42.1%) | 22.7% (-68.5%–89.4%) |

FIGURE 2 | The performance of the MM-BHAP model, ISS stage, and R-ISS stage for predicting OS in the development cohort. (A) The calibration curve of the MM-BHAP model for predicting 2-year OS. (B) The time-dependent AUC of the ROC in the three models. (C) The time-dependent Harrell's C-index in the three models. (D) The DCA was used to estimate clinical usefulness of the three models for predicting 2-year OS. The improvement in prediction of the MM-BHAP model was compared to the ISS stage (E) or R-ISS stage (F). The IDI was the value of the difference in area between red and blue zones (E, F). The continuous NRI was the value of the distance between two black dots (E, F).
improvement compared to the ISS stage according to the IDI values (range from 4.8% to 11.8%) and continuous NRI values (range from 30.2% to 56.0%) in predicting OS at different time points (Figure 6E, Supplemental Figure S8 and Table 5). Also, the MM-BHAP model had great calibration for predicting PFS (Supplemental Figure S9A). At the same time, it performed better in all the analyses of time-dependent AUC and C-index, DCA, IDI, and continuous NRI compared to the ISS stage (Supplemental Figures S9B–E).

The MM-BHAP stratification also categorized patients into three groups well as follows (Figure 7A): stage I (38.8%), stage II (40.8%), and stage III (20.4%), with median overall survivals of 71, 44.9, and 22.8 months, respectively. Also, our MM-BHAP stratification had a good stratification for PFS (Figure 7B). However, the ISS stage did not perform well in stratifying patients between three stages whether to predict OS ($p = 0.18$) or PFS ($p = 0.11$) (Figures 7C, D). Actually, the curves of the ISS stage were superimposable when time was more than approximately 60 months for OS (Figure 7C) and when time was more than about 50 months for PFS (Figure 7D), which suggested it was not enough for predicting long-term survival to merely utilize the ISS stage.

In the validation cohort, we did not draw comparison of the predictive capacity between the MM-BHAP model and R-ISS stage. This is because the patients were diagnosed between January 2013 and December 2015 in which our center did not completely perform FISH analysis with CD138-purified plasma cells, resulting in the problem taking no unified thresholds to identify HRCA.

**Distribution of Transplant-Eligible Patients**
Given that the new stratification included the HCT-CI that was used to assess transplant eligibility, we exploredly analyzed the association between the MM-BHAP stratification at diagnosis

| Stage | MM-BHAP stratification | Median OS (months) | Median PFS (months) |
|-------|-------------------------|--------------------|---------------------|
| I     | NR                      | 29.8               |                     |
| II    | 50.1                    | 22.3               |                     |
| III   | 26.2                    | 10.9               |                     |

| Stage | ISS stage | Median OS (months) | Median PFS (months) |
|-------|-----------|--------------------|---------------------|
| I     | NR        | 42.8               |                     |
| II    | NR        | 29.9               |                     |
| III   | 46.1      | 14.0               |                     |

| Stage | R-ISS stage | Median OS (months) | Median PFS (months) |
|-------|-------------|--------------------|---------------------|
| I     | NR          | 35.8               |                     |
| II    | NR          | 24.5               |                     |
| III   | 33.4        | 11.9               |                     |

**FIGURE 3** | Kaplan–Meier survival curves in the development cohort. OS of MM patients was stratified by the MM-BHAP stratification (A), ISS stage (B), and R-ISS stage (C). PFS was also classified by the three risk stratifications (D–F).
with the probability of receiving ASCT afterward. In the whole cohort (N = 385), 47 (12.2%) patients underwent ASCT, which were distributed across three stages of the MM-BHAP stratification as follows: stage I (25/149; 16.8%), stage II (18/168; 10.7%), and stage III (4/68; 5.9%). With multiple-comparison analysis by using the Bonferroni method, we found that the proportions between stage I (16.8%) and stage III (5.9%) had statistical difference (p < 0.05), suggesting the patients with stage III had less probability to fulfill the eligibility for transplantation in the future compared to patients with stage I.

**DISCUSSION**

In this study, we developed and validated a new simple prognostic model (MM-BHAP) integrating comorbidities and tumor burden, for predicting the probability of 6-month, 1-year, 2-year, and 4-year OS of NDMM patients. The novel MM-BHAP model performed well in terms of calibration, discrimination, clinical usefulness, and improvement in prediction, suggesting a good prognostic value for OS and PFS of NDMM patients. Moreover, the performance of the MM-BHAP model was superior to the ISS stage in both development and validation cohorts, while it was partially better at least not worse than the R-ISS stage in the development cohort. Notably, the MM-BHAP stratification categorized patients into three subgroups with clearly different OS or PFS, which was superior to the ISS stage and partially better than the R-ISS stage. Also, it identified a group of high-risk patients with shorter median OS (26.2 months) than that of the ISS stage (46.1 months) and R-ISS stage (33.4 months), indicating an advantage of defining truly high-risk patients in the real-world population. Furthermore, it enhanced the differential power of the ISS stage and R-ISS stage, with reclassifications in patients with ISS stage III or R-ISS stage II.

The MM-BHAP model was composed of four widely accessible factors of β2-MG, HCT-CI, ALB, and PBPC. β2-MG and ALB were typically prognostic factors integrated into the ISS stage (5). In contrast, we divided the value of ALB into three levels (≤24.5; >24.5 and ≤35; >35 g/l), based on the restricted cubic splines. With regard to PBPC, its importance was second only to HCT-CI in the MM-BHAP model. Multiple studies demonstrated that high levels of circulating plasma cells (CPCs) in MM patients were associated with worse prognosis (36–40) and even the prognosis of the MM patients with ≥5% CPCs was equivalent to that of the patients with plasma cell leukemia (41). In addition, it was evidenced that HCT-CI was independently associated with poor OS of MM patients undergoing HSCT (26, 27). It was of concern that HCT-CI
FIGURE 5 | Subgroup analyses about the three risk stratiﬁcations in the development cohort. (A) The distribution and co-occurrence of the patients respectively classiﬁed by the MM-BHAP stratiﬁcation, R-ISS stage, and ISS stage in the development cohort were displayed. Dots with connected lines represented that the patients coexisted in corresponding different subgroups and the vertical bar graphs reﬂected the number of these patients. Also, the blue, green, and orange dots respectively represented the co-occurrence of the patients classiﬁed by stage I, stage II, and stage II of the MM-BHAP stratiﬁcation with other subgroups. (B) OS curves in the subgroup with ISS stage III stratiﬁed by the MM-BHAP model. (C) OS curves in the subgroup with R-ISS stage II stratiﬁed by the MM-BHAP model.

FIGURE 6 | The performance of the MM-BHAP model and ISS stage for predicting OS in the validation cohort. (A) The calibration curve of the MM-BHAP model for predicting 2-year OS. (B) The time-dependent AUC of the ROC in the two models. (C) The time-dependent Harrell’s C-index in the two models. (D) The DCA for predicting 2-year OS. (E) The IDI and continuous NRI of the MM-BHAP model compared to the ISS stage.
also had potential to predict outcomes of the patients with newly diagnosed hematologic malignancies, not merely the patients after HSCT. A multicenter study firstly developed and validated a prognostic model incorporating HCT-CI at diagnosis, to estimate risks of mortality in acute myeloid leukemia (42). Nevertheless, little is known about the prognostic impact of HCT-CI at the time of diagnosis on the outcomes of NDMM patients. Herein, we systematically examined the prognostic value of HCT-CI at diagnosis which contributed the most to our MM-BHAP model. To our knowledge, the present study is the first to develop and validate a novel prognostic model integrating comorbidities measured by HCT-CI and tumor burden reflected by β2-MG, ALB, and PBPC.

Although HRCA was not selected into our model, it had robust prognostic implications. However, the application of novel agents and ASCT seemed to improve the poor outcomes of the patients with certain HRCA. For instance, the adverse impact of t(4;14) could be partly overcome by bortezomib (43) and the inferior outcome of del(17p) could be improved by maintenance therapy and ASCT especially in transplant eligible patients (44, 45), suggesting highly heterogeneous outcomes of the patients with HRCA. Therefore, there was an urgent need to further identify the ultra-high-risk subgroup within the patients harboring HRCA. In our cohort, the overall survival of the patients with HRCA or double/triple-hit HRCA could be further differentiated by our MM-BHAP stratification, maybe partially because of diverse therapy regimens and treatment intensity resulting from the burden of different comorbidities.

It was noteworthy that the R-ISS stage was inclined to present a better prognostic stratification for PFS compared to OS. As we all know, the R-ISS stage was completely composed of disease-related characteristics of myeloma (46–49). However, the overall survival of MM was dependent on not only the disease progression but also the host features such as comorbidities, performance status, treatment intention, and socioeconomic support, which were more likely to affect the drug accessibility and treatment integrity. Moreover, the R-ISS stage was developed

| TABLE 5 | Comprehensive evaluations of the different models in the validation cohort. |
|----------|-------------------|-------------------|-------------------|
|          | 6 months          | 12 months         | 24 months         |
| AUC, n (95% CI) | 0.739 (0.649–0.829) | 0.766 (0.675–0.856) | 0.735 (0.623–0.847) |
| MM-BHAP  | 0.606 (0.512–0.703) | 0.646 (0.546–0.748) | 0.611 (0.477–0.744) |
| ISS stage | 0.674             | 0.677             | 0.661             |
| C-index, n | 0.569             | 0.574             | 0.565             |
| MM-BHAP  | 26.62%–96.14%     | 33.15%–98.04%     | 38.98%–99.20%     |
| ISS stage | 38.58%–58.06%     | 45.87%–66.52%     | 52.25%–73.23%     |
| AUDC, n  | 0.1070            | 0.1290            | 0.1202            |
| MM-BHAP  | 0.0235            | 0.0308            | 0.0290            |
| ISS stage | 11.7% (3.6%–18.7%) | 11.8% (3.3%–18.7%) | 10.6% (1.9%–16.9%) |
| IDI, n (95% CI), p value | 0.012             | 0.020             | 0.052             |
| Continuous NRI, n (95% CI), p value | 0.012             | 0.020             | 0.052             |
in the highly selected cohorts within clinical trials (6) in which the patients with comorbidities were not underrepresented. Therefore, the R-ISS stage might be more applicable to the patients defined as “fit,” and it was essential to combine disease-related and patient-related prognostic factors in predicting overall survival in the real-world unselected population.

It was well known that the intermediate-risk group of the R-ISS stage was an exclusionary definition, which was an indication of the high heterogeneity of genetic background and outcomes. Further subdivision may redefine the outcomes of some patients, thereby achieving precise and personalized management. In our study, about 60% of the patients were among R-ISS stage II, which was similar to a previous R-ISS study (62%) (6). Moreover, among these patients, our model could identify a subgroup of patients with favorable outcomes. Furthermore, the MM-BHAP stratification could distinguish a group of patients with higher risk in the subgroup with ISS stage III, which further verified its applicability in some specific subsets.

Yet, there were still some limitations in our study. This study only included patients from a single center and needed further confirmations from larger multicenter cohorts. Moreover, this is a real-world retrospective cohort study, showing some missing data inevitably, but we handled missing data by using MICE to minimize the impact of data deficiency. HRCA was a known prognostic factor, and our subgroup analysis also suggested that patients with HRCA had poor median OS. However, it was finally not selected into our model, maybe because the impact was undervalued due to data deficiency, or it was knocked out since other variables contribute more in LASSO analysis. Besides, the comparison between MM-BHAP model and R-ISS stage was established only in the development cohort, requiring further validation. However, the effect of missing value on our model was minor because the finally selected variables included in the nomogram were all complete data in both development cohort and validation cohort. Moreover, the MM-BHAP stratification could definitely categorize patients into different groups with distinct outcomes.

In conclusion, our study was the first to use HCT-CI at diagnosis to predict the outcome of NDMM patients. Utilizing widely available prognostic factors, we constructed a novel simple MM-BHAP stratification combining tumor burden with comorbidities for better differentiation of real-world unselected patients with NDMM. It was expected to have a more accurate prediction for outcome unifying the MM-BHAP stratification and the current risk stratifications.

**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.
The study was reviewed and approved by the Ethics Committee of Xijing Hospital of Air Force Medical University.

GG, HT, SJ, and LB designed and conducted the research. JF, LX, TZ, HG, LY, QB, and RL enrolled the patients. SJ, LB, YC, XL, BT, and YG collected the data. SJ and HT analyzed the data and wrote the paper. GG and HT reviewed the manuscript. All authors contributed to the article and approved the submitted version.

This study was supported by the National Natural Science Foundation of China (No. 81907190 for GG; No. 82100218 for HT; No. 81900207 for LX), the Translation Research Grant of NCRCH (No. 2020ZKMC01 for GG), and the Key Research and Development Program in Shaanxi Province (No. 2019ZDLSF02-02 for GG).

The authors thank the support from all the medical and nursing staff, Department of Hematology, Xijing Hospital, Air Force Medical University, Xi’an, Shaanxi, China.

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/onc.2022.805702/full#supplementary-material
New Tool for Risk Assessment Before Allogeneic HCT. Blood (2005) 106:2912–9. doi: 10.1182/blood-2005-05-2004

23. Djebbi F, Pantzas F, Eyre TA, Prodder C, Davies F, Burton K, et al. Infection-Related Morbidity in a Large Study of Transplant non-Eligible Newly Diagnosed Myeloma Patients Treated With UK Standard of Care. Haematologica (2020) 105:e474–9. doi: 10.3324/haematol.2019.240762

24. Sorror ML, Logan BR, Zhu X, Rizzo JD, Cooke KR, McCarthy PL, et al. Prospective Validation of the Predictive Power of the Hematopoietic Cell Transplantation Comorbidity Index: A Center for International Blood and Marrow Transplant Research Study. Biol Blood Marrow Transplant (2015) 21:1479–87. doi: 10.1016/j.bbmt.2015.04.004

25. Thakar MS, Broglie L, Logan B, Artz A, Bunin N, Burroughs LM, et al. The Hematopoietic Cell Transplant Comorbidity Index Predicts Survival After Allogeneic Transplant for Nonmalignant Diseases. Blood (2019) 133:754–62. doi: 10.1182/blood-2018-09-876284

26. Saad A, Mahindra A, Zhang MJ, Zhong X, Costa LJ, Dispenzieri A, et al. Hematopoietic Cell Transplant Comorbidity Index is Predictive of Survival After Autologous Hematopoietic Transplantation in Multiple Myeloma. Biol Blood Marrow Transplant (2014) 20:402–8. doi: 10.1016/j.bbmt.2013.12.557

27. Obiozor C, Subramanian DP, Divine C, Shune L, Singh AK, Lin TL, et al. Evaluation of Performance Status and Hematopoietic Cell Transplantation Specific Comorbidity Index on Unplanned Admission Rates in Patients With Multiple Myeloma Undergoing Outpatient Autologous Stem Cell Transplantation. Blood (2017) 130:1641–5. doi: 10.1182/blood.2017.06.001

28. Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos MV, et al. International Myeloma Working Group Updated Criteria for the Diagnosis of Multiple Myeloma. Lancet Oncol (2014) 15:e538–48. doi: 10.1016/S1470-2045(14)70442-5

29. Ross FM, Aret-Loiseau H, Ameye G, Gutierrez NC, Liebisch P, Connor S, et al. Report From the European Myeloma Network on Interphase FISH in Multiple Myeloma and Related Disorders. Haematologica (2012) 97:1272–7. doi: 10.3324/haematol.2011.056176

30. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Software (2010) 33:1–22. doi: 10.18637/jss.v033.i01

31. Fitzgerald M, Saville BR, Lewis RJ. Decision Curve Analysis. JAMA (2015) 313:409–10. doi: 10.1001/jama.2015.37

32. Camp RL, Dolled-Filhart M, Rimm DL. X-Tile: A New Bio-Informatics Tool via Cation of Circulating Clonal Plasma Cells as a Marker for Drug Resistance and Short Survival in Multiple Myeloma: A Reappraisal. J Clin Oncol (2016) 91:E473–7. doi: 10.1002/ajh.24533

33. Fonseca R, Bergsagel PL, Drach J, Shaughnessy J, Gutierrez N, Stewart AK, et al. International Myeloma Working Group Molecular Classification of Multiple Myeloma: Spotlight Review. Leukemia (2009) 23:2212–21. doi: 10.1038/leu.2009.174

34. Cassuto JP, Krebs BP, Virot G, Dujardin P, Masseyeff R. Beta 2 Microglobulin, a Tumour Marker of Lymphoproliferative Disorders. Lancet (1978) 2:108–9. doi: 10.1016/S0140-6736(78)91428-9

35. Bataille R, Durie BG, Grenier J, Sany J. Prognostic Factors and Staging in Multiple Myeloma: A Reappraisal. J Clin Oncol (1986) 4:480–7. doi: 10.1200/JCO.1986.4.1.80

36. Dimopoulos MA, Barlogie B, Smith TL, Alexanian R. High Serum Lactate Dehydrogenase Level as a Marker for Drug Resistance and Short Survival in Multiple Myeloma. Ann Intern Med (1991) 115:931–5. doi: 10.1053/ajmo.2012.237

37. Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos MV, et al. International Myeloma Working Group Updated Criteria for the Diagnosis of Multiple Myeloma. Arch Intern Med (2013) 173:1330–7. doi: 10.1001/jama.2013.515

38. Gonsalves WI, Rajkumar SV, Dispenzieri A, Dingli D, Timm MM, Morice WG, et al. Quantification of Circulating Clonal Plasma Cells via Multiparametric Flow Cytometry Identifies Patients With Smoldering Multiple Myeloma at High Risk of Progression. Leukemia (2017) 31:130–5. doi: 10.1038/leu.2016.205

39. Rajkumar SV, Muzlar E, Kumar SK, Jevremovic D, Buadi FK, Dingli D, et al. Serial Measurements of Circulating Plasma Cells Before and After Induction Therapy Have an Independent Prognostic Impact in Patients With Multiple Myeloma Undergoing Upfront Autologous Transplantation. Haematologica (2017) 102:1439–45. doi: 10.3324/haematol.2017.166629

40. Bianchi G, Kyle RA, Larson DR, Witzig TE, Kumar S, Dispenzieri A, et al. High Levels of Peripheral Blood Circulating Plasma Cells as a Specific Risk Factor for Progression of Smoldering Multiple Myeloma. Leukemia (2013) 27:680–5. doi: 10.1038/leu.2012.237

41. Ravi P, Kumar SK, Roeker L, Gonsalves W, Buadi F, Lacy MQ, et al. Revised Diagnostic Criteria for Plasma Cell Leukemia: Results of a Mayo Clinic Study With Comparison of Outcomes to Multiple Myeloma. Blood Cancer J (2018) 8:116. doi: 10.1046/1410-0818-0140-1

42. Sorror ML, Storer BE, Fathi AT, Gerds AT, Medeiros BC, Shami P, et al. Development and Validation of a Novel Acute Myeloid Leukemia-Composite Model to Estimate Risks of Mortality. JAMA Oncol (2017) 3:1675–82. doi: 10.1001/jamaoncol.2017.2714

43. Aret-Loiseau H, Leelu 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) 28:4930–4. doi: 10.1200/JCO.2010.28.3945

44. Neben K, Lokhorst HM, Jauch A, Bertsch U, HielScher T, van der Holt B, et al. Administration of Bortezomib Before and After Autologous Stem Cell Transplantation Improves Outcome in Multiple Myeloma Patients With Deletion 17p. Blood (2012) 119:940–8. doi: 10.1182/blood-2011-09-379164

45. Merz M, HielScher T, Seckinger A, Hose D, Mai EK, Raab MS, et al. Baseline Characteristics, Chromosomal Alterations, and Treatment Affecting Prognosis of Deletion 17p in Newly Diagnosed Myeloma. Am J Hematol (2016) 91:E473–7. doi: 10.1002/ajh.24533

46. Neben K, Bertsch U, van der Holt B, HielScher T, van der Holt B, et al. Administration of Bortezomib Before and After Autologous Stem Cell Transplantation Improves Outcome in Multiple Myeloma Patients With Deletion 17p. Blood (2012) 119:940–8. doi: 10.1182/blood-2011-09-379164

47. Dimopoulos MA, Barlogie B, Smith TL, Alexanian R. High Serum Lactate Dehydrogenase Level as a Marker for Drug Resistance and Short Survival in Multiple Myeloma. Ann Intern Med (1991) 115:931–5. doi: 10.1053/ajmo.2012.237

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.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.