



LYMPHOID NEOPLASIA

# RNA-seq of newly diagnosed patients in the PADIMAC study leads to a bortezomib/lenalidomide decision signature

Michael A. Chapman,<sup>1,2</sup> Jonathan Sive,<sup>3</sup> John Ambrose,<sup>4</sup> Claire Roddie,<sup>5</sup> Nicholas Counsell,<sup>6</sup> Anna Lach,<sup>7</sup> Mahnaz Abbasian,<sup>7</sup> Rakesh Popat,<sup>5</sup> Jamie D. Cavenagh,<sup>3</sup> Heather Oakervee,<sup>3</sup> Matthew J. Streetly,<sup>8</sup> Stephen Schey,<sup>9</sup> Mickey Koh,<sup>10</sup> Fenella Willis,<sup>10</sup> Andres E. Virchis,<sup>11</sup> Josephine Crowe,<sup>12</sup> Michael F. Quinn,<sup>13</sup> Gordon Cook,<sup>14</sup> Charles R. Crawley,<sup>2</sup> Guy Pratt,<sup>15</sup> Mark Cook,<sup>15</sup> Nivette Braganza,<sup>6</sup> Toyin Adedayo,<sup>6</sup> Paul Smith,<sup>6</sup> Laura Clifton-Hadley,<sup>6</sup> Roger G. Owen,<sup>16</sup> Pieter Sonneveld,<sup>17</sup> Jonathan J. Keats,<sup>18</sup> Javier Herrero,<sup>4</sup> and Kwee Yong<sup>7</sup>

<sup>1</sup>Department of Haematology, University of Cambridge, Cambridge, United Kingdom; <sup>2</sup>Department of Haematology, Addenbrookes Hospital, Cambridge, United Kingdom; <sup>3</sup>Department of Haemato-oncology, St. Bartholomew's Hospital, Barts Health NHS Trust, London, United Kingdom; <sup>4</sup>Bill Lyons Informatics Centre, UCL Cancer Institute, University College London, London, United Kingdom; <sup>5</sup>Department of Haematology, University College London Hospitals, London, United Kingdom; <sup>6</sup>Cancer Research UK and UCL Cancer Trials Centre, London, United Kingdom; <sup>7</sup>Department of Haematology, UCL Cancer Institute, University College London, London, United Kingdom; <sup>8</sup>Department of Haematology, Guys and St. Thomas' Hospital, London, United Kingdom; <sup>9</sup>Department of Haematology, Kings College Hospital, London, United Kingdom; <sup>10</sup>Department of Haematology, St. George's Hospital, London, United Kingdom; <sup>11</sup>Department of Haematology, Royal Free London, Barnet and Chase Farm Hospitals, London, United Kingdom; <sup>12</sup>Department of Haematology, Royal United Hospitals Bath, Bath, United Kingdom; <sup>13</sup>Department of Haematology, Belfast City Hospital, Belfast, United Kingdom; <sup>14</sup>Department of Haematology, St. James's University Hospital, Leeds, United Kingdom; <sup>15</sup>Centre for Clinical Haematology, University Hospitals Birmingham, Birmingham, United Kingdom; <sup>16</sup>Haematological Malignancy Diagnostic Service, St. James' University Hospital, Leeds, United Kingdom; <sup>17</sup>Erasmus Medical Centre, Rotterdam, The Netherlands; and <sup>18</sup>Integrated Cancer Genomics Division, Translational Genomics Research Institute, Phoenix, AZ

KEY POINTS

- A 7-gene signature is derived that can identify myeloma patients who respond better to bortezomib- or lenalidomide-based therapy.
- Treatment according to the signature is noninferior to treatment with combined bortezomib, lenalidomide, and dexamethasone.

**Improving outcomes in multiple myeloma will involve not only development of new therapies but also better use of existing treatments. We performed RNA sequencing on samples from newly diagnosed patients enrolled in the phase 2 PADIMAC (Bortezomib, Adriamycin, and Dexamethasone Therapy for Previously Untreated Patients with Multiple Myeloma: Impact of Minimal Residual Disease in Patients with Deferred ASCT) study. Using synthetic annealing and the large margin nearest neighbor algorithm, we developed and trained a 7-gene signature to predict treatment outcome. We tested the signature in independent cohorts treated with bortezomib- and lenalidomide-based therapies. The signature was capable of distinguishing which patients would respond better to which regimen. In the CoMMpass data set, patients who were treated correctly according to the signature had a better progression-free survival (median, 20.1 months vs not reached; hazard ratio [HR], 0.40; confidence interval [CI], 0.23-0.72;  $P = .0012$ ) and overall survival (median, 30.7 months vs not reached; HR, 0.41; CI, 0.21-0.80;  $P = .0049$ ) than those who were not. Indeed, the outcome for these correctly treated patients was noninferior to that for those treated with combined bortezomib, lenalidomide, and dexamethasone, arguably the**

**standard of care in the United States but not widely available elsewhere. The small size of the signature will facilitate clinical translation, thus enabling more targeted drug regimens to be delivered in myeloma. (*Blood*. 2018;132(20):2154-2165)**

## Introduction

Multiple myeloma is a plasma-cell neoplasm characterized by lytic bone lesions, hypercalcemia, renal impairment, and bone marrow failure. Although outcomes have improved in recent years with the introduction of novel agents, the disease remains incurable, and clinical responses display considerable heterogeneity.<sup>1,2</sup> Additional improvements will come not only from introduction of new drugs but also from better use of existing drugs. Younger, fitter patients are usually treated with a drug combination involving a proteasome inhibitor (PI)

and/or an immunomodulatory drug (IMiD) followed by high-dose melphalan therapy with autologous stem-cell transplantation (ASCT). For transplantation-ineligible patients, recent trial data suggest that the treatment of choice may be a combination of the PI bortezomib (Velcade; Millennium Pharmaceuticals, Cambridge, MA), the IMiD lenalidomide (Revlimid; Celgene, Summit, NJ), and dexamethasone (VRD).<sup>3</sup> However, this combination is expensive and is not funded in most countries outside the United States. Furthermore, for frail patients, 3-drug combinations may prove too toxic.

It is possible that treatment outcomes in myeloma might be improved by the application of precision medicine (ie, the rational selection of drugs based on the biology of each patient's tumor). Several studies have demonstrated the potential of using transcriptomic data to derive prognostic information in myeloma.<sup>4-6</sup> Signatures can be usefully combined<sup>7</sup> but are generally agnostic to treatment,<sup>4,5,8-12</sup> and their main clinical utility is likely to be the identification of patients who may benefit from trials for high-risk disease. We sought to derive a signature that could predict responses to specific therapies.

The phase 2 PADIMAC (Bortezomib, Adriamycin, and Dexamethasone [PAD] Therapy for Previously Untreated Patients With Multiple Myeloma: Impact of Minimal Residual Disease in Patients With Deferred ASCT) study was designed to examine whether patients with good responses to PAD could safely avoid upfront ASCT. We employed RNA-sequencing (RNA-seq) on available good-quality RNA from enrolled patients and derived a training data set from patients with sustained deep responses in the absence of ASCT. We thus generated a signature for predicting bortezomib responsiveness in patients with myeloma not receiving ASCT. When tested in independent data sets, the signature performed well, identifying patients who benefited from bortezomib-based treatment in the absence of an IMiD. Furthermore, when tested on lenalidomide plus dexamethasone (RD)-treated patients, the signature performed in a reciprocal fashion, suggesting that it could be used as a binary classifier to choose between bortezomib-based treatment and RD. Patients who had been treated correctly according to the signature classification had a superior survival to those who had not. Indeed, in the related CoMMpass (Clinical Outcomes in Multiple Myeloma to Personal Assessment of Genetic Profile) data set, correctly treated patients receiving either bortezomib-based therapy (without IMiD) or receiving RD (without bortezomib) had a noninferior survival to those treated with VRD. This suggests that our signature could be employed to improve the safety and cost effectiveness of myeloma therapy without compromising outcomes.

## Materials and methods

### Sample accrual and processing and data generation

**Sample accrual and RNA isolation** RNA of sufficient quality for RNA-seq was available from 44 patients treated in the PADIMAC trial (ISRCTN03381785). The trial protocol is described in the supplemental Materials, available on the *Blood* Web site. PADIMAC was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines and was approved by the National Health Service National Research Ethics Service. Participants provided written informed consent. Patient registration and trial management were performed by the Cancer Research UK and University College London Cancer Trials Centre. All patients had newly diagnosed untreated myeloma and Eastern Cooperative Oncology Group performance status of 0 to 3 and were eligible for ASCT. Total RNA was isolated using standard methodology, as described in the supplemental Materials.

**Identification of mutations and gene expression** Standard methods were used to identify mutations and determine gene expression. Detailed methodology is described in the supplemental Materials and supplemental Table 1. Briefly, reads were mapped with TopHat<sup>13</sup> and aligned with Samtools.<sup>14</sup> Single-nucleotide variants (SNVs) and small indels were identified

using VarScan,<sup>15,16</sup> and RNA fusions were identified using FusionCatcher.<sup>17</sup> Read counts were generated with the Rsubread package.<sup>18,19</sup> Raw and count-level data have been uploaded to Gene Expression Omnibus (GEO; GEO reference GSE116324). Differentially expressed genes were identified using DESeq2,<sup>20-22</sup> and the Gage<sup>23</sup> and Pathview<sup>24,25</sup> packages were used for pathway analysis.

## Machine learning

**Selection of test data sets** Test RNA-seq data sets were derived from CoMMpass (<https://research.themmmf.org/>). Microarray test sets were obtained for relapsed/refractory patients treated with bortezomib<sup>26</sup> (GEO reference GSE9782), plasma-cell leukemia (PCL) patients treated with RD<sup>27</sup> (GEO reference GSE39925), and newly diagnosed myeloma patients treated with PAD followed by ASCT<sup>28</sup> (GEO reference GSE19784). We refer to these data as the Millennium, PCL, and HOVON/GMMG data sets, respectively.

**Data preprocessing, training, validation, and testing** RNA-seq counts were normalized and corrected for heteroscedasticity according to published methods.<sup>29-35</sup> Potential signature genes were identified from the PADIMAC data set by an empirical Bayes method,<sup>36</sup> then selected as described using synthetic annealing,<sup>37</sup> with an error rate determined by a support vector machine implemented from the e1071 package (<https://cran.r-project.org/web/packages/e1071/index.html>).

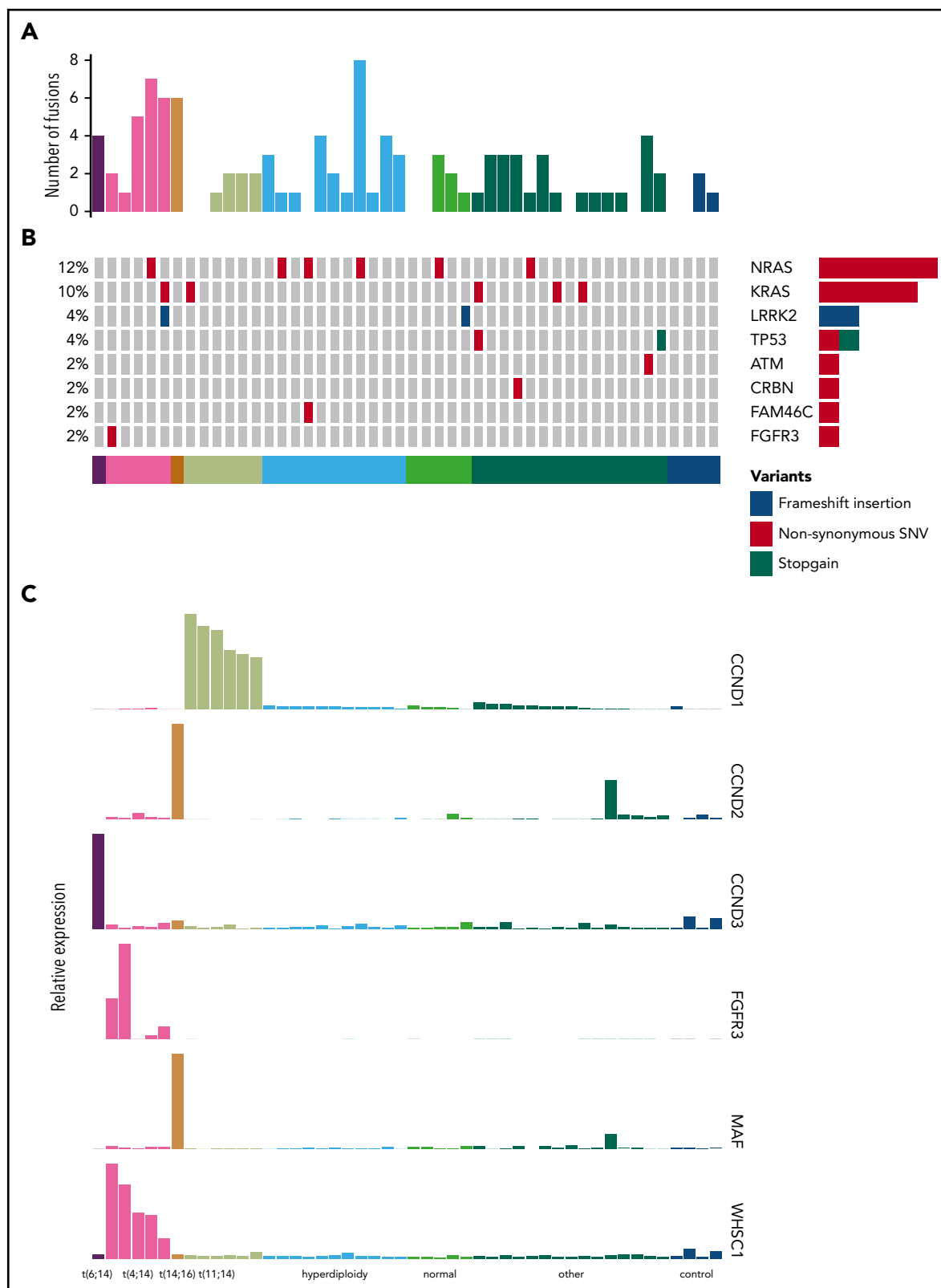
Signature assignments were made using the large margin nearest neighbor (LMNN) algorithm.<sup>38</sup> Performance within the PADIMAC data set was checked by 10-fold cross-validation. For external testing, all PADIMAC data were used for training, with an initial 50:50 split into training and internal validation sets fixed for all testing. R and Matlab scripts replicating this process have been included with the supplemental Materials. All CoMMpass, Millennium, PCL, or HOVON/GMMG data were used for testing. To determine the robustness of the signature performance in each case, a form of permutation testing was used, as described in the supplemental Materials.

**Statistical considerations** Null and observed assignments were compared using the Mann-Whitney-Wilcoxon test. Survival was compared using the Cox proportional hazards model. *P* values  $\leq .05$  were considered significant.

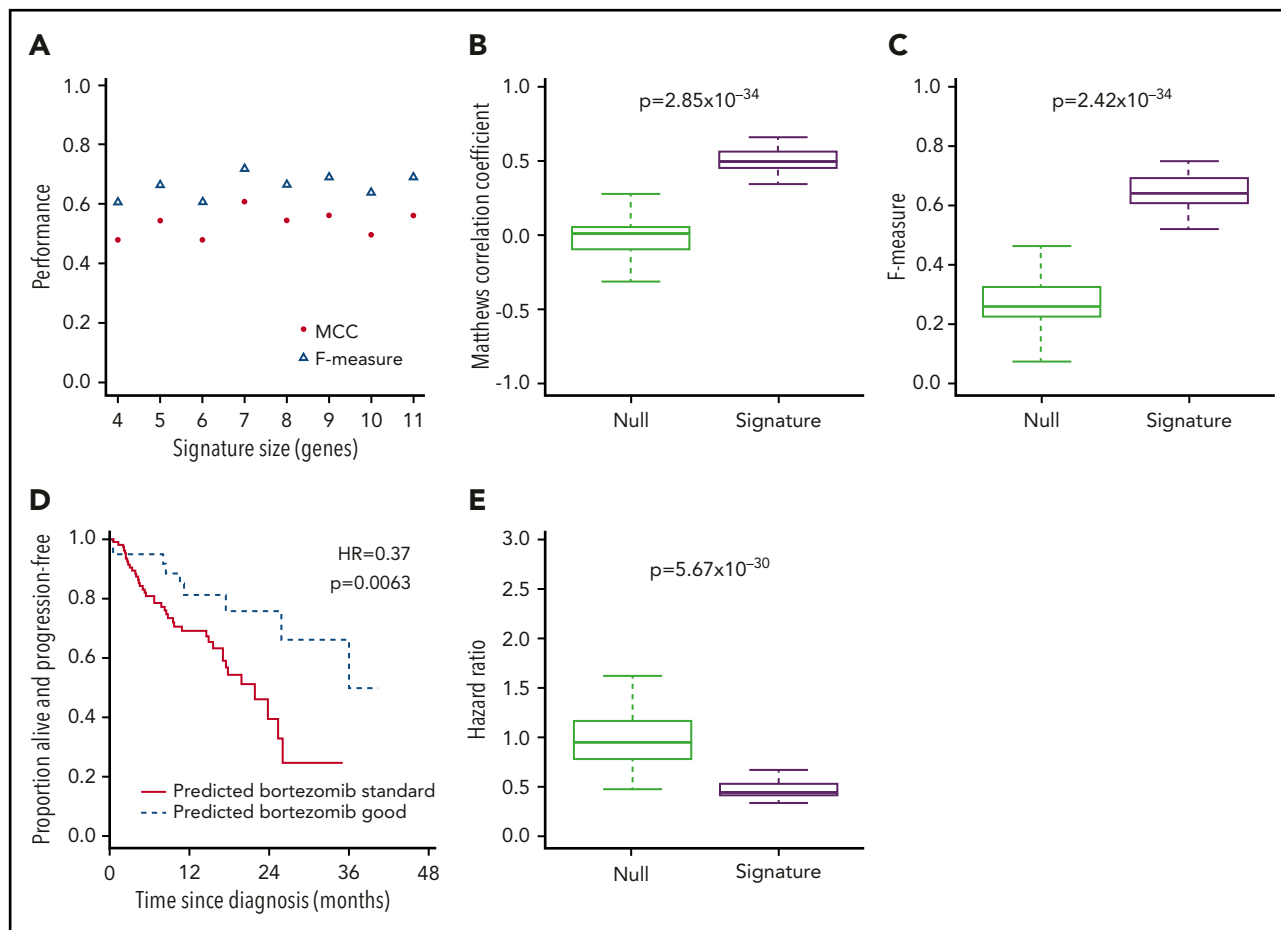
## Results

### Initial assessment of PADIMAC data excludes a mutation-based classifier

We performed RNA-seq on purified CD138<sup>+</sup> plasma cells from a cohort of 44 patients treated in the PADIMAC trial. Clinical data are shown in supplemental Table 2. We first explored the possibility of using a mutation-based classifier for bortezomib responsiveness. We identified fusion and SNV transcripts from the RNA-seq data (Figure 1A-B; supplemental Tables 3 and 4). There were 0 to 8 fusions in each sample, with a median of 1 (Figure 1A). Expected *IgH-WHSC1* fusions were detected from t(4;14) patients (supplemental Table 3). There was a median of 9 SNVs per patient in coding regions, which is lower than seen in previous DNA sequencing studies.<sup>39-42</sup> This may reflect reduced expression from mutant alleles<sup>43</sup> as well as a failure to detect mutations in the furthest 5' regions of some genes.



**Figure 1. Fusion events, key SNVs, and translocation/cyclin D (TC) classification of PADIMAC data.** Each column represents a single sample. Samples are arranged into and color coded by their TC classes (shown at the bottom). (A) Number of fusion events in each sample. (B) Key SNVs in each sample. (C) Expression of genes whose dysregulation is associated with TC classification in each sample.



**Figure 2. A 7-gene signature accurately predicts response to bortezomib-based therapy in PADIMAC and the independent CoMMpass data sets.** (A) MCCs and F measures of bortezomib-good assignments by 4 to 11 gene signatures derived from synthetic annealing after cross-validation within the PADIMAC data set. (B-C) MCCs and F measures of bortezomib-good assignments by the 7-gene signature after multiple ( $n = 100$ ) cross-validations within the PADIMAC data set (signature) compared with the MCCs and F measures of permuted assignments (null). The  $P$  values are those of the Wilcoxon-Mann-Whitney test, under the null hypothesis that the distributions of observed and null performances are the same. (D) Kaplan-Meier plot showing the progression-free survival (PFS) of patients who received bortezomib-based therapy within CoMMpass ( $n = 147$ ) and who were predicted to benefit ( $n = 39$ ; broken line) or not ( $n = 108$ ; solid line) from bortezomib-based therapy by the 7-gene signature after training in PADIMAC. The  $P$  value and hazard ratios (HRs) are those obtained from Cox regression analysis. (E) HRs for disease progression of bortezomib-good vs bortezomib-standard patients who received bortezomib-based therapy in CoMMpass. Predictions were made by the 7-gene signature, trained in PADIMAC, and followed repeated ( $n = 100$ ) training/validation splits (signature). The HRs are compared with a null data set of HRs obtained after permutations of the assignments (null). The  $P$  value is that of the Wilcoxon-Mann-Whitney test, under the null hypothesis that the distributions of observed and null performances are the same.

Nevertheless, we identified many of the known driver mutations in myeloma (Figure 1B; supplemental Table 4). Other known drivers were not mutated in our cohort, which may be related to sample size or may reflect lack of expression of mutated alleles.<sup>43</sup>

Overall, 45.5% of patients in the cohort achieved very good partial remission or better after PAD induction (supplemental Table 2). We defined a bortezomib-good group, namely patients who achieved very good partial remission or better and who were progression free at 1 year without ASCT (13 [29.5%] of 44). We termed the remaining patients bortezomib standard (31 [70.5%] of 44). There were no associations between these groups and age, International Staging System (ISS), or myeloma type (supplemental Table 5). We also saw no significant associations between bortezomib responsiveness and presence of key cytogenetic, SNV, or translocation events (supplemental Table 6). There were trends toward significant associations between the bortezomib-good group and both the presence of any translocation and the presence of a  $\beta 2$  microglobulin

translocation (supplemental Table 6). However, we did not feel that these associations were sufficiently strong for predicting clinical outcomes. We therefore turned to expression profiling.

### Derivation of a 7-gene bortezomib response signature

Expression of target genes known to be differentially expressed using microarray<sup>44</sup> and quantitative polymerase chain reaction<sup>45</sup> technologies in the translocation/cyclin D classification was consistent with that previously described (Figure 1C),<sup>44</sup> confirming the utility of RNA-seq for measuring relative gene expression in myeloma. We therefore proceeded to identify a gene signature for bortezomib responsiveness. We ranked potential genes using synthetic annealing<sup>37</sup> (supplemental Figures 1 and 2). Derived signatures comprising 4 to 11 genes performed better than permuted assignments in cross-validation of PADIMAC data using the LMNN algorithm (supplemental Figures 3 and 4; Matthews correlation coefficient [MCC]: median, 0.55 vs  $-0.045$ ; Mann-Whitney  $U$ , 0; 2-tailed  $P = .00090$ ; supplemental Figure 4A;

**Table 1. Cox regression results for survival comparisons**

Treatment (data set)	Comparison	n1	n2	Median survival 1, mo	Median survival 2, mo	HR* (CI)	P
Bortezomib based (C)	Bortezomib standard vs bortezomib good (PFS)	108	39	21.9	36.2	0.37 (0.17-0.81)	.0063
RD (C)	Bortezomib standard vs bortezomib good (PFS)	29	11	Not reached	18.8	2.35 (0.76-7.63)	.16
Bortezomib based or RD (C)	Incorrectly treated vs correctly treated (PFS)	119	68	20.1	Not reached	0.40 (0.23-0.72)	.0012
Bortezomib based or RD (C)	Incorrectly treated vs correctly treated (OS)	119	68	30.7	Not reached	0.41 (0.21-0.80)	.0049
Bortezomib based or RD or VRD (C)	All VRD vs all non-VRD (PFS)	208	187	43.7	26.0	1.50 (1.06-2.13)	.02
Bortezomib based or RD or VRD (C)	All VRD vs all non-VRD (OS)	208	187	Not reached	37.8	2.05 (1.34-3.13)	.00084
Bortezomib based or RD or VRD (C)	All VRD vs non-VRD correctly treated (OS)	208	68	Not reached	Not reached	1.12 (0.58-2.13)	.74
Bortezomib based or RD or VRD (C)	All VRD vs non-VRD correctly treated (PFS)	208	68	43.7	Not reached	0.86 (0.50-1.47)	.57
Bortezomib based ± ASCT	Bortezomib correctly treated vs ASCT (PFS)	47	39	36.6	36.2	1.16 (0.47-2.87)	.75
Bortezomib based ± ASCT	Bortezomib correctly treated vs ASCT (OS)	47	39	43.7	Not reached	3.10 (0.80-12.01)	.08
VRD (C)	Lenalidomide best vs bortezomib best (OS)	164	44	Not reached	Not reached	1.41 (0.69-2.90)	.36
VRD (C)	Lenalidomide best vs bortezomib best (PFS)	164	44	43.7	45.4	1.11 (0.62-1.97)	.74
Bortezomib alone (M)	Lenalidomide best vs bortezomib best (PFS)	135	38	4.14	4.77	0.66 (0.43-1.0)	.04
Bortezomib alone (M)	Lenalidomide best vs bortezomib best (OS)	148	40	15.2	21.8	0.57 (0.53-0.91)	.01
RD (PCL)	Lenalidomide best vs bortezomib best (PFS)	8	10	Not reached	1.0	Not defined†	$2.52 \times 10^{-5}$
RD (PCL)	Lenalidomide best vs bortezomib best (OS)	8	10	Not reached	12.5	Not defined†	.0013
PAD/ASCT (H)	Lenalidomide best vs bortezomib best (PFS)	84	57	26.6	31.8	0.91 (0.60-1.37)	.63
PAD/ASCT (H)	Lenalidomide best vs bortezomib best (OS)	85	58	Not reached	Not reached	0.73 (0.40-1.34)	.30

All signature assignments were based on the 7-gene signature. Correctly treated patients (rows 3, 4, and 7-10) were those predicted by the signature as lenalidomide best and who were treated with RD or those predicted as bortezomib best and treated with bortezomib-based therapy.

C, CoMMpass data set; CI, confidence interval; H, HOVON/GMMG data set; M, Millennium data set; n1, first group in comparison; n2, second group in comparison; OS, overall survival.

\*HR: second group vs first group.

†HR for the PCL group is not defined because all patients in group 2 experienced progression before any progressions in group 1 (row 11) or because there were no deaths in group 1 (row 12).

F measure: median, 0.67 vs 0.25; Mann-Whitney U, 0; 2-tailed  $P = .00090$ ; supplemental Figure 4B). Of these signatures, the best performing was the 7-gene signature (Figure 2A). Because training of the LMNN algorithm parameters involves splitting the training set into training and internal validation sets (supplemental Figure 2), we checked that the 7-gene signature was robust by performing multiple ( $n = 100$ ) training/validation splits

and comparing performance with permuted assignments ( $n = 100$ ) during cross-validation. The observed assignments had higher MCCs and F measures than the null assignments (MCC: median, 0.50 vs 0.0054; Mann-Whitney U, 8; 2-tailed  $P = 2.85 \times 10^{-34}$ ; Figure 2B; F measure: median, 0.64 vs 0.26; Mann-Whitney U, 3; 2-tailed  $P = 2.42 \times 10^{-34}$ ; Figure 2C), confirming that the signature performed well regardless of the training/validation split.

**Table 2. Mann-Whitney-Wilcoxon results testing the robustness of the 7-gene signature to assignments**

Treatment (data set)	Comparison	Median HR* observed	Median HR* null	Mann-Whitney U	P
Bortezomib based (C)	Bortezomib standard vs bortezomib good (PFS)	0.44	0.94	348	$5.67 \times 10^{-30}$
RD (C)	Bortezomib standard vs bortezomib good (PFS)	2.01	0.93	1526	$1.91 \times 10^{-17}$
Bortezomib based or RD (C)	Incorrectly treated vs correctly treated (PFS)	0.48	0.92	185	$5.44 \times 10^{-32}$
Bortezomib based or RD (C)	Incorrectly treated vs correctly treated (OS)	0.46	0.85	3	$2.52 \times 10^{-34}$
Bortezomib alone (M)	Lenalidomide best vs bortezomib best (PFS)	0.68	0.97	489	$2.75 \times 10^{-28}$
Bortezomib alone (M)	Lenalidomide best vs bortezomib best (OS)	0.53	1.00	108	$5.67 \times 10^{-33}$
RD (PCL)	Bortezomib best vs lenalidomide best (PFS)	0.24	0.83	1334	$2.84 \times 10^{-19}$
RD (PCL)	Bortezomib best vs lenalidomide best (OS)	0.20	0.96	1117	$1.98 \times 10^{-21}$

Mann-Whitney-Wilcoxon results testing the robustness of the 7-gene signature to assignments across multiple ( $n = 100$ ) training/validation data splits; 100 assignments were made in each data set by the 7-gene signature after random training/validation splits (observed). Each of these assignments was then permuted to maintain assignment ratios (null). The performance of the observed and null assignments for predicting PFS and OS was compared in terms of HRs by the Mann-Whitney-Wilcoxon test.

C, CoMMpass data set; H, HOVON/GMMG data set; M, Millennium data set.

\*HR: second group vs first group.

The genes comprising the signature are *EMC9*, *FAM171B*, *PLEK*, *MYO9B*, *RCN3*, *FLNB*, and *KIF1C* (supplemental Table 7). We did not see enrichment of these genes within the pathway gene sets from the Molecular Signatures Database<sup>46-48</sup> (<http://software.broadinstitute.org/gsea/msigdb/annotate.jsp>; C2 collection). However, at least 3 of the proteins (*EMC9*, *RCN3*, and *KIF1C*) are associated with the endoplasmic reticulum, and 3 others (*PLEK*, *MYO9B*, and *FLNB*) interact with actin filaments. Furthermore, 3 genes (*EMC9*, *MYO9B*, and *KIF1C*) associate positively with proliferation in myeloma.<sup>10</sup> Despite the lack of objective pathway enrichment in our signature, supervised analysis of the RNA-seq data as a whole did reveal pathways up-regulated in bortezomib-good patients (supplemental Table 8).

### The 7-gene signature is predictive of outcome in bortezomib-treated patients in the independent CoMMpass data set

To enable testing of our signature in an independent external data set, we extracted RNA-seq data from CoMMpass. We selected previously untreated patients who did not proceed to ASCT, because none of the bortezomib-good patients had undergone transplantation. There were 147 such bortezomib-treated patients (who had received no IMiD), 40 RD patients, and 208 VRD patients for whom RNA-seq data were available (supplemental Tables 9-11). There were a few differences in clinical features between the groups. PADIMAC patients, being transplantation eligible, were younger than all the CoMMpass cohorts (supplemental Figures 5-7). VRD-treated patients in CoMMpass were younger than the bortezomib-treated and RD cohorts (supplemental Figures 9 and 10). RD-treated patients had higher rates of del13 than PADIMAC or bortezomib-treated patients and a lower rate of t(4;14) than bortezomib-treated patients (supplemental Figures 6 and 8). There were

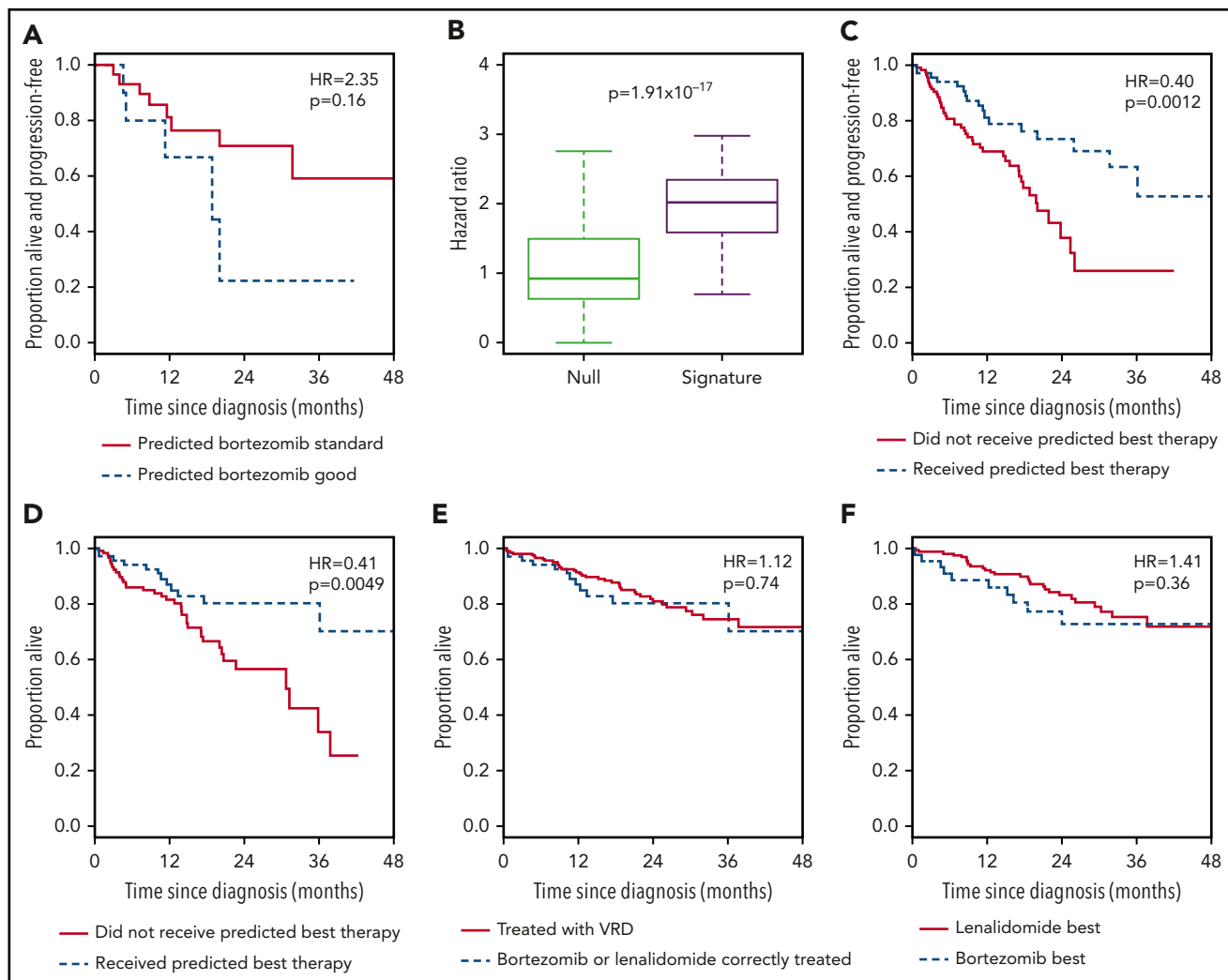
no differences in ISS stage between the groups (supplemental Figures 5-10).

We trained our 7-gene signature on the PADIMAC data and tested its ability to identify patients who would benefit from bortezomib-based therapy within CoMMpass (supplemental Figure 11). Patients who received bortezomib-based therapy and were assigned to the bortezomib-good group had a better PFS than those assigned to the bortezomib-standard group (Figure 2D; Table 1, row 1). The randomization seed for the PADIMAC training/validation split had been fixed before testing. To ensure that the predictive ability of the signature was robust, we performed multiple additional training/validation splits of the PADIMAC training set and compared the resulting assignments in the CoMMpass test set with permuted assignments that formed a null data set. As expected for a robust signature, HRs for the predicted bortezomib-good patients were lower than HRs from random predictions (Figure 2E; Table 2, row 1).

### The 7-gene signature has reciprocal performance in RD-treated patients and has the potential to select therapy

To distinguish between the signature acting as a general predictor of good-prognosis disease and as a specific predictor of bortezomib-sensitive disease, we tested it in RD-treated patients. We reasoned that, if the signature were bortezomib specific, the survival of RD-patients assigned to the bortezomib-good group would be no better than the survival of those assigned to the bortezomib-standard group. To our surprise, RD-treated patients assigned to the bortezomib-good group in fact had an inferior PFS to those assigned to the bortezomib-standard group (Figure 3A; Table 1, row 2). Although the difference was not significant, those assigned to the bortezomib-good





**Figure 3. The 7-gene signature can be used as a classifier to select between bortezomib-based therapy and RD in the independent CoMMpass data set.** (A) Kaplan-Meier plot showing the PFS of patients who received RD therapy ( $n = 40$ ) within CoMMpass and who were predicted to benefit ( $n = 11$ ; broken line) or not ( $n = 29$ ; solid line) from bortezomib-based therapy by the 7-gene signature after training in PADIMAC. The  $P$  value and HR are those obtained from Cox regression analysis. (B) HRs for disease progression of bortezomib-good vs bortezomib-standard patients who received RD in CoMMpass. Predictions were made by the 7-gene signature, trained in PADIMAC, and followed repeated ( $n = 100$ ) training/validation splits (signature). The HRs are compared with a null data set of HRs obtained after permutations of the assignments (null). The  $P$  value is that of the Wilcoxon-Mann-Whitney test, under the null hypothesis that the distributions of observed and null performances are the same. (C) Kaplan-Meier plot showing the PFS of patients who received bortezomib-based therapy or RD within CoMMpass ( $n = 187$ ) and who received the correct ( $n = 68$ ; broken line) or incorrect ( $n = 119$ ; solid line) therapy predicted by the 7-gene signature after training in PADIMAC. The  $P$  value and HR are those obtained from Cox regression analysis. (D) Kaplan-Meier plot showing the OS of patients who received bortezomib-based therapy or RD within CoMMpass ( $n = 187$ ) and who received the correct ( $n = 68$ ; broken line) or incorrect ( $n = 119$ ; solid line) therapy predicted by the 7-gene signature after training in PADIMAC. The  $P$  value and HR are those obtained from Cox regression analysis. (E) Kaplan-Meier plot showing the OS of patients ( $n = 276$ ) who received VRD ( $n = 208$ ; solid line) or who received bortezomib-based therapy or RD within CoMMpass and who received the correct therapy predicted by the 7-gene signature ( $n = 68$ ; broken line) after training in PADIMAC. The  $P$  value and HR are those obtained from Cox regression analysis. (F) Kaplan-Meier plot showing the OS of patients who received VRD in CoMMpass ( $n = 208$ ; solid line) and who were predicted to benefit from RD ( $n = 164$ ; solid line) or from bortezomib-based therapy ( $n = 44$ ; broken line) by the 7-gene signature after training in PADIMAC. The  $P$  value and HR are those obtained from Cox regression analysis.

group across repeated training/validation splits had consistently lower survival, with higher HRs than those obtained by permuting the assignments (Figure 3B; Table 2, row 2).

The implication of these findings is that patients predicted to do well with bortezomib by our signature do poorly when treated with RD and vice versa. Hence, the 7-gene signature could be used as a binary classifier to rationally choose between bortezomib-based therapy and RD. To test this, we selected CoMMpass patients treated with bortezomib-based therapy or with RD and assigned each to a bortezomib-best or lenalidomide-best group. We then compared survival between those patients who

received the predicted best treatment and those who did not. Patients who received the correct therapy had a superior PFS (Figure 3C; Table 1, row 3) and OS (Figure 3D; Table 1, row 4). The incorrectly treated patients had a median PFS of 20.1 months and a median OS of 31.2 months, whereas the median PFS and OS were not reached for correctly treated patients. These predictions were again robust to the initial training/validation split of the PADIMAC data set (supplemental Figure 12; Table 2, rows 3 and 4).

We excluded the possibility that the signature was acting as a surrogate for clinical features. We saw no association between

signature assignment and key cytogenetic events ( $P = .13$ ; Fisher's exact test; supplemental Table 12), and in multivariate Cox regression, ideal treatment according to the signature retained significance for survival when age, ISS, or myeloma subtype were taken into account (supplemental Table, 13 rows 1 and 2).

Because of the finding of proliferative genes in our signature, we also wanted to check that it was not acting as a surrogate for the gene expression-based proliferation index (GPI50) signature<sup>10</sup> or other prognostic signatures. Because these signatures have not, to our knowledge, been applied to RNA-seq data previously, we first tested that they could be applied to the CoMMpass data set. Indeed, when all of the GPI50, University of Arkansas for Medical Sciences (UAMS70), Erasmus University Medical Centre (EMC92), and Intergroupe Francophone du Myélome (IFM15) signatures<sup>4-6,10</sup> were applied to CoMMpass, the distribution of scores was similar to that seen in microarray data sets<sup>5</sup> (supplemental Figure 13A,C,E,G). Furthermore, all signatures retained prognostic significance using thresholds equivalent to those previously published<sup>5</sup> (supplemental Figure 13B, D,F,H). Having established that these prognostic signatures were effective in RNA-seq data, we examined whether there was any association between their assignments and the assignments of our 7-gene signature. None was seen (supplemental Table 14). Furthermore, receiving ideal treatment according to the 7-gene signature retained its prognostic significance even in a multivariate analysis including these signatures (supplemental Table 15).

A recent trial reported the superiority of VRD over RD in transplantation-ineligible patients,<sup>3</sup> but VRD treatment is not currently funded widely outside the United States. We wondered whether rationally selected therapy could be a cost-effective alternative to this gold-standard treatment. We first demonstrated that VRD was superior to unselected bortezomib-based treatment and RD in CoMMpass (supplemental Figure 14; Table 1, rows 5 and 6). We then compared the survival of patients treated correctly according to our signature with the survival of all patients treated with VRD in CoMMpass. There was no statistically significant difference in OS (Figure 3E; Table 1, row 7) or PFS (supplemental Figure 15A; Table 1, row 8) between patients treated correctly with bortezomib or RD and those treated with VRD. This was also true in a multivariate analysis incorporating clinical features (supplemental Table 13, rows 3 and 4). We also compared the outcomes of CoMMpass patients treated correctly with bortezomib according to the signature (without transplantation) with those of all patients receiving bortezomib-based induction followed by ASCT. Interestingly, there was no significant difference in survival between the 2 groups (supplemental Figure 16; Table 1, rows 9 and 10), although there was weak evidence of an effect implying longer OS with transplantation.

We hypothesized that the 7-gene signature should have minimal predictive ability in VRD-treated patients. As expected, we saw no difference between the outcomes for patients assigned to the bortezomib-best or lenalidomide-best group when those patients were treated with VRD. This was true both for OS (Figure 3F; Table 1, row 11) and for PFS (supplemental Figure 15B; Table 1, row 12). This lack of predictive ability was also seen in multivariate analyses incorporating clinical features (supplemental Table 13, rows 5 and 6).

## The 7-gene signature performs well in other independent data sets

We were keen to test how our signature would perform in other nontransplantation settings, such as relapsed disease. However, being limited by the availability of publicly available RNA-seq data, we turned to microarray data. There were 2 suitable data sets available. One comprised samples from patients with relapsed/refractory myeloma treated with single-agent bortezomib<sup>26</sup> (the Millennium data set). The second contained transcriptomic data from a small series of patients with PCL treated with RD<sup>27</sup> (the PCL data set). We reasoned that, within the Millennium data set, patients assigned to the bortezomib-best class should have a better survival, whereas within the PCL data set, those assigned to the lenalidomide-best class should have a superior outcome.

Signature assignments behaved as predicted. In the Millennium data set, the bortezomib-best group had a superior PFS (Figure 4A; Table 1, row 13) and OS (Figure 4B; Table 1, row 14) compared with the lenalidomide-best group. These results were robust to training/validation splits (Figure 4C; Table 2, rows 5 and 6). In the PCL data set, those predicted to be in the lenalidomide-best group had a superior PFS (Figure 4D; Table 1, row 15) and OS (Figure 4E; Table 1, row 16) compared with patients assigned to the bortezomib-best group. Again, the signature was robust, with little influence from the training/validation split (Figure 4F; Table 2, rows 7 and 8).

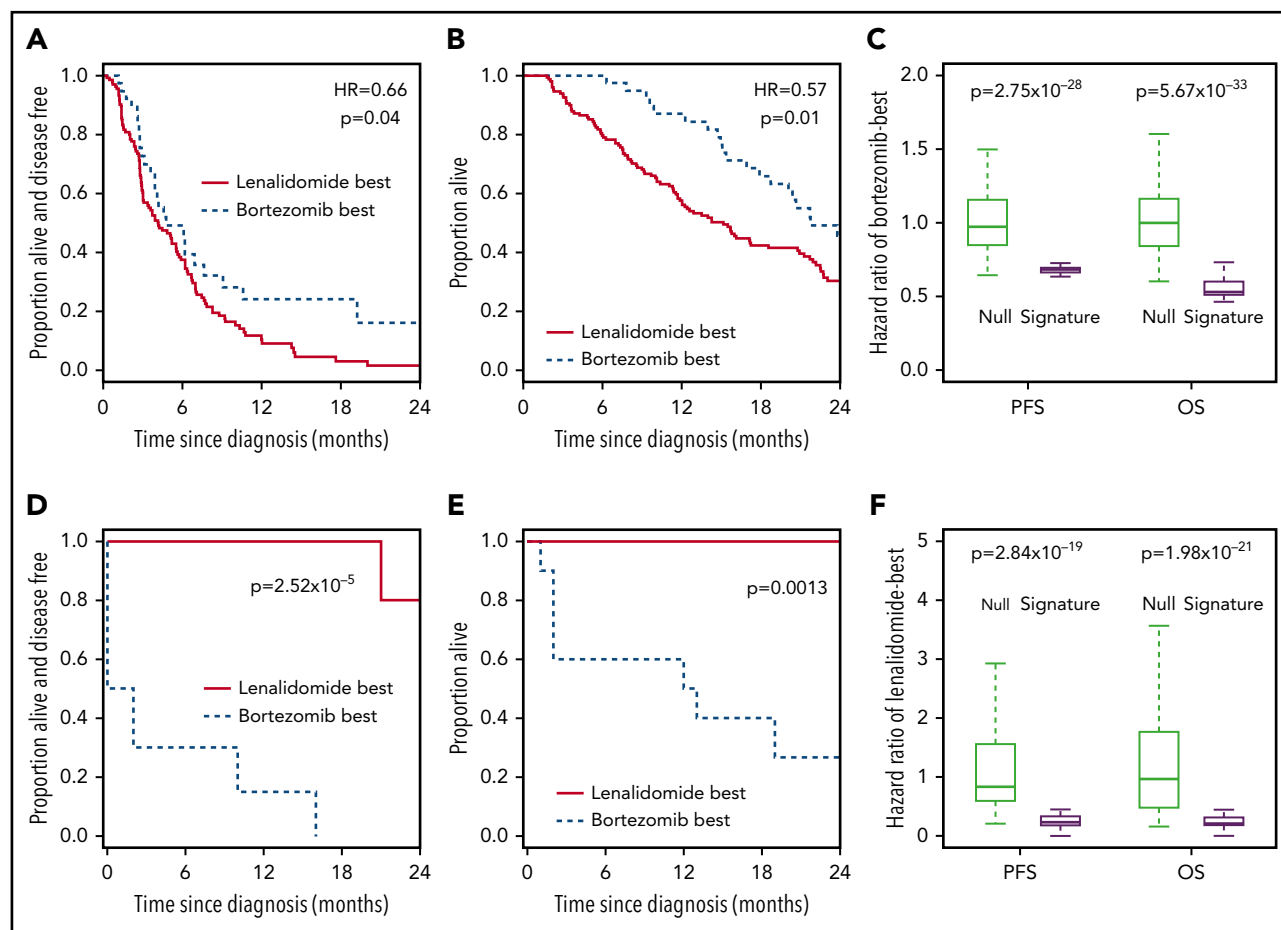
## The 7-gene signature loses predictive power in patients proceeding to ASCT

The bortezomib-good patients in the PADIMAC training set avoided ASCT because of their good response, according to trial protocol. We had thus far confined testing in external data sets to patients who had not undergone ASCT. We wondered whether the signature would retain its predictive power in patients proceeding to ASCT or whether transplantation would overcome the survival differences between correctly and incorrectly treated patients. The HOVON-64/GMMG-HD4 phase 3 trial<sup>49</sup> compared patients with newly diagnosed myeloma treated with conventional chemotherapy vs those treated with PAD. Both groups of patients proceeded to ASCT.

We used our signature to make bortezomib-best and lenalidomide-best assignments in patients who had received PAD. We reasoned that, if the signature retained its predictive power in the ASCT setting, we would see superior survival in those patients assigned to the bortezomib-best group. However, this was not the case, and we saw no significant difference in either PFS (Figure 5A; Table 1, row 17) or OS (Figure 5B; Table 1, row 18) between the different signature assignments.

As a further check for the specificity of the signature, we also tested its predictive value in the dexamethasone-only arm. As anticipated, there was no difference in PFS between those patients predicted to be bortezomib good and those predicted to be lenalidomide good (supplemental Figure 17A). However, patients predicted to be bortezomib good had a superior OS in this arm (supplemental Figure 17B). This is likely because patients receiving dexamethasone were eligible to receive crossover bortezomib upon disease progression.





**Figure 4. The 7-gene signature accurately predicts bortezomib or lenalidomide responsiveness in further independent data sets.** (A) Kaplan-Meier plot showing the PFS of patients who received single-agent bortezomib within the Millennium studies ( $n = 173$ ) and who were predicted to benefit from bortezomib-based therapy ( $n = 38$ ; broken line) or from RD therapy ( $n = 135$ ; solid line) by the 7-gene signature after training in PADIMAC. The  $P$  value and HR are those obtained from Cox regression analysis. (B) Kaplan-Meier plot showing the OS of patients who received single-agent bortezomib within the Millennium studies ( $n = 188$ ) and who were predicted to benefit from bortezomib-based therapy ( $n = 40$ ; broken line) or RD therapy ( $n = 148$ ; solid line) by the 7-gene signature after training in PADIMAC. The  $P$  value and HR are those obtained from Cox regression analysis. (C) HRs for PFS and OS of patients predicted to benefit from bortezomib-based therapy who received bortezomib in the Millennium studies. Predictions were made by the 7-gene signature, trained in PADIMAC with repeated ( $n = 100$ ) training/validation splits (signature). The HRs are compared with a null data set of HRs obtained after permutations of the assignments (null). The  $P$  values are those of the Wilcoxon-Mann-Whitney test, under the null hypothesis that the distributions of observed and null performances are the same. (D) Kaplan-Meier plot showing the PFS of patients who received RD within the PCL study ( $n = 18$ ) and who were predicted to benefit ( $n = 8$ ; solid line) or not ( $n = 10$ ; broken line) from RD therapy by the 7-gene signature after training in PADIMAC. The  $P$  value is that obtained from Cox regression analysis. (E) Kaplan-Meier plot showing the OS of patients who received RD within the PCL study ( $n = 18$ ) and who were predicted to benefit ( $n = 8$ ; solid line) or not ( $n = 10$ ; broken line) from RD therapy by the 7-gene signature after training in PADIMAC. The  $P$  value is that obtained from Cox regression analysis. (F) HRs for PFS and OS of patients predicted to benefit from lenalidomide-based therapy and who received RD in the PCL study. Predictions were made by the 7-gene signature, trained in PADIMAC with repeated ( $n = 100$ ) training/validation splits (signature). The HRs are compared with a null data set of HRs obtained after permutations of the assignments (null). The  $P$  value is that of the Wilcoxon-Mann-Whitney test, under the null hypothesis that the distributions of observed and null performances are the same.

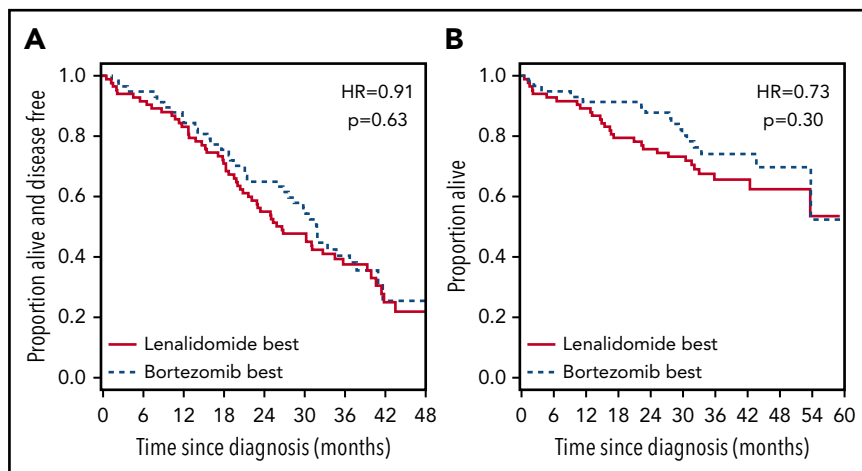
## Discussion

If it can be realized, precision cancer medicine will benefit patients in terms of improved efficacy and reduced toxicity and will benefit society in terms of better management of limited drug budgets. Transcriptomics has considerable promise in this area.<sup>50</sup> There are signatures that predict for overall prognosis in cancer,<sup>51-62</sup> including myeloma,<sup>4-12</sup> and signatures that predict response to individual therapies.<sup>63-67</sup> However, we are not aware of any published signature that can be used to rationally select between different active cancer therapies. Remarkable improvements in myeloma outcome over recent years have been seen thanks to the introduction of multiple novel agents, but this has been associated with increasing costs of treatment.<sup>68</sup> Therefore, precision medicine is arguably of particular importance

in this disease to help navigate through the increasing armamentarium of available therapies.

Here, we describe the derivation and testing of a 7-gene signature that can be used to select between bortezomib-based or RD therapy in myeloma patients not undergoing ASCT. Patients treated correctly according to the signature in the CoMMpass data set had a 69.7% 3-year OS, similar to the outcome of patients treated with VRD, probably the current standard of care for transplantation-ineligible patients.<sup>3</sup> These comparisons have to be viewed with caution, because CoMMpass is not a clinical trial. Nevertheless, there was no evidence that VRD-treated patients represented a poor-prognosis cohort in CoMMpass; there were no significant differences between the rates of poor-risk cytogenetics

**Figure 5. The 7-gene signature loses predictive ability in the transplantation setting.** (A) Kaplan-Meier plot showing the PFS of patients who received PAD within the HOVON-65/GMMG-HD4 study (n = 143) before ASCT and who were predicted to benefit from bortezomib-based therapy (n = 58; broken line) or from RD therapy (n = 85; solid line) by the 7-gene signature after training in PADIMAC. The P value and HR are those obtained from Cox regression analysis. (B) Kaplan-Meier plot showing the OS of patients who received PAD within the HOVON-65/GMMG-HD4 study (n = 143) before ASCT and who were predicted to benefit from bortezomib-based therapy (n = 58; broken line) or RD therapy (n = 85; solid line) by the 7-gene signature after training in PADIMAC. The P value and HR are those obtained from Cox regression analysis.



or ISS in the VRD group and the bortezomib- or RD-treated group, and noninferior survival was maintained in a multivariate model incorporating ISS.

If the outcome with rationally selected bortezomib- or lenalidomide-based therapy is equivalent to that with VRD, it would be important to consider why this might be. It may be that many patients treated with VRD are predominantly benefiting from just the bortezomib or the lenalidomide. Alternatively, it might be that any gains in combining the drugs are offset by increased toxicity, particularly in older or frailer patients. It is important to note that the reciprocal performance that we observed is an intrinsic property of the signature and not simply because bortezomib sensitivity is automatically associated with lenalidomide resistance (and vice versa). This is clearly not the case in clinical practice, nor is it consistent with the existence of multiple treatment-agnostic prognostic signatures in myeloma.

Although PADIMAC was a trial for transplantation-eligible patients, the signature was trained on patients who had had a good response in the absence of ASCT. Therefore, our initial test set comprised patients who were transplantation ineligible (there were no data sets from transplantation-eligible patients who did not proceed to transplantation). Transplantation-ineligible patients would be the most obvious to benefit after successful translation of the signature to the clinic. When tested in transplantation-eligible patients who had received PAD and ASCT in the HOVON-64/GMMG-HD4 trial, the signature lost its predictive ability, implying that transplantation can overcome the effect of not receiving the correct treatment. An interesting question would be whether receiving the correct predicted treatment without transplantation is equivalent to ASCT. We saw no survival difference between transplantation-ineligible patients treated correctly with bortezomib and all transplantation-eligible patients treated with bortezomib followed by ASCT. However, there was weak evidence of an effect suggesting better OS for patients who underwent transplantation.

There are limitations of our signature that need to be overcome before employment in a clinical trial. The signature assigns approximately one quarter to one third of patients to the bortezomib-best group and the remainder to the lenalidomide-

best group by default. It may be that this larger group is heterogeneous, with some patients having poorer prognosis or multiclonal disease and thus requiring VRD or the addition of other novel agents. Others may benefit equally from bortezomib- or lenalidomide-based treatments, regardless of signature assignment. Our signature was not capable of identifying these different groups (data not shown). Our external test cohorts were fairly small because of the lack of publicly available and appropriate test data sets, and prospective validation of our signature will be needed. This will probably require the development of a quantitative polymerase chain reaction or focused sequencing panel, although the small number of genes in the signature means that this should be feasible. Finally, it is not clear to what extent our signature represents a drug effect or a class effect, because of the lack of publicly available test data sets involving patients treated with alternative PIs or IMiDs. Future data may become available from clinical trials in which expression profiling is incorporated into the protocol.

We believe that our signature has the potential to move the myeloma field toward rational therapy decisions for transplantation-ineligible patients in the future. It is essential that myeloma genomic data sets with relevant clinical outcome data continue to be made publicly available to allow refinement and prospective validation of these approaches. This will require the ongoing support of the myeloma research community.

## Acknowledgments

The authors thank the patients and their physicians who participated in the PADIMAC and CoMMpass studies and are extremely grateful to the reviewers for their comments and suggestions, which have undoubtedly strengthened the paper.

The PADIMAC trial was funded by a Leukaemia and Lymphoma Research (now Bloodwise) grant (LRF/10018) and a University College London Biomedical Research Centres grant. It was managed by the Cancer Research UK and University College London Cancer Trials Centre. It was overseen by the National Cancer Research Institute Haemato-oncology Clinical Study Group. J.A. was funded by the Cancer Research UK—University College London Centre Award (C416/A18088 and C416/A25145). M.A.C. is a Wellcome-Beit Intermediate Fellow. PADIMAC was also funded in part by research funding from Janssen to the following authors: J.D.C., M.J.S., S.S., M.K., F.W., A.E.V., J.C., M.F.Q., G.C., C.R.C., G.P., R.G.O., and K.Y.

## Authorship

Contribution: M.A.C., J.S., J.A., C.R., R.P., J.J.K., J.H., and K.W. conceived the present study and performed data analysis; and all authors contributed to data collection and writing of the manuscript.

Conflict-of-interest disclosure: Alexion: honoraria (M.K.). Amgen: travel bursaries (G.C.), advisory role/consultancy (K.Y.), honoraria (R.P., G.P., K.Y.), speaker's bureau (G.C.), and research (F.W., G.C., L.C.-H., K.Y.). Autolus: advisory role/consultancy (K.Y.). Bayer: honoraria (J.C.). Celgene: travel bursaries (C.R., R.P., F.W., G.C., R.G.O.), advisory role/consultancy (R.G.O.), honoraria (G.P., R.G.O.), speaker's bureau (M.J.S., G.C., G.P.), and research (K.Y.). Chugai: advisory role/consultancy (K.Y.) and research (K.Y.). Gilead: travel bursaries (J.J.K.), honoraria (M.K., G.P.) and speaker's bureau (M.K., G.P.). Janssen: travel bursaries (G.C., G.P.), advisory role/consultancy (R.P., M.J.S., G.C., R.G.O., K.Y.), honoraria (M.A.C., R.P., M.F.Q., G.C., G.P., R.G.O., K.Y.), and speaker's bureau (M.J.S., G.C., G.P., K.Y.). Novartis: travel bursaries (A.E.V.). Pfizer: travel bursaries (J.C.), honoraria (A.E.V.), and research (L.C.-H.). Sanofi: travel bursaries (G.C.) and speaker's bureau (G.C.). Seattle Genetics: advisory role/consultancy (G.C.). Takeda: travel bursaries (M.A.C., G.C., G.P.), advisory role/consultancy (R.P., G.C., K.Y.), honoraria (M.A.C., R.P., R.G.O., G.C., G.P., K.Y.), speaker's bureau (G.C., G.P.), and research (G.C.). The remaining authors declare no competing financial interests.

ORCID profiles: M.A.C., 0000-0001-8342-0606; J.H., 0000-0001-7313-717X.

Correspondence: Michael A. Chapman, Department of Haematology, University of Cambridge, NHS Blood and Transplant, Long Rd, Cambridge CB2 0PT, United Kingdom; e-mail: mac54@cam.ac.uk.

## Footnotes

Submitted 9 May 2018; accepted 28 July 2018. Prepublished online as *Blood* First Edition paper, 4 September 2018; DOI 10.1182/blood-2018-05-849893.

The online version of this article contains a data supplement.

There is a *Blood* Commentary on this article in this issue.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

## REFERENCES

1. Avet-Loiseau H, Durie BG, Cavo M, et al; International Myeloma Working Group. Combining fluorescent in situ hybridization data with ISS staging improves risk assessment in myeloma: an International Myeloma Working Group collaborative project. *Leukemia*. 2013;27(3):711-717.
2. Palumbo A, Avet-Loiseau H, Oliva S, et al. Revised International Staging System for multiple myeloma: A report from International Myeloma Working Group. *J Clin Oncol*. 2015; 33(26):2863-2869.
3. Durie BG, Hoering A, Abidi MH, et al. Bortezomib with lenalidomide and dexamethasone versus lenalidomide and dexamethasone alone in patients with newly diagnosed myeloma without intent for immediate autologous stem-cell transplant (SWOG S0777): a randomised, open-label, phase 3 trial. *Lancet*. 2017;389(10068): 519-527.
4. Decaux O, Lodé L, Magrangeas F, et al; Intergroupe Francophone du Myélome. Prediction of survival in multiple myeloma based on gene expression profiles reveals cell cycle and chromosomal instability signatures in high-risk patients and hyperdiploid signatures in low-risk patients: a study of the Intergroupe Francophone du Myélome. *J Clin Oncol*. 2008;26(29):4798-4805.
5. Kuiper R, Broyl A, de Knecht Y, et al. A gene expression signature for high-risk multiple myeloma [published correction appears in *Leukemia*. 2014;28(5):1178-1180]. *Leukemia*. 2012;26(11):2406-2413.
6. Shaughnessy JD Jr, Zhan F, Burington BE, et al. A validated gene expression model of high-risk multiple myeloma is defined by deregulated expression of genes mapping to chromosome 1. *Blood*. 2007;109(6): 2276-2284.
7. Chng WJ, Chung TH, Kumar S, et al. Gene signature combinations improve prognostic stratification of multiple myeloma patients. *Leukemia*. 2016;30(5):1071-1078.
8. Chung TH, Mulligan G, Fonseca R, Chng WJ. A novel measure of chromosome instability can account for prognostic difference in multiple myeloma. *PLoS One*. 2013;8(6): e66361.
9. Dickens NJ, Walker BA, Leone PE, et al. Homozygous deletion mapping in myeloma samples identifies genes and an expression signature relevant to pathogenesis and outcome. *Clin Cancer Res*. 2010;16(6): 1856-1864.
10. Hose D, Rème T, Hielscher T, et al. Proliferation is a central independent prognostic factor and target for personalized and risk-adapted treatment in multiple myeloma. *Haematologica*. 2011;96(1):87-95.
11. Moreaux J, Klein B, Bataille R, et al. A high-risk signature for patients with multiple myeloma established from the molecular classification of human myeloma cell lines. *Haematologica*. 2011;96(4):574-582.
12. Shaughnessy JD Jr, Qu P, Usmani S, et al. Pharmacogenomics of bortezomib test-dosing identifies hyperexpression of proteasome genes, especially PSMD4, as novel high-risk feature in myeloma treated with Total Therapy 3. *Blood*. 2011;118(13):3512-3524.
13. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol*. 2013;14(4):R36.
14. Li H, Handsaker B, Wysoker A, et al; 1000 Genome Project Data Processing Subgroup. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009;25(16): 2078-2079.
15. Koboldt DC, Larson DE, Wilson RK. Using VarScan 2 for germline variant calling and somatic mutation detection. *Curr Protoc Bioinformatics*. 2013;44:15.4.1-17.
16. Koboldt DC, Zhang Q, Larson DE, et al. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res*. 2012;22(3): 568-576.
17. Nicorici D, Satalan M, Edgren H, et al. FusionCatcher - a tool for finding somatic fusion genes in paired-end RNA-sequencing data. *bioRxiv*. doi:10.1101/011650.
18. Liao Y, Smyth GK, Shi W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res*. 2013; 41(10):e108.
19. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. 2014;30(7):923-930.
20. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol*. 2010;11(10):R106.
21. Anders S, McCarthy DJ, Chen Y, et al. Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat Protoc*. 2013;8(9):1765-1786.
22. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
23. Luo W, Friedman MS, Shedden K, Hankenson KD, Woolf PJ. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics*. 2009;10:161.
24. Luo W, Pant G, Bhavnani YK, Blanchard SG Jr, Brouwer C. Pathview Web: user friendly pathway visualization and data integration. *Nucleic Acids Res*. 2017;45(W1):W501-W508.
25. Luo W, Brouwer C. Pathview: an R/Bioconductor package for pathway-based data integration and visualization. *Bioinformatics*. 2013;29(14):1830-1831.
26. Mulligan G, Mitsiades C, Bryant B, et al. Gene expression profiling and correlation with outcome in clinical trials of the proteasome inhibitor bortezomib. *Blood*. 2007;109(8): 3177-3188.
27. Todoerti K, Agnelli L, Fabris S, et al. Transcriptional characterization of a prospective series of primary plasma cell leukemia revealed signatures associated with tumor

- progression and poorer outcome. *Clin Cancer Res*. 2013;19(12):3247-3258.
28. Broyl A, Hose D, Lokhorst H, et al. Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. *Blood*. 2010;116(14):2543-2553.
  29. Law CW, Chen Y, Shi W, Smyth GK. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;15(2):R29.
  30. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res*. 2012;40(10):4288-4297.
  31. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
  32. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26(1):139-140.
  33. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010;11(3):R25.
  34. Robinson MD, Smyth GK. Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics*. 2007;23(21):2881-2887.
  35. Robinson MD, Smyth GK. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics*. 2008;9(2):321-332.
  36. Phipson B, Lee S, Majewski IJ, Alexander WS, Smyth GK. Robust Hyperparameter Estimation Protects against Hypervariable Genes and Improves Power to Detect Differential Expression. *Ann Appl Stat*. 2016;10(2):946-963.
  37. Filippone M, Masulli F, Rovetta S. Simulated annealing for supervised gene selection. *Soft Comput*. 2011;15(8):1471-1482.
  38. Weinberger KQ, Saul LK. Distance metric learning for large margin nearest neighbor classification. *J Mach Learn Res*. 2009;10:207-244.
  39. Bolli N, Avet-Loiseau H, Wedge DC, et al. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nat Commun*. 2014;5:2997.
  40. Chapman MA, Lawrence MS, Keats JJ, et al. Initial genome sequencing and analysis of multiple myeloma. *Nature*. 2011;471(7339):467-472.
  41. Lohr JG, Stojanov P, Carter SL, et al; Multiple Myeloma Research Consortium. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. *Cancer Cell*. 2014;25(1):91-101.
  42. Walker BA, Boyle EM, Wardell CP, et al. Mutational spectrum, copy number changes, and outcome: results of a sequencing study of patients with newly diagnosed myeloma. *J Clin Oncol*. 2015;33(33):3911-3920.
  43. Rashid NU, Sperling AS, Bolli N, et al. Differential and limited expression of mutant alleles in multiple myeloma. *Blood*. 2014;124(20):3110-3117.
  44. Bergsagel PL, Kuehl WM. Molecular pathogenesis and a consequent classification of multiple myeloma. *J Clin Oncol*. 2005;23(26):6333-6338.
  45. Kaiser MF, Walker BA, Hockley SL, et al. A TC classification-based predictor for multiple myeloma using multiplexed real-time quantitative PCR. *Leukemia*. 2013;27(8):1754-1757.
  46. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. 2011;27(12):1739-1740.
  47. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*. 2005;102(43):15545-15550.
  48. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst*. 2015;1(6):417-425.
  49. Sonneveld P, Schmidt-Wolf IG, van der Holt B, et al. Bortezomib induction and maintenance treatment in patients with newly diagnosed multiple myeloma: results of the randomized phase III HOVON-65/ GMMG-HD4 trial. *J Clin Oncol*. 2012;30(24):2946-2955.
  50. Jamieson NB, Maker AV. Gene-expression profiling to predict responsiveness to immunotherapy. *Cancer Gene Ther*. 2017;24(3):134-140.
  51. Cardoso F, van't Veer LJ, Bogaerts J, et al; MINDACT Investigators. 70-Gene signature as an aid to treatment decisions in early-stage breast cancer. *N Engl J Med*. 2016;375(8):717-729.
  52. Cullen J, Rosner IL, Brand TC, et al. A biopsy-based 17-gene genomic prostate score predicts recurrence after radical prostatectomy and adverse surgical pathology in a racially diverse population of men with clinically low- and intermediate-risk prostate cancer. *Eur Urol*. 2015;68(1):123-131.
  53. Dave SS, Wright G, Tan B, et al. Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells. *N Engl J Med*. 2004;351(21):2159-2169.
  54. Galon J, Costes A, Sanchez-Cabo F, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science*. 2006;313(5795):1960-1964.
  55. Gray RG, Quirke P, Handley K, et al. Validation study of a quantitative multigene reverse transcriptase-polymerase chain reaction assay for assessment of recurrence risk in patients with stage II colon cancer. *J Clin Oncol*. 2011;29(35):4611-4619.
  56. Klein EA, Cooperberg MR, Magi-Galluzzi C, et al. A 17-gene assay to predict prostate cancer aggressiveness in the context of Gleason grade heterogeneity, tumor multifocality, and biopsy undersampling. *Eur Urol*. 2014;66(3):550-560.
  57. Kratz JR, He J, Van Den Eeden SK, et al. A practical molecular assay to predict survival in resected non-squamous, non-small-cell lung cancer: development and international validation studies. *Lancet*. 2012;379(9818):823-832.
  58. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453-457.
  59. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med*. 2004;351(27):2817-2826.
  60. van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 2002;415(6871):530-536.
  61. Venook AP, Niedzwiecki D, Lopatin M, et al. Biologic determinants of tumor recurrence in stage II colon cancer: validation study of the 12-gene recurrence score in cancer and leukemia group B (CALGB) 9581. *J Clin Oncol*. 2013;31(14):1775-1781.
  62. Yothers G, O'Connell MJ, Lee M, et al. Validation of the 12-gene colon cancer recurrence score in NSABP C-07 as a predictor of recurrence in patients with stage II and III colon cancer treated with fluorouracil and leucovorin (FU/LV) and FU/LV plus oxaliplatin. *J Clin Oncol*. 2013;31(36):4512-4519.
  63. Baker JB, Dutta D, Watson D, et al. Tumour gene expression predicts response to cetuximab in patients with KRAS wild-type metastatic colorectal cancer. *Br J Cancer*. 2011;104(3):488-495.
  64. Fehrenbacher L, Spira A, Ballinger M, et al; POPLAR Study Group. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. *Lancet*. 2016;387(10030):1837-1846.
  65. Kim HK, Choi IJ, Kim CG, et al. Three-gene predictor of clinical outcome for gastric cancer patients treated with chemotherapy. *Pharmacogenomics J*. 2012;12(2):119-127.
  66. Ulloa-Montoya F, Louahed J, Dizier B, et al. Predictive gene signature in MAGE-A3 antigen-specific cancer immunotherapy. *J Clin Oncol*. 2013;31(19):2388-2395.
  67. Weiss GR, Grosh WW, Chianese-Bullock KA, et al. Molecular insights on the peripheral and intratumoral effects of systemic high-dose rIL-2 (aldesleukin) administration for the treatment of metastatic melanoma. *Clin Cancer Res*. 2011;17(23):7440-7450.
  68. Fonseca R, Abouzaid S, Bonafede M, et al. Trends in overall survival and costs of multiple myeloma, 2000-2014. *Leukemia*. 2017;31(9):1915-1921.