# Supplementary Materials and Methods

## Trial protocol, RNA isolation, and data generation

### PADIMAC trial protocol

Bortezomib, Adriamycin, and Dexamethasone (**PAD**) Therapy for Previously Untreated Patients with Multiple Myeloma: **I**mpact of **M**inimal Residual Disease (MRD) in Patients with Deferred **A**S**C**T (PADIMAC) was a Phase 2 single arm trial designed to determine the two-year progression-free survival (PFS) for patients who, having achieved very good partial response (VGPR) or better following PAD induction therapy, were stratified to a no-ASCT arm. Patients received PAD chemotherapy (bortezomib 1.3mg/m2 days 1, 4, 8, 11; doxorubicin 9mg/m2 days 1-4 and dexamethasone 40mg days 1-4 [and days 8-11 and 15-18 for the first cycle only]) for up to six cycles (and for a minimum of four). Bortezomib was initially given intravenously, but once approved this was switched to subcutaneous. Those failing to achieve partial response (PR) were offered salvage therapy off protocol.  All others had peripheral blood stem cell mobilisation using cyclophosphamide and GCSF. Depending on disease response, patients were then stratified to ASCT (in PR) or no further treatment (in VGPR or better). Responses were assessed by intent-to-treat analysis, using International Myeloma Working Group uniform response criteria1. Cytogenetic analysis for high risk features was performed by Fluorescence In Situ Hybridization (FISH) on CD138-selected cells.

### Generation of RNA-Sequencing (RNA-Seq) data

Bone marrow samples were obtained from patients prior to starting therapy. CD138+ myeloma cells were isolated using magnetic activated cell sorting CD138 MicroBeads (Miltenyi Biotech), resulting in cell suspensions with purity of greater than 90%. Total RNA isolation was carried out using PureLink RNA Micro Kit RNAqueous kit (Invitrogen). Samples were only processed if they had an RNA integrity number of greater than 6, as determined by the RNA 6000 Nano Kit (Agilent). Libraries were prepared from 250ng total RNA with the TruSeq Stranded mRNA LT sample preparation kit (Illumina) according to the manufacturer’s instructions, apart from a 10-minute fragmentation and 14 cycles of PCR. The average insert size was 150bp. Samples were sequenced on the NextSeq 500 instrument (Illumina) at 3.0pM using version 1 chemistry, resulting in more than 15 million 80bp reads per sample.

### Sequencing quality metrics

We obtained a median of 18823499 (IQR: 17128130 - 21247616) 81bp paired-end reads per sample. Across all samples, a median of 87.2% of sequences had FastQC quality scores of 30 or higher (IQR: 86.0 – 89.0). RNA-SeQC2 was used to assess the extent to which 3' bias was present in the samples, for the 1000 most highly expressed, and 1000 most lowly expressed genes respectively. The 3' bias score was found to positively correlate with the per sample RNA integrity number (with R2 values of 0.71 and 0.64 respectively).

### Identification of mutations

Sequencing run quality was assessed by evaluating cluster density, number of reads, read quality across all cycles, percentage mapping to the genome, and Phred score.3 RNA-SeQC2 was used to assess 3’ bias.Sequence read quality was measured using FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Sequencing quality metrics are described in the Supplementary Materials. Reads were mapped using TopHat4 against the NCBI reference transcriptome.5 Bam files were filtered to remove duplicate reads using the Picard MarkDuplicates tool (broadinstitute.github.io/picard/). Reads were aligned with Samtools.6 Single nucleotide variants (SNVs) and small indels were identified using VarScan,7,8 as described in Koboldt *et al*.7 (alternative and support protocol 1) and annotated using Annovar.9 RNA fusions were identified using FusionCatcher10 referenced against Ensembl80.11 Mutations were filtered by reference to four non-CD138 selected bone marrow harvests from normal healthy volunteers, the COSMIC database,12 the 1000 Genomes Project,13 the Exome Sequencing Project,14 and the Exome Aggregation Consortium.15

### Gene expression

Read counts were generated by mapping the bam files against the NCBI build 37.2 of the human transcriptome using the Rsubread package.16,17 Raw and count level data have been uploaded to GEO (GSE116324). Differentially expressed genes were identified using DESeq2,18-20 requiring a log2 fold change of ±1 and an adjusted p-value of less than 0.05. The Gage21 and Pathview22,23 packages were used for pathway analysis and core gene set identification. Pathways were identified with the false discovery rate controlled at 0.05.

## Machine learning

### Selection of orthologous datasets for testing

Test RNA-Seq datasets were derived from the CoMMpass project (<https://research.themmrf.org/>). They were from patients treated first-line with bortezomib-based therapy (i.e. bortezomib combined with alkylators, anthracyclines, and/or steroids, but not IMiDs); with RD; and with VRD. We excluded patients who had had an ASCT. Determination of myeloma sub-type is described below. Microarray test sets were obtained for relapsed/refractory patients treated with bortezomib24 (Gene Expression Omnibus [GEO] reference GSE9782), plasma cell leukemia (PCL) patients treated with RD25 (GEO GSE39925), and newly diagnosed myeloma patients treated with PAD followed by ASCT26 (GEO GSE19784). We refer to these data as the Millennium, PCL, and HOVON/GMMG datasets, respectively.

### Determination of myeloma sub-type in CoMMpass data

Clinical information was downloaded from the relating **C**linical **o**utcomes in **M**ultiple **M**yeloma to **p**ersonal **ass**essment of Genetic Profile (CoMMpass) project website (<https://research.themmrf.org/>). To determine subtype of myeloma (IgG, IgA, or light chain), we used the annotated status where it was supplied. In the cases it was absent, we judged IgH status by the presence of a monoclonal band and an IgG of greater than 16g/l with an IgA of less than 4g/l (for IgG myeloma) or an IgA of greater than 4g/l and an IgG less than 16g/l (for IgA myeloma). Cases without a significant monoclonal band but a kappa:lambda light chain ratio less than 0.26 or greater than 1.65 were defined as light chain myeloma.

### Pre-processing of RNA-Seq data

RNA-Seq counts were normalized by library size, converted to log counts per million (lcpm), and corrected for heteroscedasity according to published methods.27-33 Potential signature genes were identified from the PADIMAC dataset by an empirical Bayes method.34 Genes were further ranked using synthetic annealing with ten-fold internal cross-validation, then selected by their vector of aged references as described.35 The error rate for determining the energy of the system was determined by a support vector machine implemented from the e1071 package (<https://cran.r-project.org/web/packages/e1071/index.html>) with default settings.

### Co-normalization of RNA-Seq and microarray data

We were limited by the availability of additional RNA-Seq datasets from appropriately treated patients for testing. We therefore turned to microarray datasets. However, we were unaware of published methods for co-normalizing RNA-Seq and microarray data, which would be required to continue to use the PADIMAC data as a training set. We anticipated significant difficulties in aligning RNA-Seq and microarray data because of the fundamental difference in the way they are generated, as reviewed extensively elsewhere.18-20,28,30,31

We believed that the correction of heteroscedasity that was part of our RNA-Seq processing pipeline (see Figure S18) would be beneficial. We also felt that it was critical that the microarray data should not cluster exclusively with one or other class in the RNA-Seq data, which almost certainly would have led to the microarray samples all being classified in the same way. We therefore took the following empirical approach. We concatenated the PADIMAC training array (seven gene rows by 44 sample columns) with the corresponding microarray data. Where multiple probe sets mapped to a single gene, we selected the probe set with the maximum median absolute deviation. We then row- and column-normalized the RNA-Seq and microarray data to a median of zero and median absolute deviation of one. We examined these data to check that the microarray data did not exclusively co-localize with the bortezomib-good or lenalidomide-good samples. If there was co-localization likely to lead to identical classification, we considered alternative normalization procedures.

In the discussion below, we have focused on the first three principal components (PCs) of the data, for ease of visualization. In all cases, the first three PCs contributed at least 60% of the variance and PCs 4-7 had no additional effect on interpretation (data not shown). The Millennium data had a wider range in all three PCs than the PADIMAC data and the samples were displaced from the PADIMAC samples along the first PC (Figure S19). However, they clustered very closely in the second and third PCs. Importantly, there was no evidence that the Millennium data clustered exclusively with either the bortezomib-good or lenalidomide-good PADIMAC samples. Compared to the Millennium data, the HOVON/GMMG samples clustered more closely with the PADIMAC samples, although there was slight displacement along both the first and second PCs (Figure S20). Again, there was no exclusive clustering with either bortezomib-good or lenalidomide-good samples in PADIMAC. We therefore proceeded to use both these normalized datasets without further modification.

After standard normalization, the PCL samples formed a very tight cluster with the lenalidomide-good samples of PADIMAC and this co-cluster was clearly distinct from the bortezomib-good samples of PADIMAC (Figure S21). We had no *a priori* reason to believe that this reflected the true situation, i.e. we would not expect all PCL patients to benefit from lenalidomide therapy in preference to bortezomib-based treatment. We therefore re-normalized the data using quantile normalization. This resulted in the microarray data being better distributed between bortezomib-good and lenalidomide-good RNA-Seq samples (figure S22). This quantile-normalized dataset was therefore used for the LMNN predictions without further modification.

### Choice of machine learning algorithm

We did not apply a testing phase for different machine learning algorithms. Rather, we chose to use the Largest Margin Nearest Neighbor (LMNN) algorithm *a priori*, based partly on theoretical considerations and partly on our previous experience with classifying microarray data where it had outperformed other methods (unpublished observations). The theoretical considerations were threefold. First, we had noted the experience of Amin *et al*.36 in their analysis of complete response (CR) and gene expression in patients treated with bortezomib-based therapy. They had employed a variety of machine learning techniques, including support vector machines (SVM), decision trees, K-nearest neighbors, linear discriminant analysis, Bayesian additive regression trees, prediction analysis of microarrays, and artificial neural networks. None of these methods yielded a statistically significant prediction of CR. Second, whilst we were attracted by the intuitive approach of geometric methods, we considered that non-Euclidean distance metrics (or pseudometrics) were likely to be required. Third, we wanted a model that might be flexible enough to be extended from a two-class to a multi-class classification in the future. Thus, whilst both SVM and LMNN were geometric approaches that did not rely on Euclidean distance metrics (by application of the kernel trick and a Mahalanobis pseudometric, respectively), extension of SVM into a multi-class problem was likely to be inherently more complex than extension of LMNN.37

It will be noted that in gene selection by simulated annealing we adopted SVM rather than LMNN. This was a practical consideration based on the fact that one round of training and testing with LMNN took in excess of a minute whereas the equivalent cycle in SVM as implemented by the e1071 package took approximately 0.002 seconds. Thus, the simulated annealing, which took around 8 hours with SVM, might have been expected to take approximately 27 years with LMNN. Use of a high-performance cluster to accelerate simulated annealing with LMNN may have been possible but would not have been trivial given the iterative nature of the simulated annealing algorithm and would have still been anticipated to require weeks to months to run.

### Training, validation, and testing

Having chosen LMNN as our preferred machine learning algorithm, performance within the PADIMAC dataset was checked by ten-fold cross-validation with 45:45 training:validation splits. For external testing, all PADIMAC data were used for training, with an initial 50:50 split into a training and internal validation set that was fixed for all testing. R and Matlab scripts that replicate this process have been included with the Supplementary Materials. All the CoMMpass, Millennium, PCL, or HOVON/GMMG data were used for testing. To determine the robustness of the signature performance in each case, 100 further training:validation splits of the PADIMAC training set were randomly obtained and used to make new assignments. Null assignments for each split were derived by randomizing these observed assignments, thus preserving the assignment frequencies. The performance of the null and observed assignments were then compared using the Mann-Whitney-Wilcoxon test. P-values of 0.05 or less were considered significant.

# References

1. Durie BG, Harousseau JL, Miguel JS, et al: International uniform response criteria for multiple myeloma. Leukemia 20:1467-73, 2006

2. DeLuca DS, Levin JZ, Sivachenko A, et al: RNA-SeQC: RNA-seq metrics for quality control and process optimization. Bioinformatics 28:1530-2, 2012

3. Ewing B, Hillier L, Wendl MC, et al: Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 8:175-85, 1998

4. Kim D, Pertea G, Trapnell C, et al: TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14:R36, 2013

5. O'Leary NA, Wright MW, Brister JR, et al: Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 44:D733-45, 2016

6. Li H, Handsaker B, Wysoker A, et al: The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-9, 2009

7. Koboldt DC, Larson DE, Wilson RK: Using VarScan 2 for Germline Variant Calling and Somatic Mutation Detection. Curr Protoc Bioinformatics 44:15 4 1-17, 2013

8. Koboldt DC, Zhang Q, Larson DE, et al: VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 22:568-76, 2012

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

10. Nicorici D, Satalan M, Edgren H, et al: FusionCatcher - a toll for finding somatic fusion genes in paired-end RNA-sequencing data. bioRxiv doi:10.1101/011650, 2014

11. Yates A, Akanni W, Amode MR, et al: Ensembl 2016. Nucleic Acids Res 44:D710-6, 2016

12. Forbes SA, Beare D, Boutselakis H, et al: COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res 45:D777-D783, 2017

13. Genomes Project C, Auton A, Brooks LD, et al: A global reference for human genetic variation. Nature 526:68-74, 2015

14. Fu W, O'Connor TD, Jun G, et al: Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. Nature 493:216-20, 2013

15. Lek M, Karczewski KJ, Minikel EV, et al: Analysis of protein-coding genetic variation in 60,706 humans. Nature 536:285-91, 2016

16. Liao Y, Smyth GK, Shi W: The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res 41:e108, 2013

17. Liao Y, Smyth GK, Shi W: featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923-30, 2014

18. Anders S, Huber W: Differential expression analysis for sequence count data. Genome Biol 11:R106, 2010

19. Anders S, McCarthy DJ, Chen Y, et al: Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. Nat Protoc 8:1765-86, 2013

20. Love MI, Huber W, Anders S: Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550, 2014

21. Luo W, Friedman MS, Shedden K, et al: GAGE: generally applicable gene set enrichment for pathway analysis. BMC Bioinformatics 10:161, 2009

22. Luo W, Pant G, Bhavnasi YK, et al: Pathview Web: user friendly pathway visualization and data integration. Nucleic Acids Res 45(W1):W501-W508, 2017

23. Luo W, Brouwer C: Pathview: an R/Bioconductor package for pathway-based data integration and visualization. Bioinformatics 29:1830-1, 2013

24. Mulligan G, Mitsiades C, Bryant B, et al: Gene expression profiling and correlation with outcome in clinical trials of the proteasome inhibitor bortezomib. Blood 109:3177-88, 2007

25. 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 19:3247-58, 2013

26. Broyl A, Hose D, Lokhorst H, et al: Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. Blood 116:2543-53, 2010

27. Law CW, Chen Y, Shi W, et al: voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 15:R29, 2014

28. McCarthy DJ, Chen Y, Smyth GK: Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. Nucleic Acids Res 40:4288-97, 2012

29. Ritchie ME, Phipson B, Wu D, et al: limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43:e47, 2015

30. Robinson MD, McCarthy DJ, Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139-40, 2010

31. Robinson MD, Oshlack A: A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol 11:R25, 2010

32. Robinson MD, Smyth GK: Moderated statistical tests for assessing differences in tag abundance. Bioinformatics 23:2881-7, 2007

33. Robinson MD, Smyth GK: Small-sample estimation of negative binomial dispersion, with applications to SAGE data. Biostatistics 9:321-32, 2008

34. Phipson B, Lee S, Majewski IJ, et al: Robust Hyperparameter Estimation Protects against Hypervariable Genes and Improves Power to Detect Differential Expression. Ann Appl Stat 10:946-963, 2016

35. Filippone M, Masulli F, Rovetta S: Simulated annealing for supervised gene selection. Soft Computing 15:1471-1482, 2011

36. Amin SB, Yip WK, Minvielle S, et al: Gene expression profile alone is inadequate in predicting complete response in multiple myeloma. Leukemia 28:2229-34, 2014

37. Weinberger KQ, Saul LK: Distance metric learning for large margin nearest neighbor classification. Journal of Machine Learning Research 10:207-244, 2009

# Supplementary Tables

|  |  |
| --- | --- |
| **Tool** | **Version** |
| 1000 Genomes Project | August 2015 |
| Annovar | 2015-03-22 |
| CoMMpass | IA10 |
| COSMIC | 70 |
| DESeq2 | 1.10.1 |
| e1071 | 1.6-8 |
| EdgeR | 3.20.5 |
| Ensembl | 80 |
| FastQC | 0.11.4 |
| FusionCatcher | 0.99.42 |
| Gage | 2.20.0 |
| Limma | 3.34.5 |
| LMNN | 3.0.0 |
| NCBI reference transcriptome | 37.2 |
| Pathview | 1.10.1 |
| Picard MarkDuplicates | 1.136 |
| RNA-SeQC | 1.1.8 |
| Samtools | 0.1.19 |
| TopHat | 2.0.10 |
| VarScan | 2.3.9 |

**Table S1 – Versions of bioinformatic tools used.**

|  |  |  |
| --- | --- | --- |
| **Median age in years (range)** |  | 56 (30-67) |
| **Ig subtype** | IgG | 26 (59.1%) |
|  | IgA | 14 (31.8%) |
|  | LC | 4 (9.1%) |
| **FISH abnormalities** | t(4:14) | 5 (11.4%) |
|  | t(11:14) | 6 (13.6%) |
|  | t(6;14) | 1 (2.3%) |
|  | t(14;16) | 1 (2.3%) |
|  | Hyperdiploidy | 11 (25.0%) |
|  | 17p- | 2 (4.5%) |
|  | 1q+ | 10 (22.7%) |
|  | 1p- | 3 (6.8%) |
|  | 13- | 4 (9.1%) |
| **ISS stage** | I | 12 (27.3%) |
|  | II | 22 (50.0%) |
|  | III | 10 (22.7%) |
| **Best response to PAD induction** | CR | 3 (6.8%) |
|  | VGPR | 17 (38.6%) |
|  | PR | 11 (25.0%) |
|  | SD | 7 (15.9%) |
|  | PD | 6 (13.6%) |
|  | **ORR** | **31 (70.5%)** |

**Table S2 – Clinical features of PADIMAC dataset.** LC: light-chain myeloma. CR: complete response. VGPR: very good partial response. PR: partial response. SD: stable disease. PD: progressive disease. ORR: overall response rate. N=44.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sample** | **FISH translocation** | **Ig fusions** | **Other fusions** | **Total fusions** | **Fusions** |
| PAD-141 | t(4:14) | 5 | 2 | 7 | IGHG1--WHSC1  IGHG3--WHSC1  BCL2L11--IGHG1  LINC-PINT--IGHG1  YWHAE--IGHG1  TPM4—UBC  OAZ1--KLF2 |
| PAD-145 |  | 2 | 5 | 7 | BLOC1S5-TXNDC5--IGKC  IGKC--TSC22D3  RRBP1--ADD3  TBC1D13--COQ4  GYPE—GYPA  CRTC1—CNFN  DOT1L--POU2F2 |
| PAD-152 | t(14;16) | 1 | 4 | 5 | IGH@--WWOX  B2M—UBC  CDKN1A—UBC  LLFOS-48D6.2--KLF2  RND3--UBC |
| PAD-140 | t(4;14) | 4 | 0 | 4 | WHSC1--IGH@  KLF2--IGHA1  HAGH--IGH@  PLD4--IGHA1 |
| PAD-130 | t(4;14) | 3 | 1 | 4 | IGH@--WHSC1  B2M--IGH@  IGH@--BLOC1S5-TXNDC5  PGM3--FGD5-AS1 |
| PAD-049 | IGH @14q | 2 | 2 | 4 | ITGB7--IGHA1  ITGB7--IGH@  TXNDC5--KLF2  RND3--KLF2 |
| PAD-056 | t(4;14) | 3 | 0 | 3 | IGHG1--WHSC1  WHSC1--IGH@  IGKC--FGFR3 |
| PAD-095 |  | 0 | 3 | 3 | TMEM156--GBA3  SMAP2--FPGT-TNNI3K  FERMT3--CKAP5 |
| PAD-120 |  | 0 | 3 | 3 | SULF2--CD40  GLCCI1--MAP3K14  ARID1A--AGBL4 |
| PAD-124 | t(11;14) | 2 | 0 | 2 | IGKV1-8—PSAP  FER1L4--IGKC |
| PAD-024 |  | 2 | 0 | 2 | IGH@--ZFP36L1  PVT1--IGKC |
| PAD-075 |  | 2 | 0 | 2 | B2M--IGKV1-8  H1FX--IGHA1 |
| PAD-039 |  | 1 | 1 | 2 | PVT1--IGLV3-1  FOSB--UBC |
| PAD-149 |  | 1 | 1 | 2 | B2M--IGH@  MFSD4--NR3C1 |
| PAD-035 |  | 0 | 2 | 2 | FOSB--KLF2  OAZ1--KLF2 |
| PAD-131 | t(4;14) | 1 | 0 | 1 | IGH@--WHSC1 |
| PAD-146 |  | 1 | 0 | 1 | B2M--IGHA2 |
| PAD-061 | t(11;14) | 0 | 1 | 1 | RCE1--PPP1R37 |
| PAD-125 | IGH @14q | 0 | 1 | 1 | C2CD5--CCDC91 |
| PAD-139 |  | 0 | 1 | 1 | LSP1--TSSC4 |
| PAD-017 |  | 0 | 1 | 1 | IL3RA--CFAP74 |
| PAD-128 |  | 0 | 1 | 1 | TXNDC11--TNFRSF17 |

**Table S3 – Gene fusions identified within the PADIMAC dataset.** Only samples in which gene fusions were identified are presented. Identified FISH translocations are recorded. Fusions were categorized as involving the light or heavy chain immunoglobulin loci (Ig fusions) or non-immunoglobulin loci (other fusions). Where reciprocal fusions were identified by FusionCatcher, both are recorded. Internal deletions within the immunoglobulin loci have been excluded. Fusions identified in the normal samples have also been excluded.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Gene** | **PADIMAC RNAseq (n=44)** | **Walker *et al.***  **(n=463)** | **Lohr *et al.***  **(n=203)** | **Bolli *et al.***  **(n=67)** |
| **ATM** | 2.3% | 3.0% | 3.9% | 2.3% |
| **BRAF** | 0.0% | 6.7% | 5.9% | 14.9% |
| **CCND1** | 0.0% | 2.1% | 4.4% | 4.5% |
| **DIS3** | 0.0% | 8.6% | 11.3% | 1.5% |
| **EGR1** | 0.0% | 3.5% | 3.9% | 6.0% |
| **FGFR3** | 2.3% | 2.6% | 2.0% | 0.0% |
| **IRF4** | 0.0% | 3.2% | 0.0% | 2.6% |
| **KRAS** | 11.4% | 21.1% | 25.6% | 47.8% |
| **NRAS** | 13.6% | 19.4% | 19.7% |
| **TP53** | 4.5% | 3.0% | 8.3% | 13.4% |
| **TRAF3** | 0.0% | 3.7% | 5.9% | 3.0% |
| **RAS pathway** | 25.0% | 43.2% |  |  |
| **NF-KB pathway** | 2.3% | 17% |  |  |

**Table S4 – Comparison of detection rates of recognized recurrent single nucleotide variants and pathways between PADIMAC RNAseq and published DNA sequencing datasets.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Bortezomib-standard (N=31)** | **Bortezomib-good (N=13)** | **p-value** | **q-value** |
| **Age < 56**  **Age >= 56** | 16  15 | 6  7 | 1.00 | 1.00 |
| **ISS I**  **ISS II**  **ISS III** | 11  15  5 | 1  7  5 | 0.10 | 0.31 |
| **IgA**  **IgG**  **Light chain** | 5  23  3 | 5  7  1 | 0.25 | 0.37 |

**Table S5 – Testing associations between bortezomib-good responders and clinical features.** Contingency tables, p-values, and q-values are shown. P-values are from Fisher’s exact test and q-values represent corrections for multiple hypothesis testing using the method of Benjamini and Hochberg. The clinical features are age (cut-off reflects the median), international staging system (ISS), and subtype.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Bortezomib- standard** | **Bortezomib-good** | **p-value** | **q-value** |
| **Standard-risk cytogenetics**  **Poor-risk cytogenetics** | 22  9 | 10  3 | 1 | 1 |
| **No RAS mutation**  **RAS mutation** | 21  10 | 12  1 | 0.13 | 0.22 |
| **No translocation**  **Any translocation** | 19  12 | 3  10 | 0.05 | 0.11 |
| **No Ig translocation**  **Any Ig translocation** | 23  8 | 7  6 | 0.29 | 0.36 |
| **No β2-microglobulin translocation**  **Any β2-microglobulin translocation** | 30  1 | 9  4 | 0.02 | 0.11 |

**Table S6 – Testing associations between bortezomib-good responders and key cytogenetic, SNV, and translocation events.** Contingency tables, p-values, and q-values are shown. P-values are from Fisher’s exact test and q-values represent corrections for multiple hypothesis testing using the method of Benjamini and Hochberg. Poor-risk cytogenetics are defined as the presence of t(4;14), MAF translocation, 1p loss, 1q gain, or 17p loss. RAS mutation refers to the presence of an NRAS or KRAS mutation. Any translocation refers to the presence of at least one translocation that was not observed in the normal samples. Any Ig translocation refers to the presence of a translocation involving the immunoglobulin heavy or light chain regions, but excluding internal Ig deletions and any fusions observed in normal samples. β2-microglobulin translocation refers to any translocation involving the β2-microglobulin gene.

|  |  |
| --- | --- |
| **Gene** | **Protein function** |
| EMC9 | Endoplasmic reticulum membrane protein complex subunit. |
| FAM171B | Unknown. |
| PLEK | Pleckstrin. Substrate of protein kinase C. Involved in exocytosis. |
| MYO9B | Myosin family motor protein. Moves along actin filaments and regulates G proteins. |
| RCN3 | CREC family endoplasmic reticulum protein. Chaperone protein in secretory pathway. |
| FLNB | Filamin. Actin-crosslinking protein involved in signal transduction. |
| KIF1C | Kinesin motor. Retrograde transport of Golgi vesicles to endoplasmic reticulum. |

**Table S7 – Genes comprising the seven-gene signature.** Brief functional descriptions are given.

|  |  |  |  |
| --- | --- | --- | --- |
| **KEGG pathway code** | **Pathway** | **p-value** | **q-value** |
| hsa04650 | Natural killer cell mediated cytotoxicity | 0.000005 | 0.001431 |
| hsa04660 | T cell receptor signaling pathway | 0.000028 | 0.003915 |
| hsa04144 | Endocytosis | 0.000251 | 0.023283 |
| hsa04062 | Chemokine signaling pathway | 0.000704 | 0.039165 |
| hsa04960 | Aldosterone-regulated sodium reabsorption | 0.000636 | 0.039165 |

**Table S8 – Pathways identified from supervised analysis of the bortezomib-good group versus the bortezomib-standard group of patients.** Redundant pathways were removed with Gage. Pathways with a q-value of less than 0.05 are shown.

|  |  |  |
| --- | --- | --- |
| **Median age in years (range)** |  | 70 (37-90) |
| **Ig subtype** | IgG | 78/137 (56.9%) |
|  | IgA | 25/137 (18.2%) |
|  | IgM | 1/137 (0.7%) |
|  | LC | 33/137 (24.1%) |
|  | Data not available | 10 |
| **FISH abnormalities** | t(4:14) | 21/89 (23.6%) |
|  | t(11:14) | 29/85 (34.1%) |
|  | t(6;14) | 1/43 (2.3%) |
|  | t(14;16) | 12/80 (15.0%) |
|  | Hyperdiploidy | 16 |
|  | 17p- | 16/89 (18.0%) |
|  | 1q+ | 24/64 (37.5%) |
|  | 1p- | 6/59 (10.2%) |
|  | 13- | 16/52 (30.8%) |
| **ISS stage** | I | 31/137 (22.6%) |
|  | II | 56/137 (40.9%) |
|  | III | 50/137 (36.5%) |
|  | Data not available | 10 |
| **Treatment** | VCD | 58 (39.5%) |
|  | BD | 47 (32.0%) |
|  | VMP | 29 (19.7%) |
|  | Single agent | 5 (3.4%) |
|  | Other combinations | 8 (5.4%) |
| **Best response to treatment** | CR | 10/125 (8.0%) |
|  | VGPR | 63/125 (50.4%) |
|  | PR | 39/125 (31.2%) |
|  | SD | 12/125 (9.6%) |
|  | PD | 1/125 (0.8%) |
|  | Data not available | 22 |
|  | **ORR** | **112/125 (89.6%)** |

**Table S9 – Clinical features of CoMMpass bortezomib-treated dataset.** LC: light-chain myeloma. CR: complete response. VGPR: very good partial response. PR: partial response. SD: stable disease. PD: progressive disease. ORR: overall response rate. N=147.

|  |  |  |
| --- | --- | --- |
| **Median age in years (range)** |  | 72.5 (51-90) |
| **Ig subtype** | IgG | 21/33 (63.6%) |
|  | IgA | 5/33 (15.2%) |
|  | LC | 7/33 (21.2%) |
|  | Data not available | 7 |
| **FISH abnormalities** | t(4:14) | 1/26 (3.8%) |
|  | t(11:14) | 5/27 (18.5%) |
|  | t(6;14) | 0/17 (0%) |
|  | t(14;16) | 3/26 (11.5%) |
|  | Hyperdiploidy | 2 |
|  | 17p- | 2/27 (7.4%) |
|  | 1q+ | 6/21 (28.6%) |
|  | 1p- | 2/20 (10.0%) |
|  | 13- | 13/31 (41.9%) |
| **ISS stage** | I | 11/36 (30.6%) |
|  | II | 15/36 (41.7%) |
|  | III | 10/36 (27.8%) |
|  | Data not available | 4 |
| **Best response to treatment** | CR | 5/37 (13.5%) |
|  | VGPR | 15/37 (40.5%) |
|  | PR | 10/37 (27.0%) |
|  | SD | 7/37 (18.9%) |
|  | PD | 0/37 (0%) |
|  | Data not available | 3 |
|  | **ORR** | **30/37 (81.1%)** |

**Table S10 – Clinical features of CoMMpass lenalidomide/dexamethasone-treated dataset.** LC: light-chain myeloma. CR: complete response. VGPR: very good partial response. PR: partial response. SD: stable disease. PD: progressive disease. ORR: overall response rate. N=40.

|  |  |  |
| --- | --- | --- |
| **Median age in years (range)** |  | 63 (27-88) |
| **Ig subtype** | IgG | 123/197 (62.4%) |
|  | IgA | 46/197 (23.4%) |
|  | IgM | 1/197 (0.8%) |
|  | LC | 27/197 (13.7%) |
|  | Data not available | 11 |
| **FISH abnormalities** | t(4:14) | 25/162 (11.9%) |
|  | t(11:14) | 40/168 (20.3%) |
|  | t(6;14) | 1/89 (0%) |
|  | t(14;16) | 12/150 (1.6%) |
|  | Hyperdiploidy | 35 |
|  | 17p- | 13/135 (9.6%) |
|  | 1q+ | 48/150 (32.0%) |
|  | 1p- | 15/140 (10.7%) |
|  | 13- | 41/164 (25.0%) |
| **ISS stage** | I | 61/202 (30.2%) |
|  | II | 84/202 (41.6%) |
|  | III | 57/202 (28.2%) |
|  | Data not available | 6 |
| **Best response to treatment** | CR | 38/203 (19.1%) |
|  | VGPR | 114/203 (45.7%) |
|  | PR | 40/203 (28.7%) |
|  | SD | 11/203 (6.4%) |
|  | PD | 0/203 (0%) |
|  | Data not available | 8 |
|  | **ORR** | **192/203 (94.6%)** |

**Table S11 – Clinical features of CoMMpass bortezomib/lenalidomide/dexamethasone- (VRD) treated dataset.** LC: light-chain myeloma. CR: complete response. VGPR: very good partial response. PR: partial response. SD: stable disease. PD: progressive disease. ORR: overall response rate. N=208.

|  |  |  |
| --- | --- | --- |
|  | **Lenalidomide-good (N=46)** | **Bortezomib-good (N=19)** |
| **Hyperdiploid** | 9 | 9 |
| **t(11;14)** | 17 | 4 |
| **t(4;14)** | 11 | 2 |
| **t(14;16)** | 9 | 4 |

**Table S12 – Associations between cytogenetics and the seven-gene signature associations in CoMMpass bortezomib- and RD-treated patients.** Only patients with available cytogenetic data have been included.

|  |  |  |  |
| --- | --- | --- | --- |
| **Row** | **Comparison** | **Univariate hazard ratio (95% CI); p-value** | **Multivariate hazard ratio (95% CI); p-value** |
| **1** | Incorrectly- *versus* correctly- treated bortezomib or RD (PFS) | 0.44 (0.25-0.78); p=0.0034 | 0.29 (0.14-0.62); p=0.0014 |
| **2** | Incorrectly- *versus* correctly- treated bortezomib or RD (OS) | 0.44 (0.23-0.84); p=0.0088 | 0.36 (0.16-0.80); p=0.01 |
| **3** | Correctly-treated bortezomib or RD *versus* VRD (OS) | 1.15 (0.60-2.21); p=0.67 | 0.90 (0.40-2.02); p=0.80 |
| **4** | Correctly-treated bortezomib or RD *versus* VRD (PFS) | 0.89 (0.52-1.52); p=0.68 | 0.58 (0.29-1.17); p=0.13 |
| **5** | VRD-treated – lenalidomide-best *versus* bortezomib-best (OS) | 1.23 (0.58-2.60); p=0.59 | 0.99 (0.42-2.32); p=0.98 |
| **6** | VRD-treated – lenalidomide-best *versus* bortezomib-best (PFS) | 0.93 (0.52-1.68); p=0.82 | 0.93 (0.49-1.77); p=0.83 |

**Table S13 – Incorporation of clinical feature into the Cox regression models does not influence outcome of survival comparisons in CoMMpass.** Additional variables in the multivariate model were age ≥ 75, International Staging System stage, and myeloma subtype (IgG, IgA, or light chain). All signature assignments were based on the seven-gene signature. Correctly-treated patients (rows 1-4) 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. The hazard ratio (HR) reflects the hazard of the second group in each comparison compared to the first group.

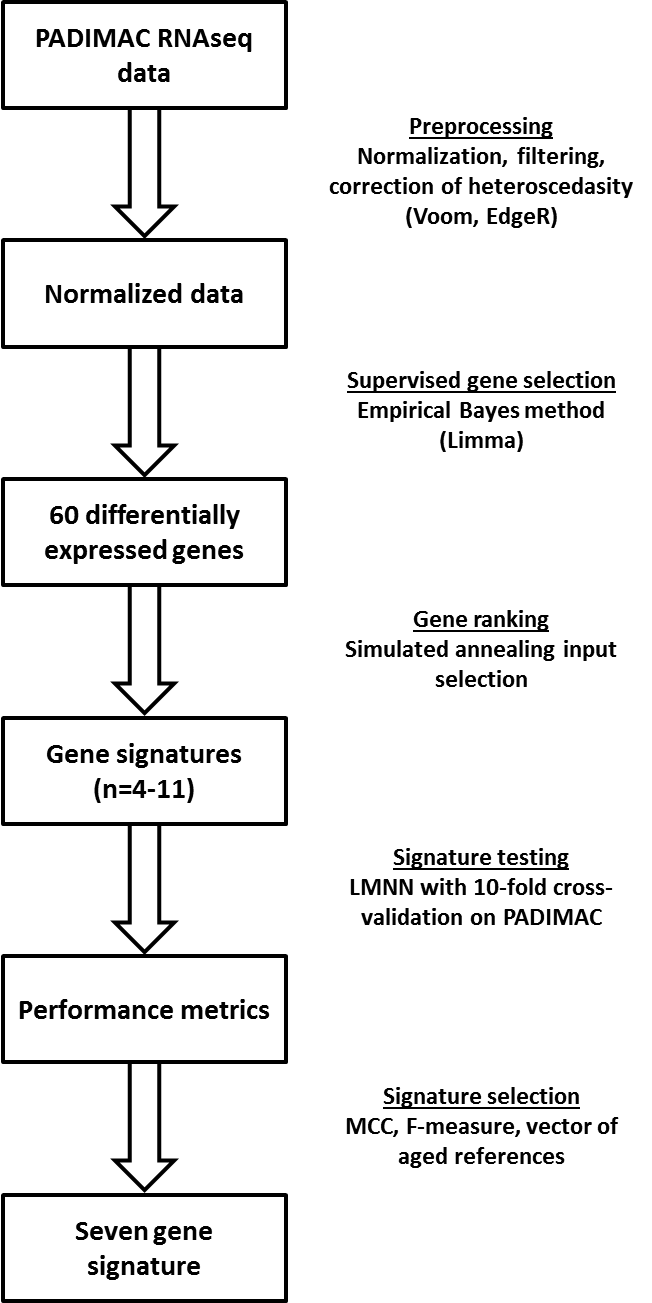
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Lenalidomide-good (N=139)** | **Bortezomib-good (N=48)** | **p-value** | **q-value** |
| **GPI50 low**  **GPI50 high** | 114  25 | 44  4 | 0.16 | 0.60 |
| **UAMS70 low**  **UAMS70 high** | 120  19 | 44  4 | 0.45 | 0.60 |
| **EMC92 low**  **EMC92 high** | 106  33 | 40  8 | 0.42 | 0.60 |
| **IFM15 low**  **IFM15 high** | 99  40 | 34  14 | 1.00 | 1.00 |

**Table S14 – Testing associations between the seven-gene model and various predictive models.** Contingency tables, p-values, and q-values are shown. P-values are from Fisher’s exact test and q-values represent corrections for multiple hypothesis testing using the method of Benjamini and Hochberg. For each predictive model, the cut-off between low and high is the same as that used in figure S14.

|  |  |  |  |
| --- | --- | --- | --- |
| **Row** | **Comparison** | **Univariate hazard ratio (95% CI); p-value** | **Multivariate hazard ratio (95% CI); p-value** |
| **1** | Incorrectly- *versus* correctly- treated bortezomib or RD (PFS) | 0.44 (0.25-0.78); p=0.0034 | 0.29 (0.14-0.62); p=0.0014 |
| **2** | Incorrectly- *versus* correctly- treated bortezomib or RD (OS) | 0.44 (0.23-0.84); p=0.0088 | 0.36 (0.16-0.80); p=0.01 |

**Table S15 – Incorporation of prognostic gene expression scores into the Cox regression models does not influence outcome of survival comparisons in CoMMpass.** Additional variables in the multivariate model were score ≥ poor prognosis threshold for each of the GPI50, EMC92, UAMS70, and IFM15 signatures. Correctly-treated patients were those predicted by the seven-gene signature as lenalidomide-best and who were treated with RD or those predicted as bortezomib-best and treated with bortezomib-based therapy. The hazard ratio (HR) reflects the hazard of the second group in each comparison compared to the first group.

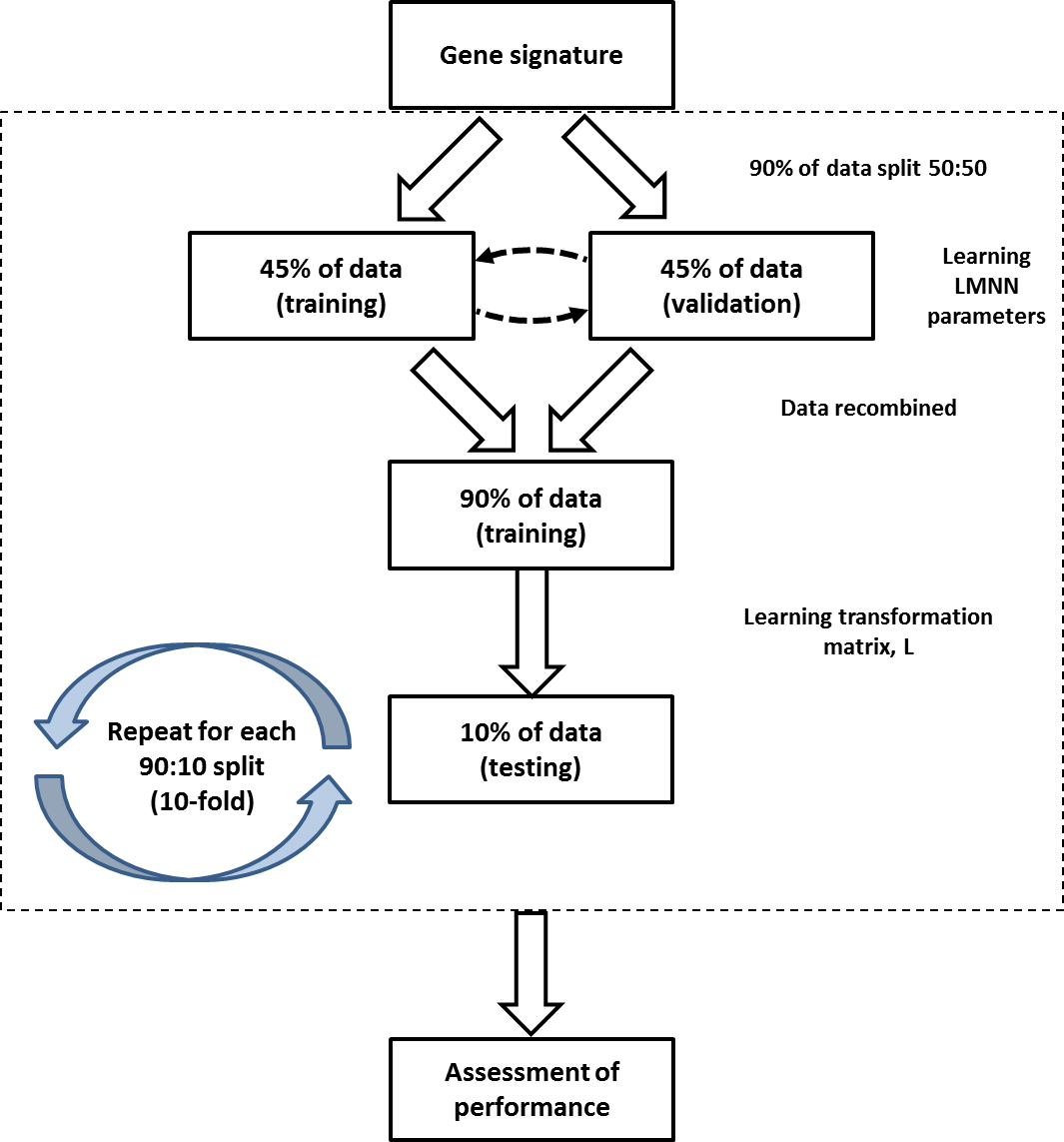
# Supplementary Figures

****

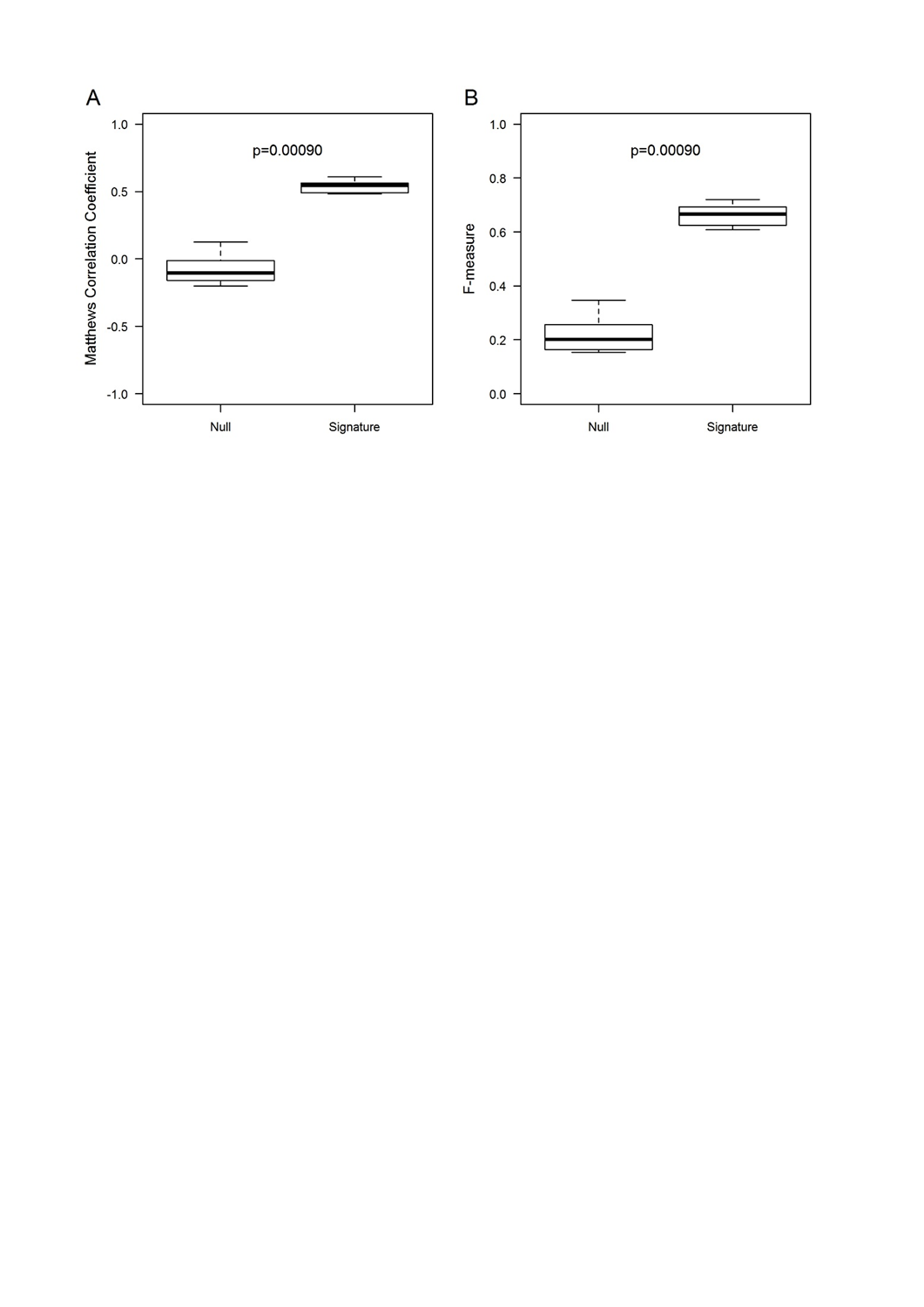
**Figure S1 – Overview of seven-gene signature.** Data were normalized, filtered and corrected for heteroscedasity. The top 60 most differentially expressed genes associated with good prognosis were selected and these were further ranked by synthetic annealing. Ten-fold cross-validation using the Largest Margin Nearest Neighbours (LMNN) algorithm was used to assess performance and the best signature was identified. MCC – Matthews Correlation Coefficient.

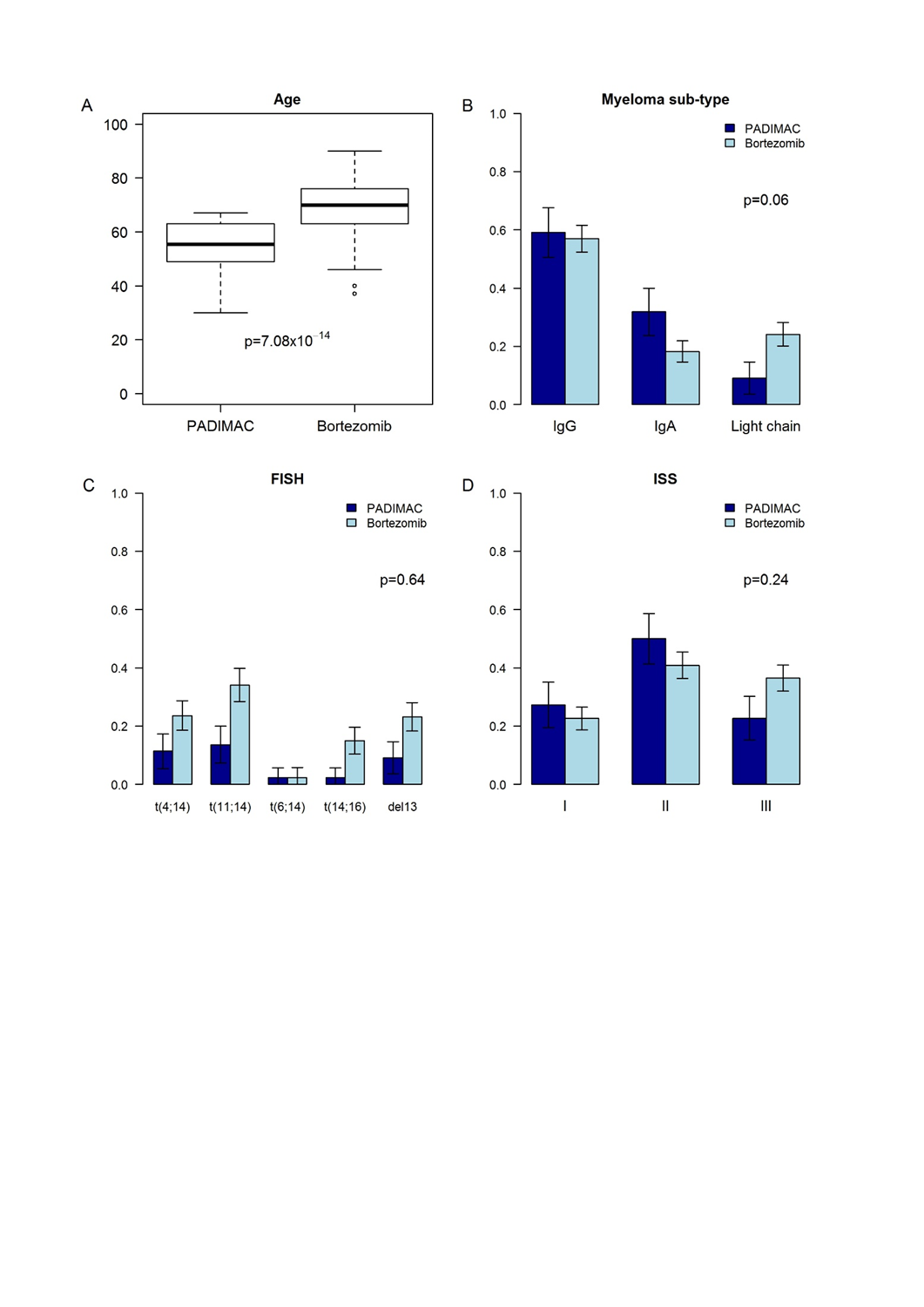
****

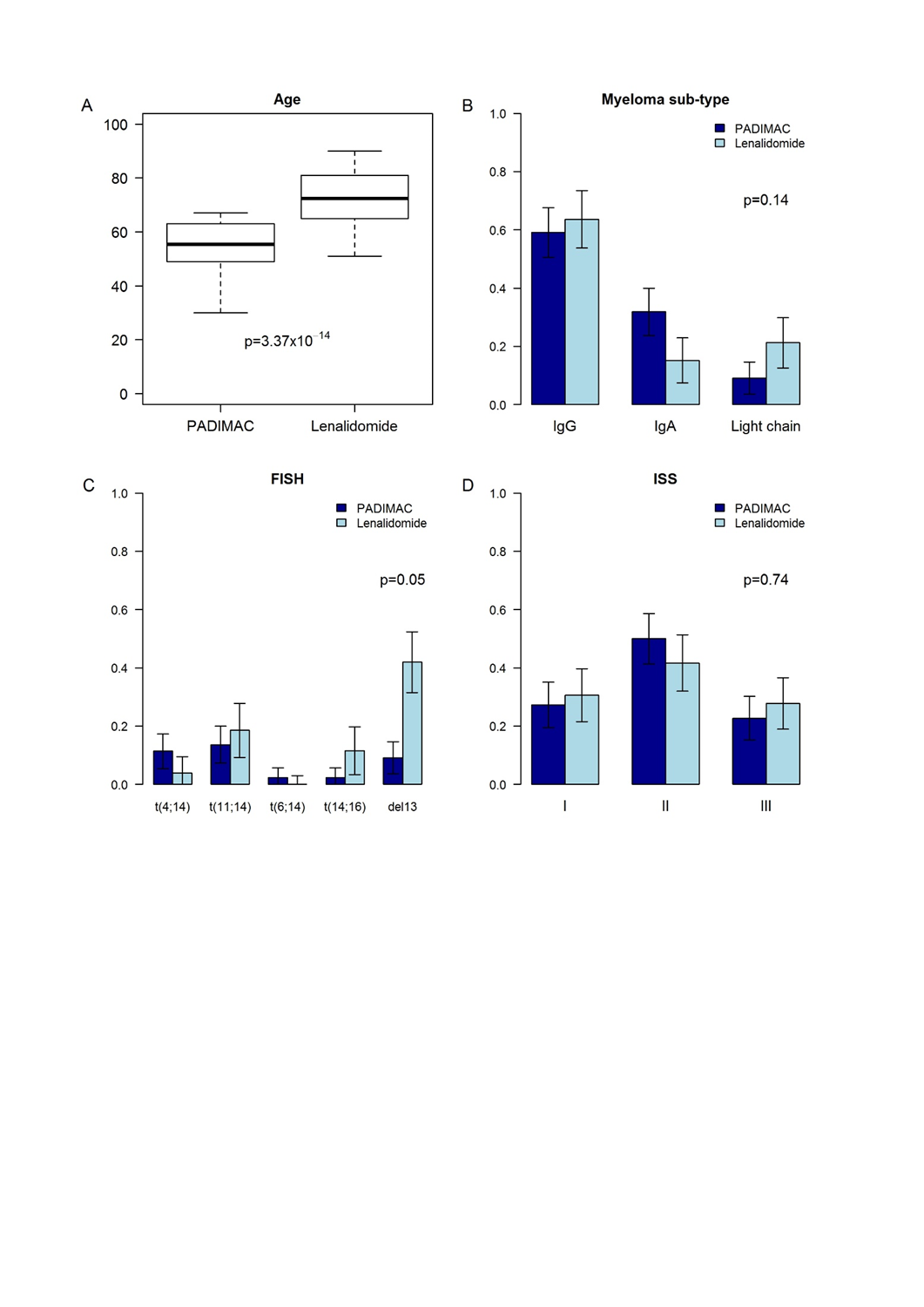
**Figure S2 – Output of synthetic annealing.** (A) Error rate of cross-validation (black lines and dots) and system temperature (red line) in arbitrary units (AU) plotted against increasing algorithm iterations. (B) Number of genes in signatures adopted by each algorithm iteration. The final signature with 100% correct assignment (iteration 414) contained seven genes (see dotted line).

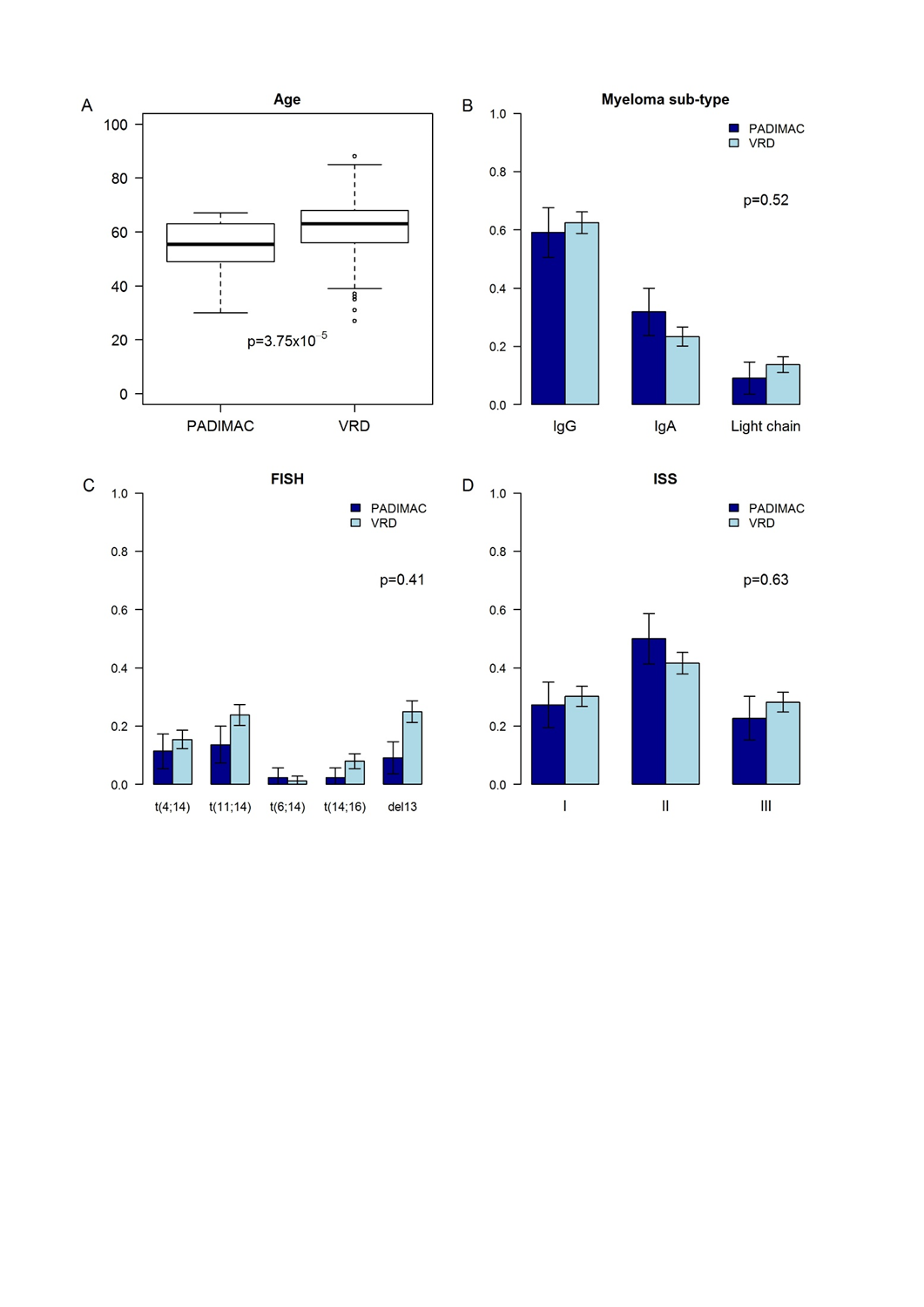
****

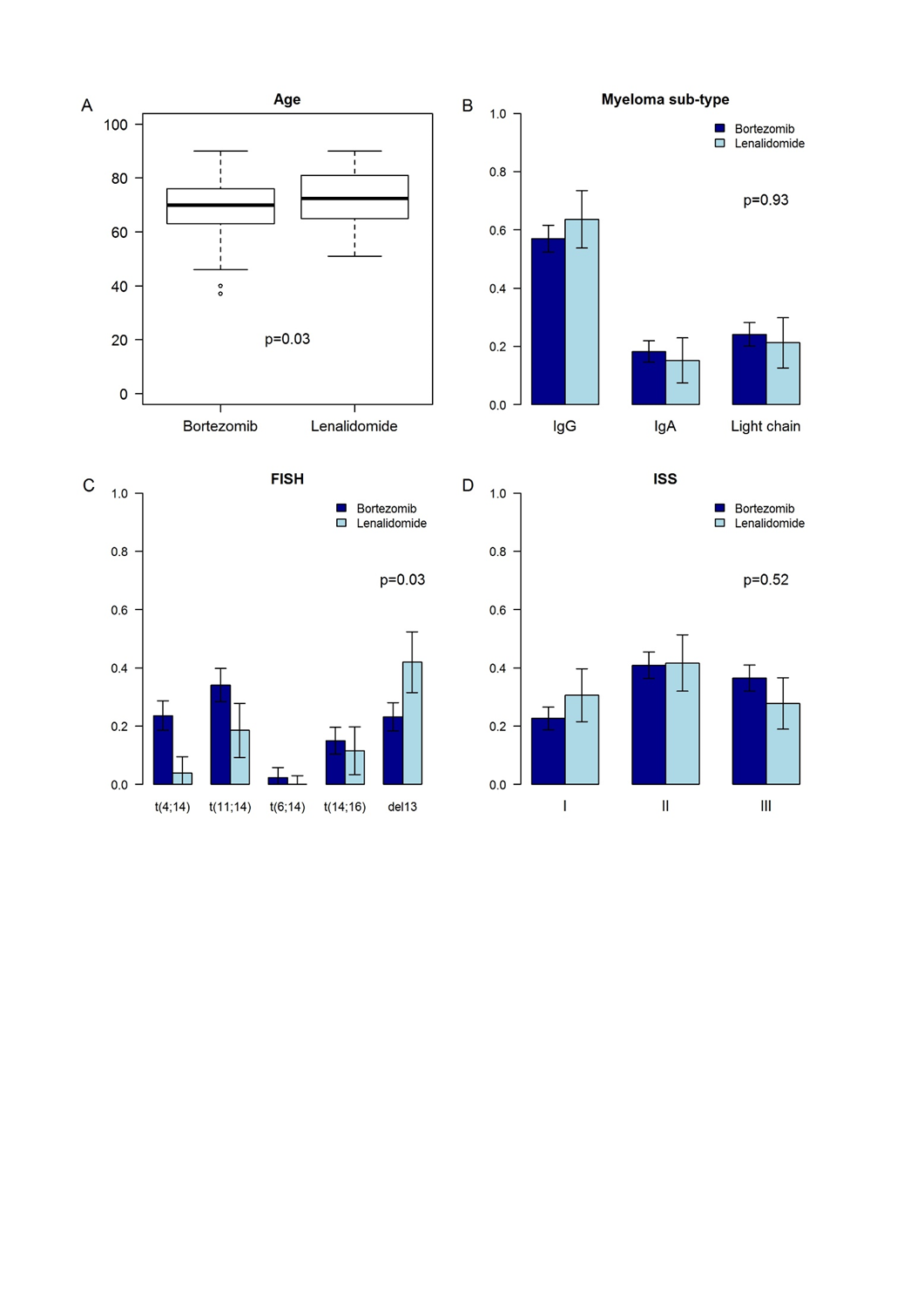
**Figure S3 – Process of ten-fold cross-validation in PADIMAC data.** Samples were split 90:10 into a training and test set. Within the training set, there was a further split into training and internal validation sets for the learning of the LMNN parameters. Data were recombined to establish a training set for learning the transformation matrix which was tested on the 10% of excluded data. Assignments were recorded and nine further mutually exclusive splits were made such that each sample was tested once. Assessments of performance (Matthews Correlation Coefficient and F-measure) were made on the recorded assignments across the whole dataset.

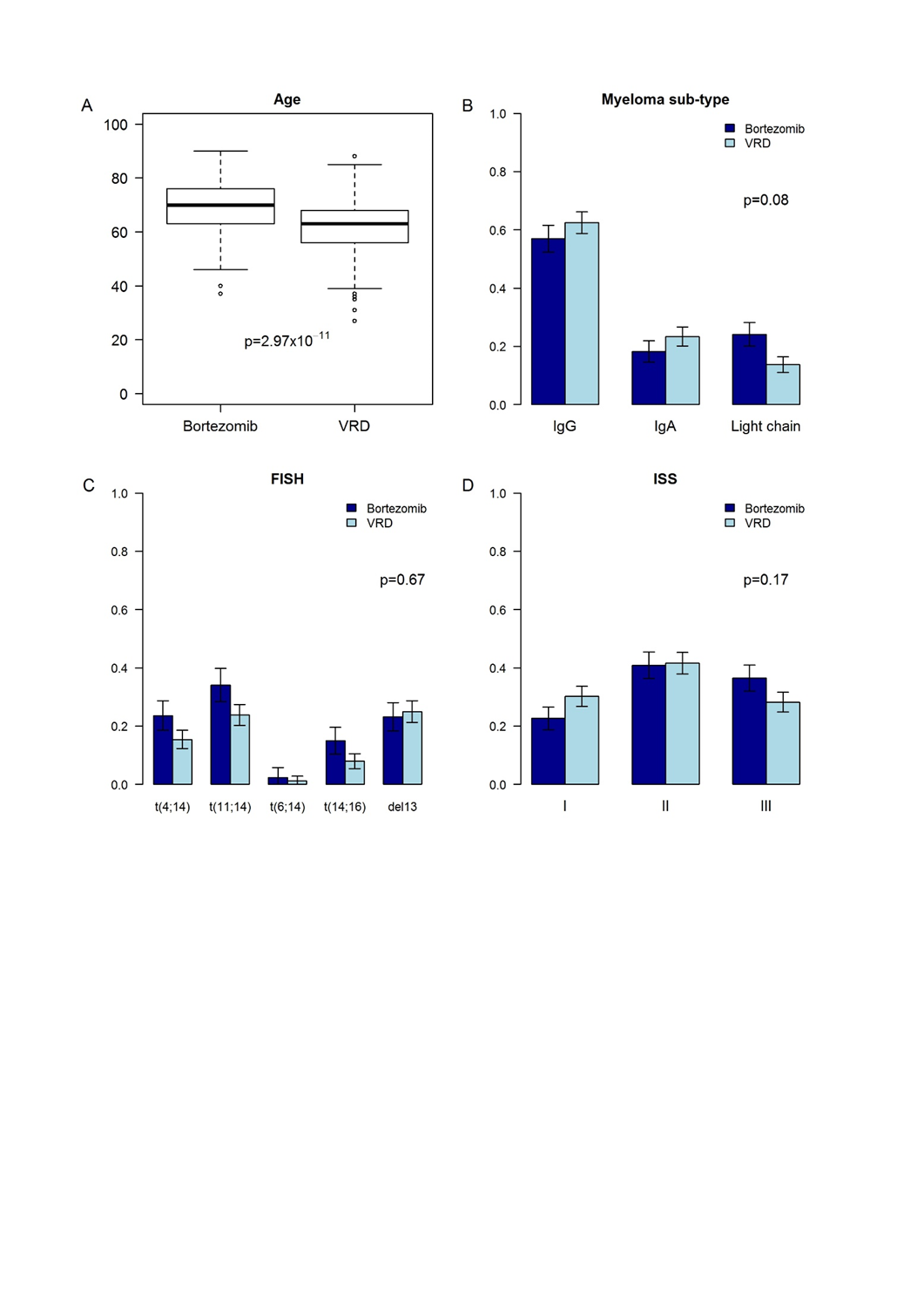
**Figure S4 – Performance of 4-11 gene signatures in ten-fold cross-validation within PADIMAC is better than expected by chance.** (A) and (B) Matthews Correlation Coefficients and F-measures of the signatures (Signatures) compared to permuted assignments (Null). Signature predictions were made with the Largest Margin Nearest Neighbors (LMNN) algorithm. 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.

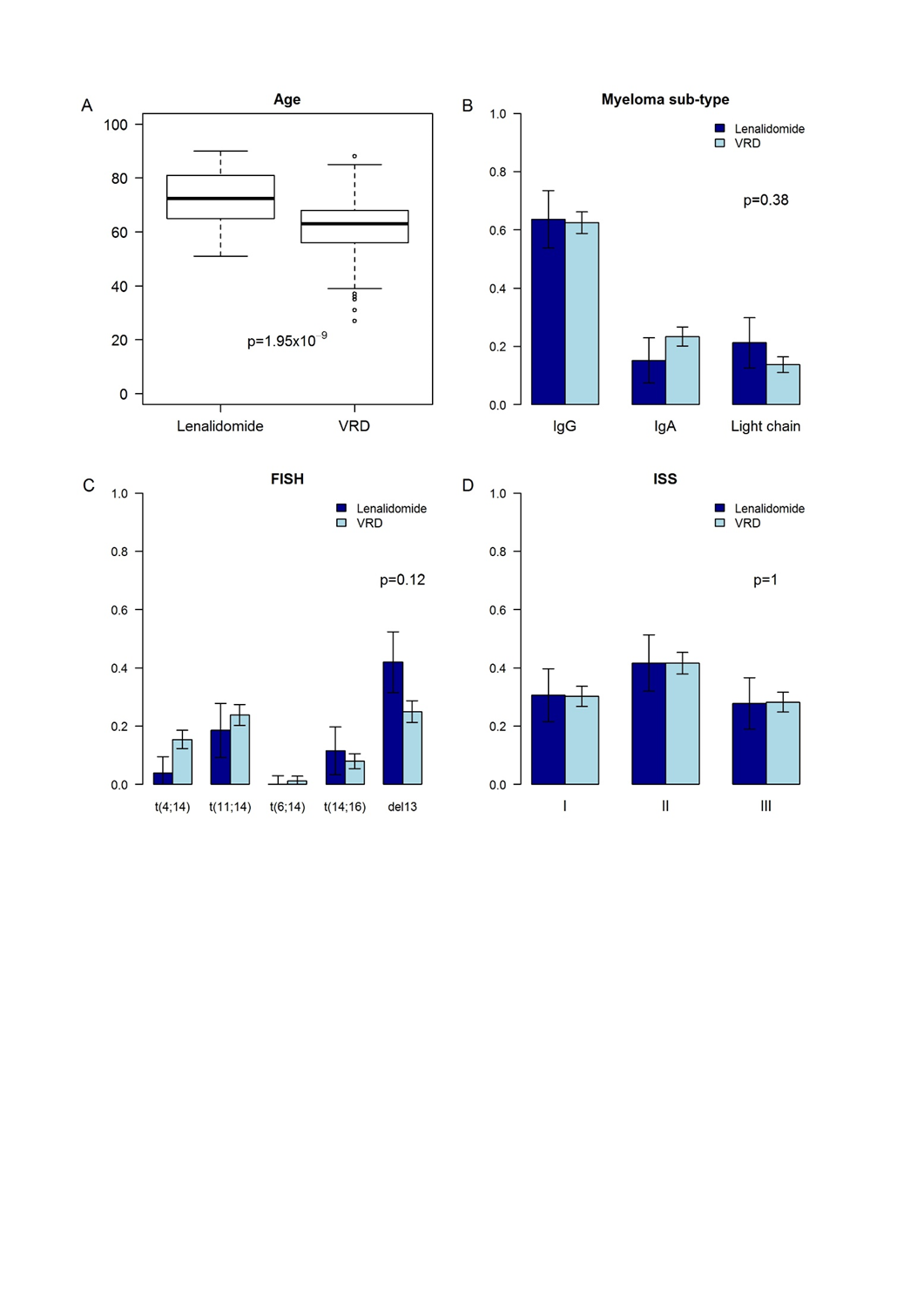
**Figure S5 – Comparisons of clinical features of PADIMAC patients with bortezomib-treated (no IMiD) patients in CoMMpass.** (A) Age of patients. The p-value is that of the Wilcoxon-Mann-Whitney test, under the null hypothesis that the distributions of ages are the same. (B) Subtype of patients. The p-value is that of the Fisher exact test. (C) Positive FISH findings. The p-value is that of the Fisher exact test. (D) International staging system (ISS) stage. The p-value is that of the Fisher exact test.

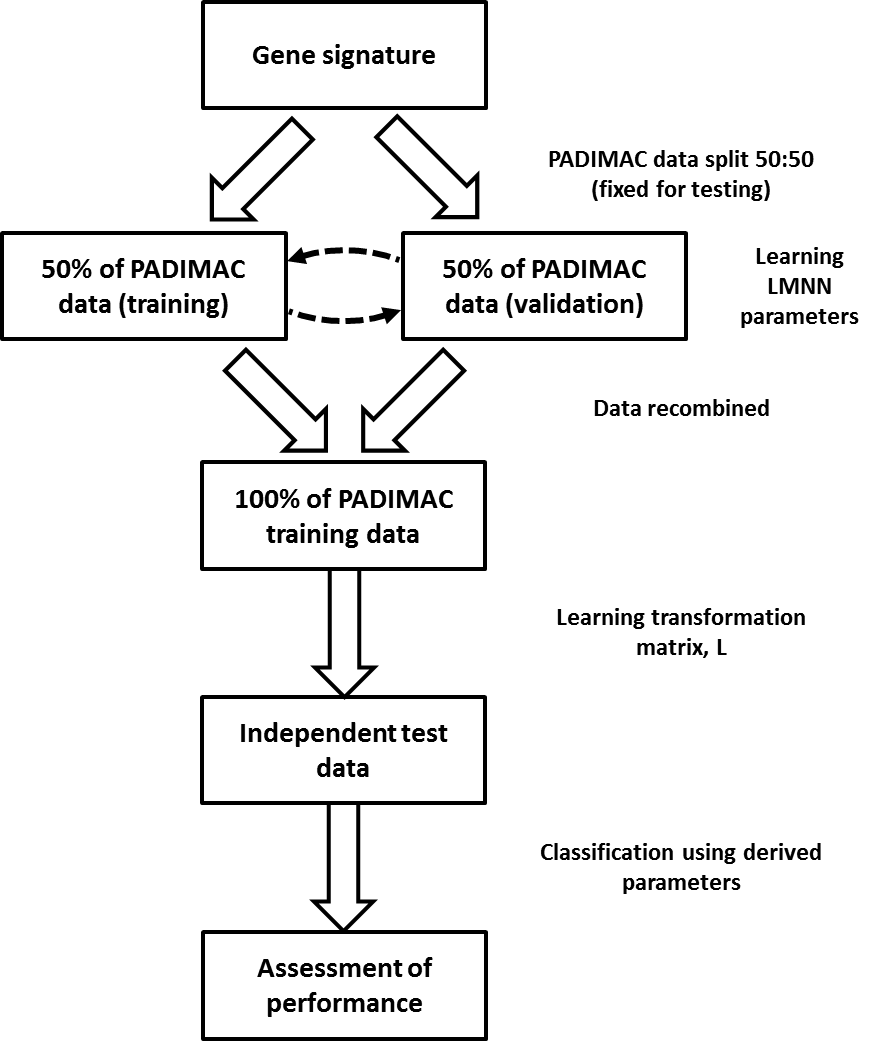
**Figure S6 – Comparisons of clinical features of PADIMAC patients with RD-treated patients in CoMMpass.** (A) Age of patients. The p-value is that of the Wilcoxon-Mann-Whitney test, under the null hypothesis that the distributions of ages are the same. (B) Subtype of patients. The p-value is that of the Fisher exact test. (C) Positive FISH findings. The p-value is that of the Fisher exact test. (D) International staging system (ISS) stage. The p-value is that of the Fisher exact test.

**Figure S7 – Comparisons of clinical features of PADIMAC patients with VRD-treated patients in CoMMpass.** (A) Age of patients. The p-value is that of the Wilcoxon-Mann-Whitney test, under the null hypothesis that the distributions of ages are the same. (B) Subtype of patients. The p-value is that of the Fisher exact test. (C) Positive FISH findings. The p-value is that of the Fisher exact test. (D) International staging system (ISS) stage. The p-value is that of the Fisher exact test.

**Figure S8 – Comparisons of clinical features of bortezomib-treated with RD-treated patients in CoMMpass.** (A) Age of patients. The p-value is that of the Wilcoxon-Mann-Whitney test, under the null hypothesis that the distributions of ages are the same. (B) Subtype of patients. The p-value is that of the Fisher exact test. (C) Positive FISH findings. The p-value is that of the Fisher exact test. (D) International staging system (ISS) stage. The p-value is that of the Fisher exact test.

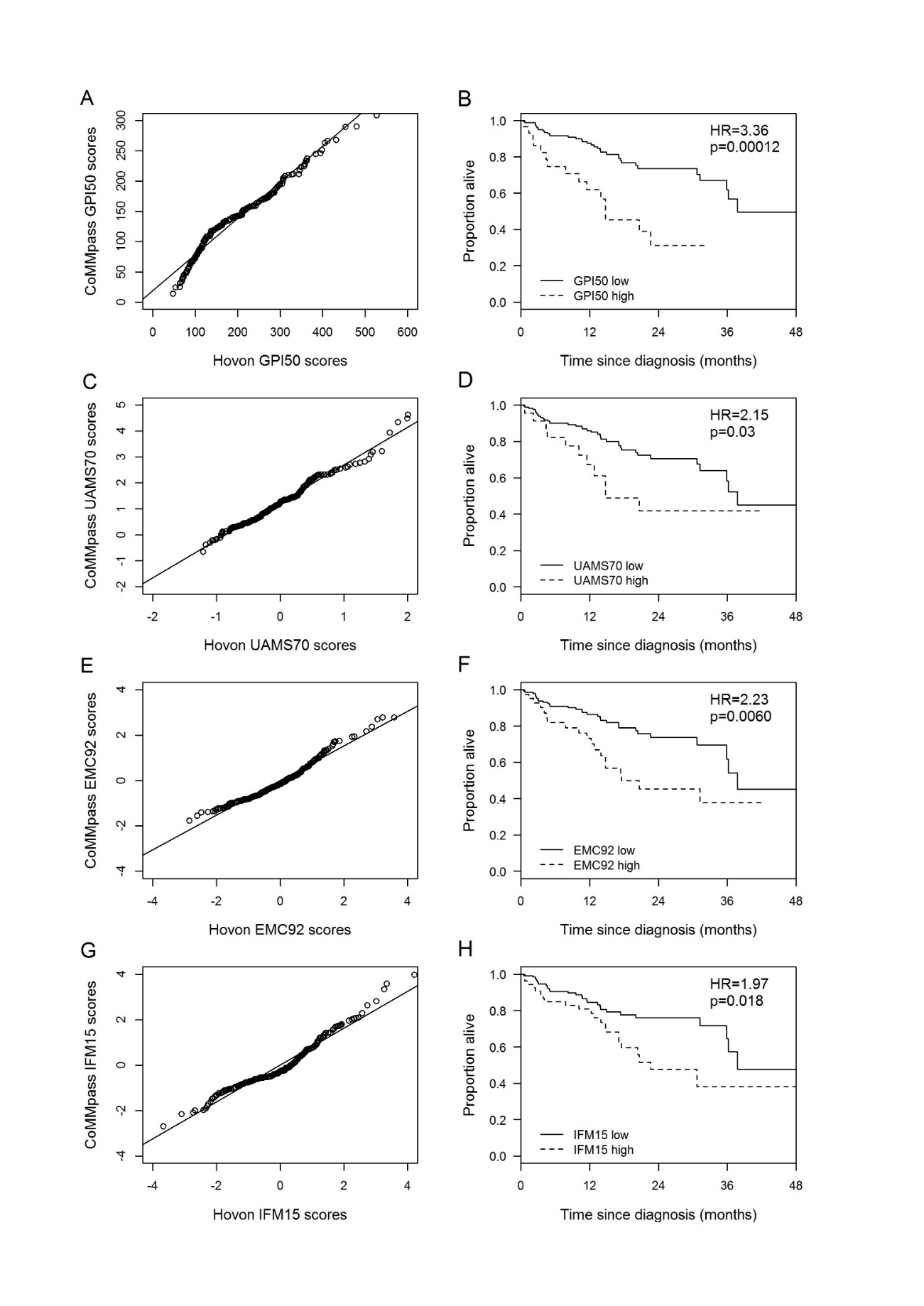
**Figure S9 – Comparisons of clinical features of bortezomib-treated with VRD-treated patients in CoMMpass.** (A) Age of patients. The p-value is that of the Wilcoxon-Mann-Whitney test, under the null hypothesis that the distributions ages are the same. (B) Subtype of patients. The p-value is that of the Fisher exact test. (C) Positive FISH findings. The p-value is that of the Fisher exact test. (D) International staging system (ISS) stage. The p-value is that of the Fisher exact test.

**Figure S10 – Comparisons of clinical features of lenalidomide-treated with VRD-treated patients in CoMMpass.** (A) Age of patients. The p-value is that of the Wilcoxon-Mann-Whitney test, under the null hypothesis that the distributions of ages are the same. (B) Subtype of patients. The p-value is that of the Fisher exact test. (C) Positive FISH findings. The p-value is that of the Fisher exact test. (D) International staging system (ISS) stage. The p-value is that of the Fisher exact test.

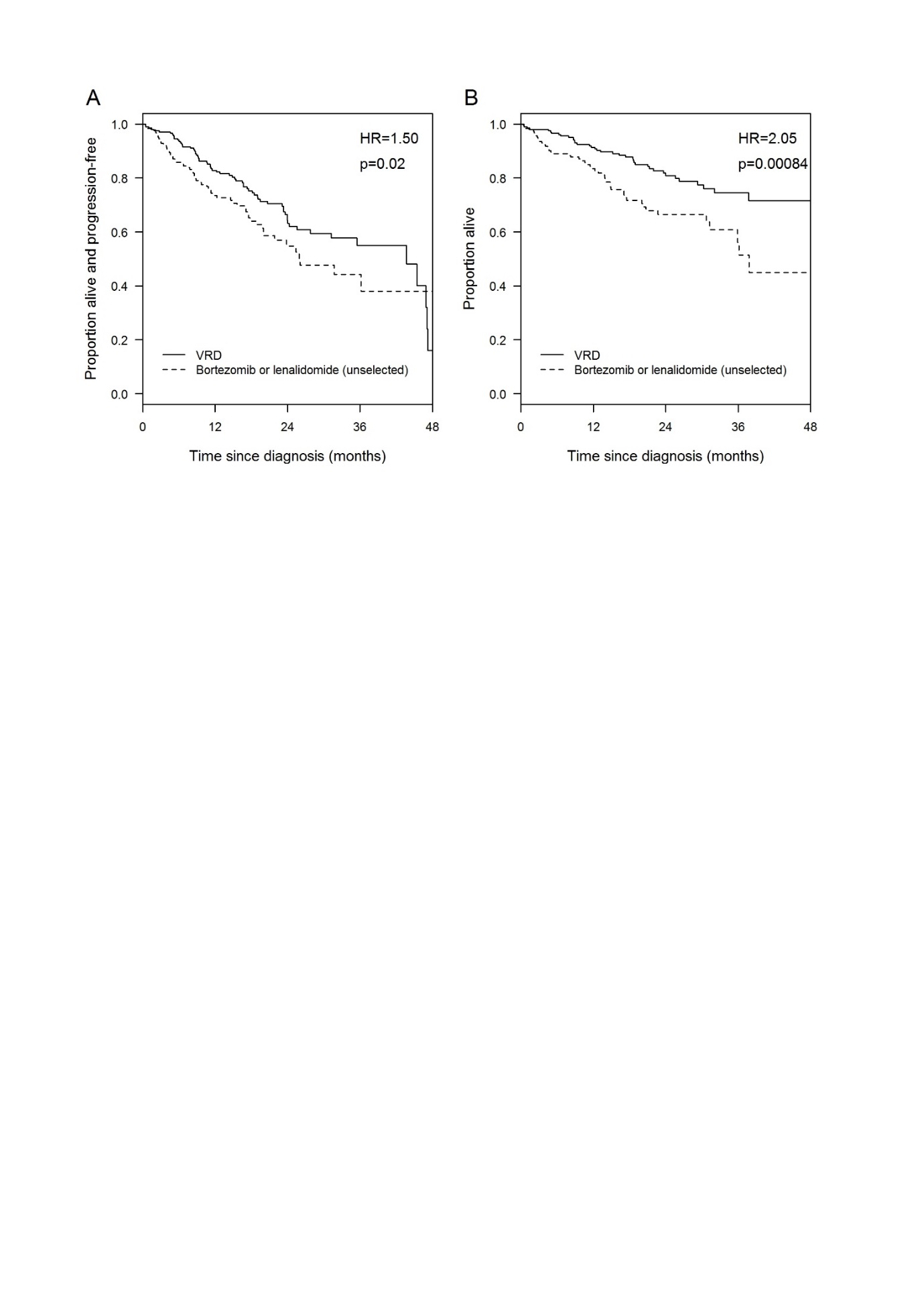
**Figure S11 – Overview of independent data testing.** The PADIMAC data were split into a 50:50 training/validation set and the LMNN parameters were learnt. The data were recombined to form a training dataset to learn the transformation matrix, L. Testing was then performed on independent datasets.

**A close up of a logo

Description generated with very high confidenceFigure S12 – Performance of the seven-gene signature in predicting response to bortezomib-based and RD therapy is robust to training/validation dataset splits.** Hazard ratios (HRs) for (A) disease progression or (B) overall survival of bortezomib- or RD-treated patients in CoMMpass who received the predicted best therapy compared to those who did not. Predictions were made by the seven-gene signature, trained in PADIMAC with repeated (n=100) training/validation splits (Signature). The HRs are compared with a null dataset of HRs obtained following 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.

****

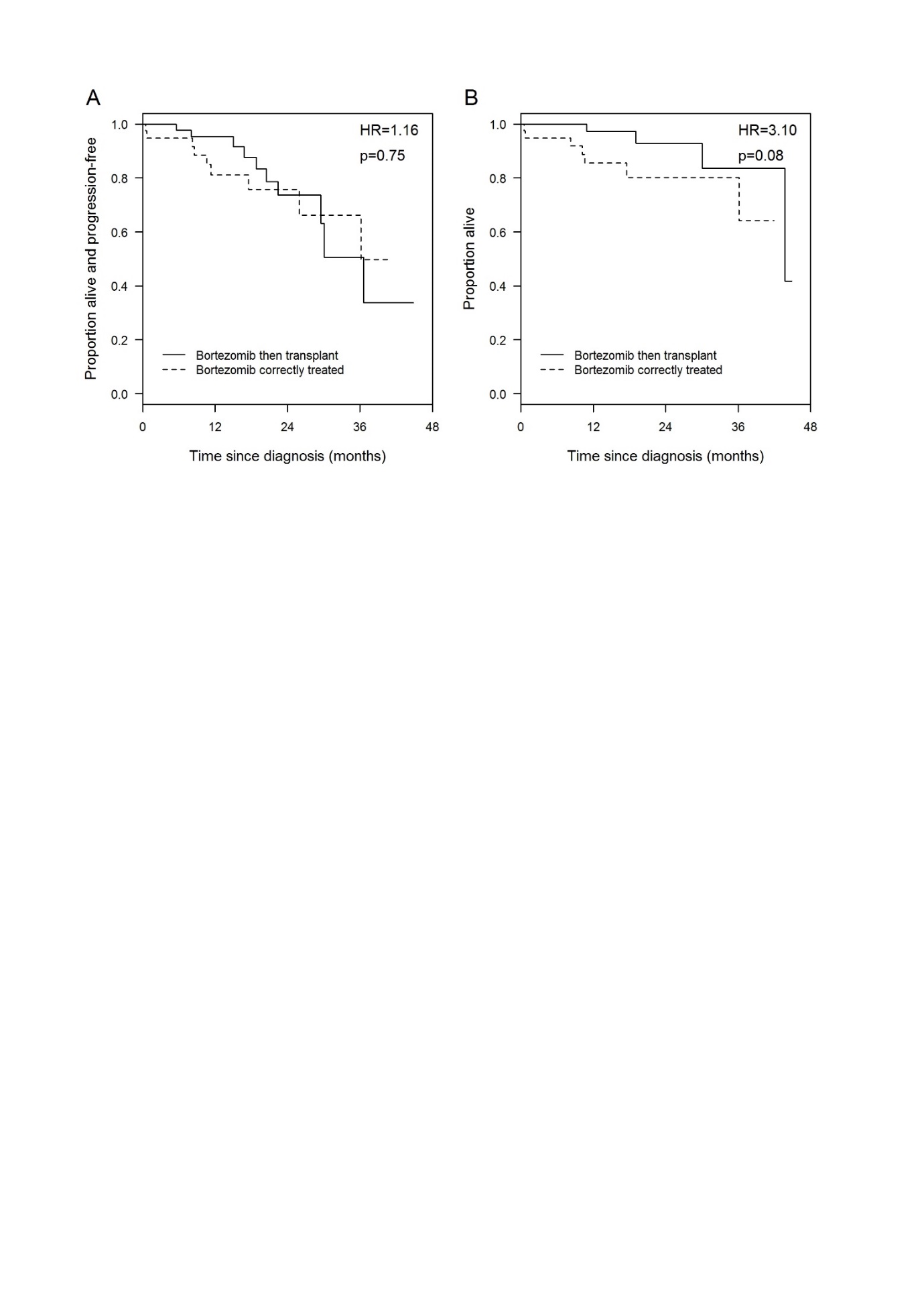
**Figure S13 – Prognostic signatures established with microarray data can be applied to RNA-Seq data.** Graphs on the left are Q-Q plots to compare the distribution of scores of each signature applied to the HOVON/GMMG microarray dataset (x-axis) and the CoMMpass RNA-Seq dataset (y-axis). Graphs on the right are Kaplan Meier curves showing overall survival in CoMMpass patients (n=187) according to whether the signatures assigned low (solid lines) or high (dotted lines) scores. The threshold for low/high score was determined by published microarray values converted to RNA-Seq values by the Q-Q plot.

****

**Figure S14 – Bortezomib, lenalidomide, and dexamethasone (VRD) therapy is superior to unselected bortezomib-based therapy or RD in the CoMMpass dataset.** (A)Kaplan-Meier plot showing the progression-free survival of patients (n=395) who received VRD (n=208; solid line) or who received bortezomib-based therapy or RD within CoMMpass, regardless of signature assignment (n=187; dotted line) following training in PADIMAC. (B)Kaplan-Meier plot showing the overall survival of patients (n=395) who received VRD (n=208; solid line) or who received bortezomib-based therapy or RD within CoMMpass, regardless of signature assignment (n=187; dotted line) following training in PADIMAC. P-values in both plots are from Cox regression analysis.

**A screenshot of a cell phone

Description generated with very high confidenceFigure S15 – Comparison of signature performance against and within VRD-treated patients.** (A)Kaplan-Meier plot showing the progression-free survival of patients (n=276) who received bortezomib/lenalidomide/dexamethasone (VRD; n=208; solid line) or who received bortezomib-based therapy or RD within CoMMpass and who received the correct therapy predicted by the seven-gene signature (n=68; dotted line) following training in PADIMAC. (B) Kaplan-Meier plot showing the progression-free survival 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; dotted line) by the seven-gene signature following training in PADIMAC. P-values in both plots are from Cox regression analysis.

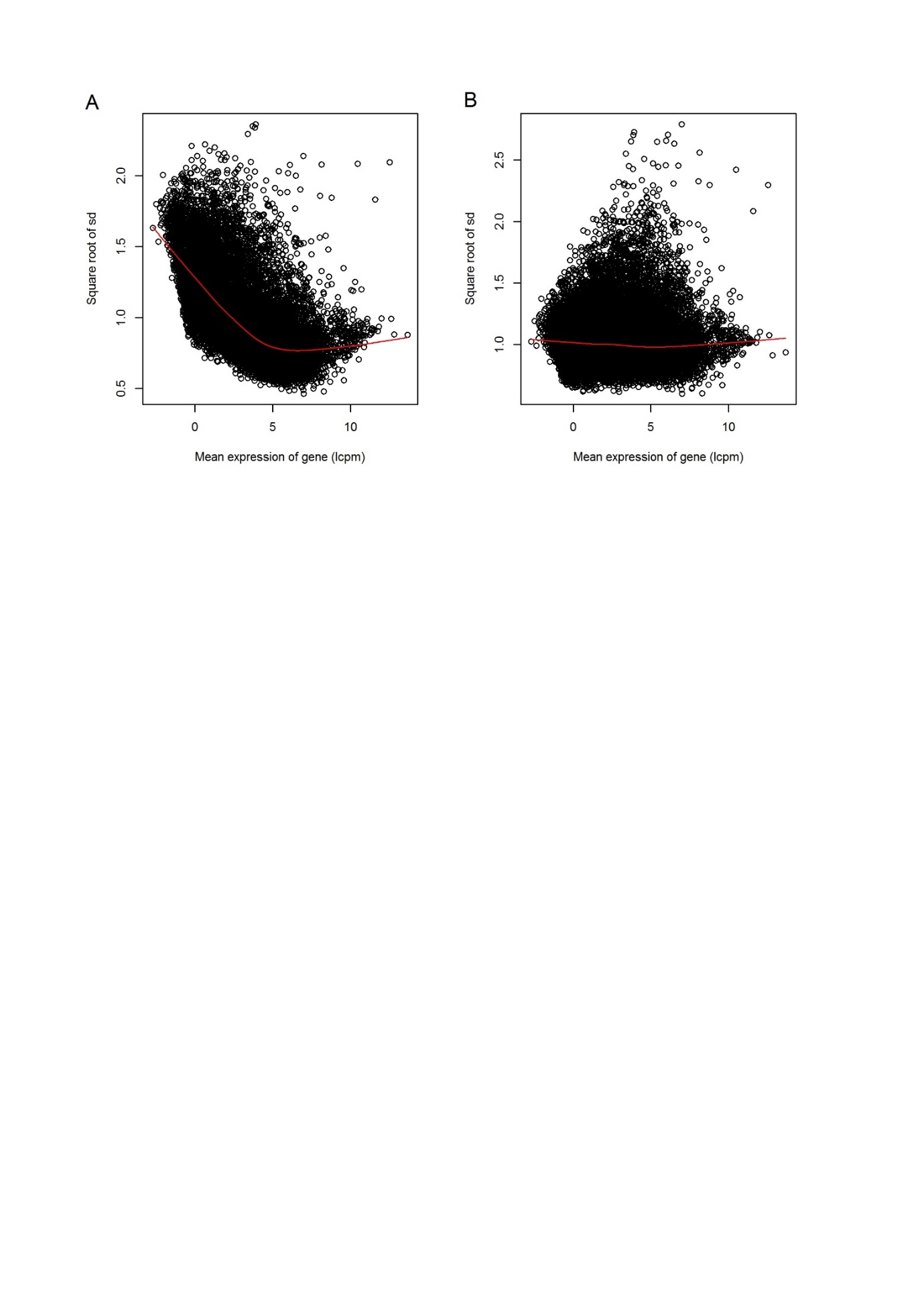


**Figure S16 – There is no significant difference in survival between patients correctly treated with bortezomib without autologous stem cell transplant (ASCT) and all patients treated with bortezomib followed by ASCT in CoMMpass.** (A)Kaplan-Meier plot showing the progression-free survival of patients (n=86) who received bortezomib-based therapy followed by ASCT (n=39; solid line) or who received bortezomib-based therapy within CoMMpass correctly according to the seven-gene signature (n=47; dotted line) following training in PADIMAC. (B) Kaplan-Meier plot showing the overall survival of patients (n=86) who received bortezomib-based therapy followed by ASCT (n=39; solid line) or who received bortezomib-based therapy within CoMMpass correctly according to the seven-gene signature (n=47; dotted line) following training in PADIMAC. P-values in both plots are from Cox regression analysis.

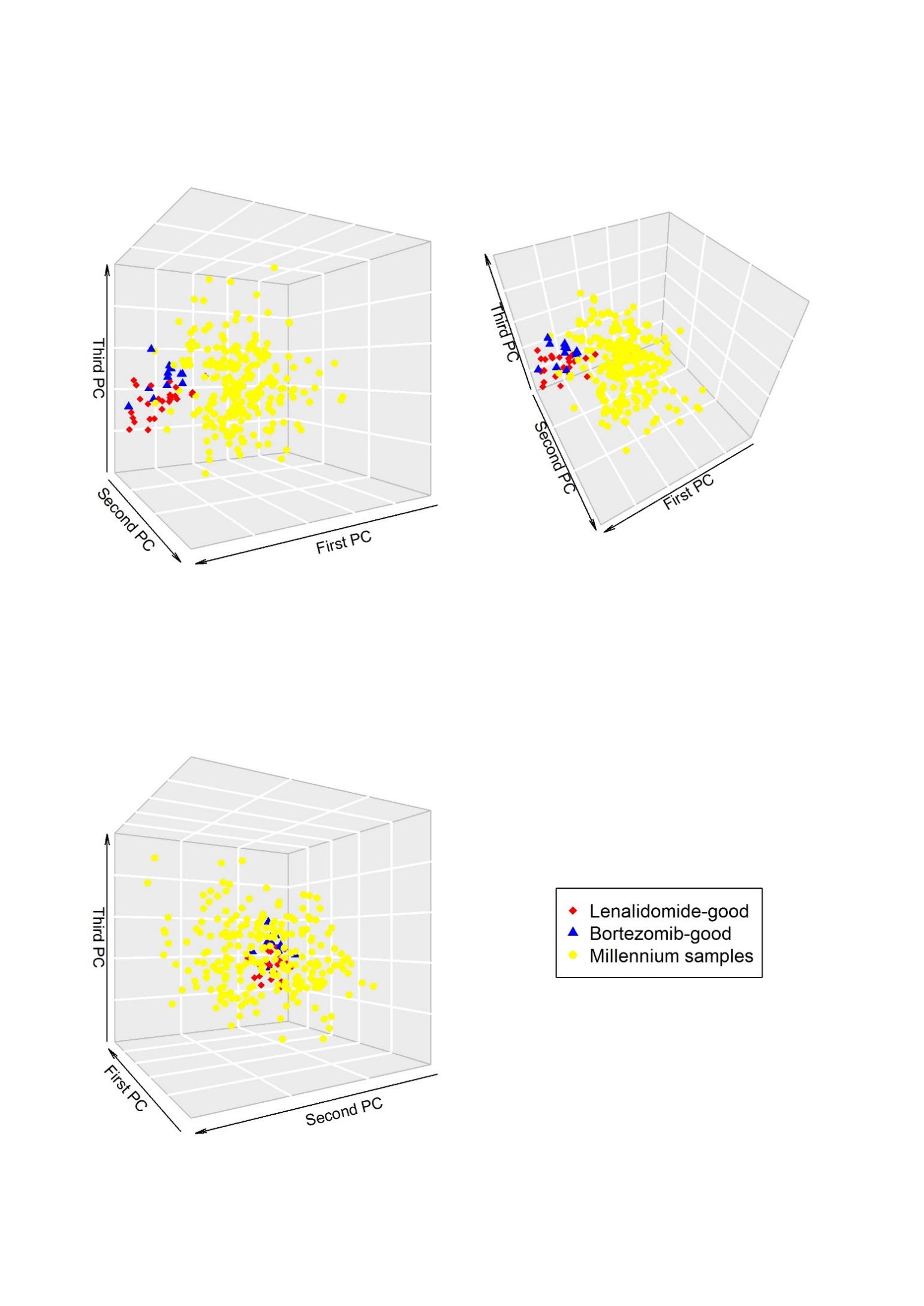
**A screenshot of a cell phone

Description generated with very high confidence**

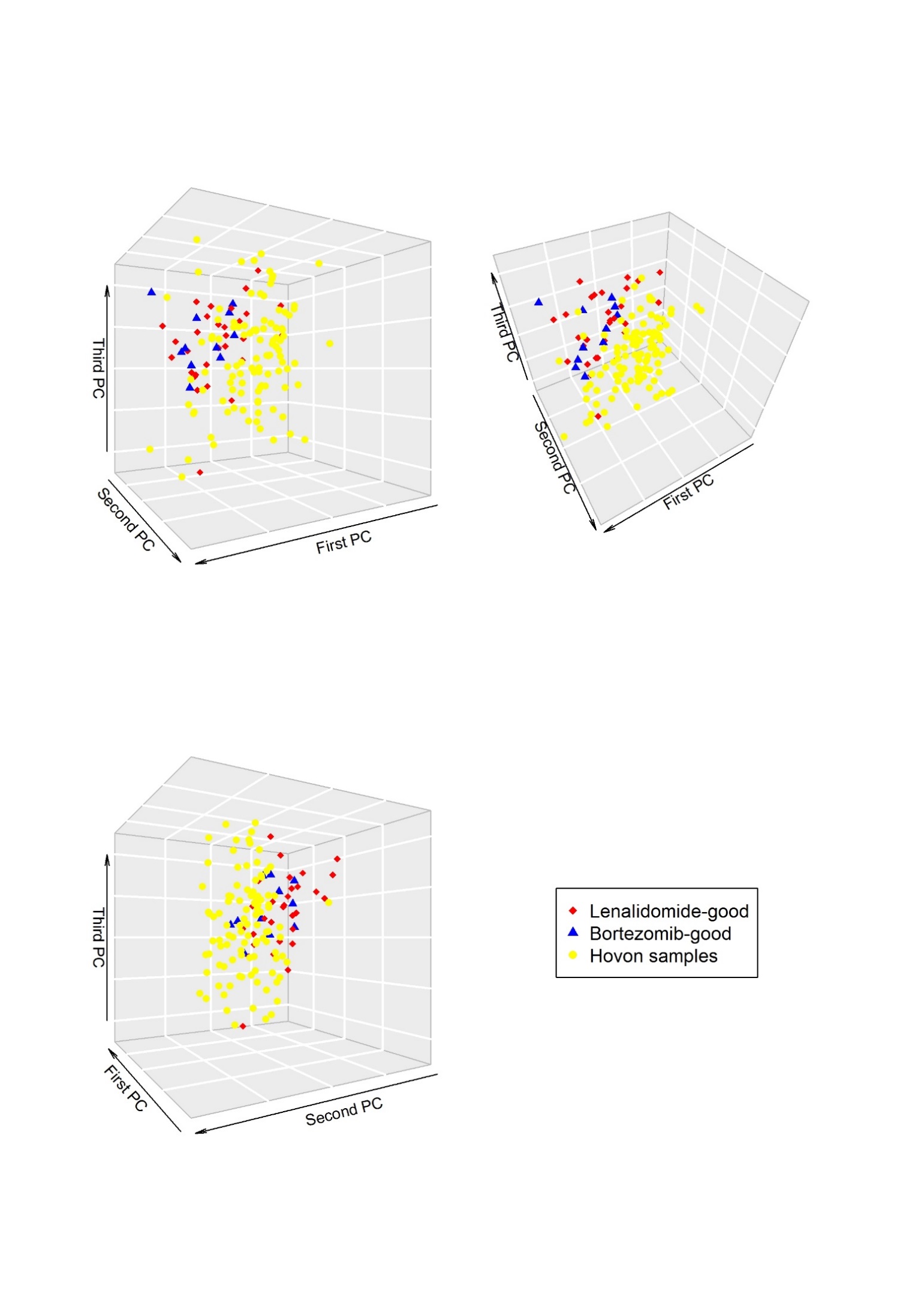
**Figure S17 – The seven-gene signature is not predictive for progression-free survival (PFS) in dexamethasone-treated patients in the Millennium dataset, but it does predict for overall survival (OS). (**A)Kaplan-Meier plot showing the PFS of patients who received dexamethasone (n=69) and who were classed as bortezomib-best (n=17; dotted line) or lenalidomide-best (n=52; solid line) by the seven-gene signature following training in PADIMAC. (B) Kaplan-Meier plot showing the PFS of patients who received dexamethasone (n=69) and who were classed as bortezomib-best (n=17; dotted line) or lenalidomide-best (n=52; solid line) by the seven-gene signature following training in PADIMAC. P-values in both plots are from Cox regression analysis.

****

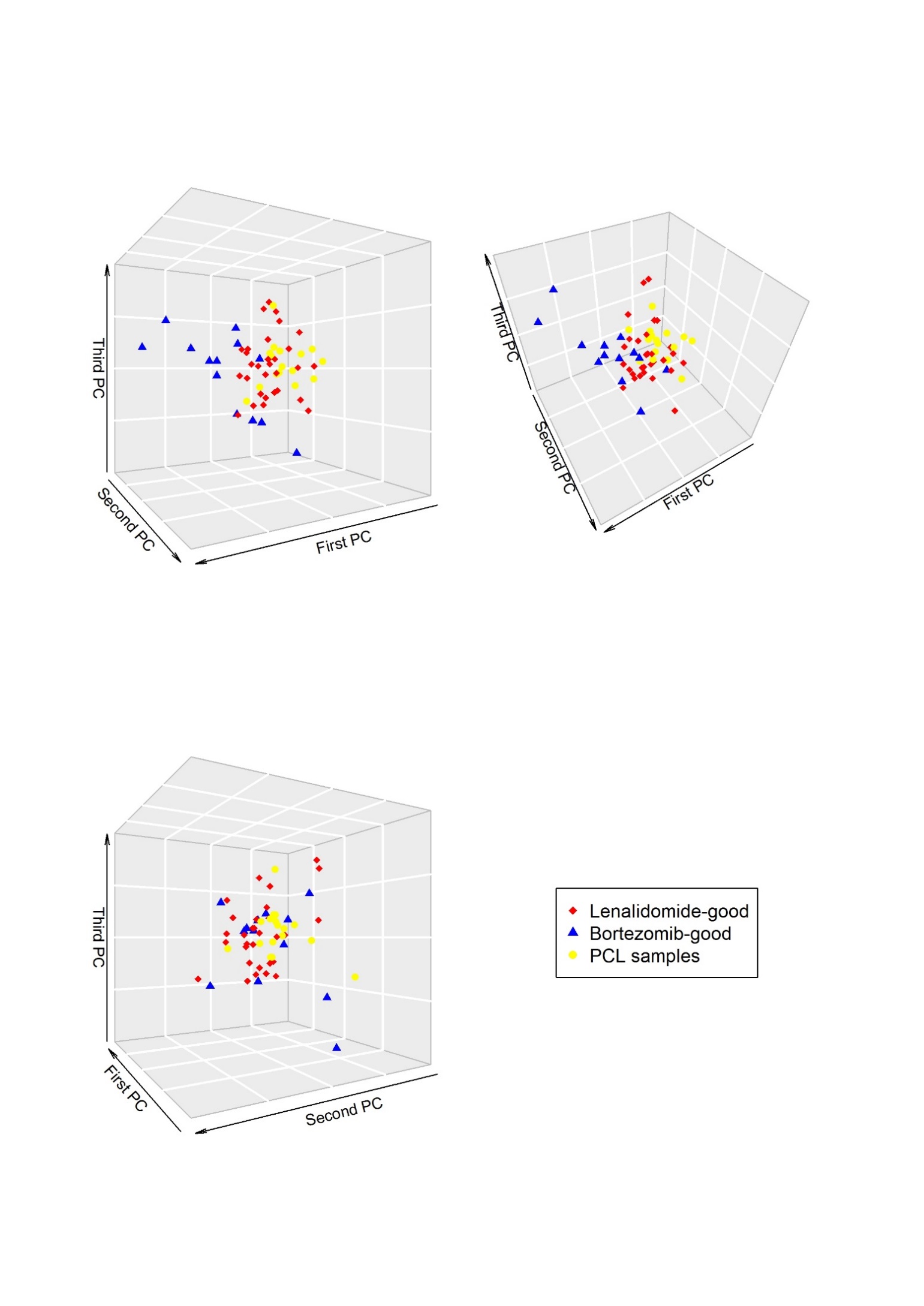
**Figure S18 – Correction of heteroscedasity as part of RNA-Seq normalization.** (A) There is inherent heteroscedasity in RNA-Seq expression data, which relate to the sampling nature of data generation. (B) We corrected this to make the data behave more like microarray data (see methods for details). In both plots, mean expression of each gene is plotted on the x-axis and is given by log counts per million (lcpm) and the square root of the standard deviation (sd) of gene expression is plotted on the y-axis. The red line is a locally weighted smoothing line.

****

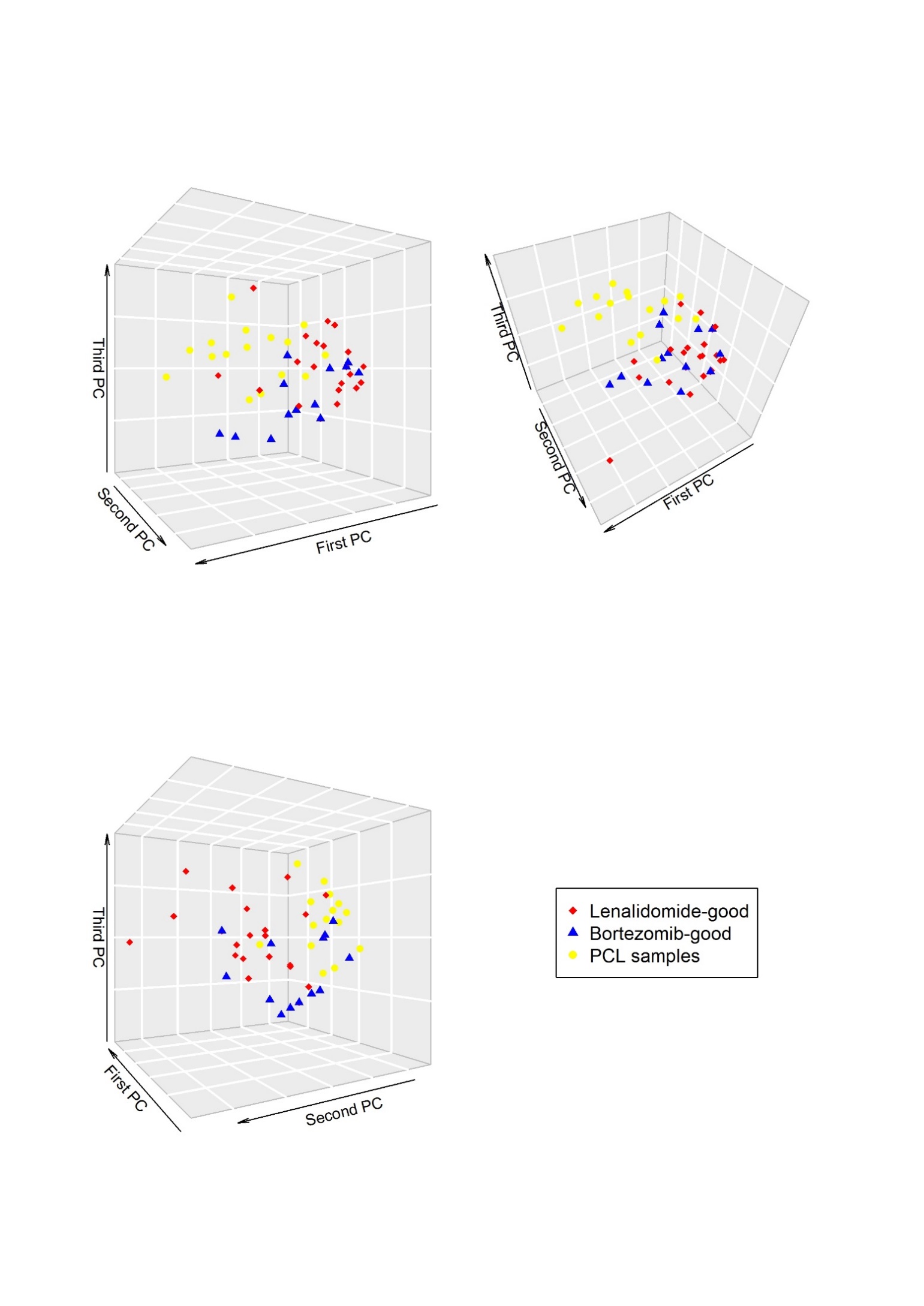
**Figure S19 – Standard normalization of Millennium microarray data followed by principal components analysis (PCA) does not imply a propensity towards a systematic classification bias.** Although the Millennium data and PADIMAC data cluster separately, there is no evidence that the Millennium data cluster exclusively with either the bortezomib-good or lenalidomide-good training samples.



**Figure S20 – Standard normalization of HOVON/GMMG microarray data followed by principal components analysis (PCA) does not imply a propensity towards a systematic classification bias.** Although the HOVON/GMMG data and PADIMAC data cluster separately, there is no evidence that the HOVON/GMMG data cluster exclusively with either the bortezomib-good or lenalidomide-good training samples.



**Figure S21 – Standard normalization of PCL data followed by microarray data suggests a propensity towards systematic bias of classification of PCL samples towards lenalidomide-good.** The PCL samples cluster almost exclusively with the lenalidomide-good samples. Geometric classification, regardless of the distance metric, would be anticipated to fail.



**Figure S22 – Quantile normalization of PCL data leads to a better admixture with PADIMAC data compared with standard normalization (see Figure S22).** Although the PCL data and PADIMAC data cluster separately, there is no evidence that the PCL data cluster exclusively with either the bortezomib-good or lenalidomide-good training samples.