

RESEARCH ARTICLE

Open Access



Predicting multi-level drug response with gene expression profile in multiple myeloma using hierarchical ordinal regression

Xinyan Zhang^{1†}, Bingzong Li^{2†}, Huiying Han³, Sha Song³, Hongxia Xu³, Yating Hong², Nengjun Yi^{4*} and Wenzhuo Zhuang^{3*}

Abstract

Background: Multiple myeloma (MM), like other cancers, is caused by the accumulation of genetic abnormalities. Heterogeneity exists in the patients' response to treatments, for example, bortezomib. This urges efforts to identify biomarkers from numerous molecular features and build predictive models for identifying patients that can benefit from a certain treatment scheme. However, previous studies treated the multi-level ordinal drug response as a binary response where only responsive and non-responsive groups are considered.

Methods: It is desirable to directly analyze the multi-level drug response, rather than combining the response to two groups. In this study, we present a novel method to identify significantly associated biomarkers and then develop ordinal genomic classifier using the hierarchical ordinal logistic model. The proposed hierarchical ordinal logistic model employs the heavy-tailed Cauchy prior on the coefficients and is fitted by an efficient quasi-Newton algorithm.

Results: We apply our hierarchical ordinal regression approach to analyze two publicly available datasets for MM with five-level drug response and numerous gene expression measures. Our results show that our method is able to identify genes associated with the multi-level drug response and to generate powerful predictive models for predicting the multi-level response.

Conclusions: The proposed method allows us to jointly fit numerous correlated predictors and thus build efficient models for predicting the multi-level drug response. The predictive model for the multi-level drug response can be more informative than the previous approaches. Thus, the proposed approach provides a powerful tool for predicting multi-level drug response and has important impact on cancer studies.

Keywords: Gene expression, Hierarchical ordinal regression, Multiple myeloma, Multi-level drug response, Prediction

Background

Multiple myeloma (MM) is a malignant plasma cell disorder characterized by the proliferation in the bone marrow of clonal plasma cells [1, 2]. Around 30,280 new multiple myeloma cases are expected to be diagnosed in

2017 [3]. Meanwhile, MM is an incurable disease using conventional treatment, which results in a median overall survival of 3 to 4 years [4, 5]. Nevertheless, treatment outcome in MM has improved significantly in the last decade, partially due to the introduction of novel agents, such as the proteasome inhibitors (e.g. bortezomib) and immunomodulatory drugs (e.g. thalidomide) [6]. In spite of that, the heterogeneity exists in the patients' response to those new treatments and molecular features responsible for the variability in response remain undefined [7–9]. It urges more efforts to identify

* Correspondence: nyj@uab.edu; zhuangwenzhuo@suda.edu.cn

[†]Equal contributors

⁴Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL 35294, USA

³Department of Cell Biology, School of Biology & Basic Medical Sciences, Soochow University, Suzhou, China

Full list of author information is available at the end of the article



biomarkers from numerous molecular features and build predictive models for identifying patients that can benefit from a certain treatment scheme [7].

MM, like other cancers, is caused by the accumulation of genetic abnormalities [2, 10]. Various molecular analyses suggest that myeloma is composed of distinct subtypes that have different molecular pathologies and prognosis [10]. For example, cytogenetic studies reveal that 60 to 80% of myeloma cases reveal chromosomal translocations involving the immunoglobulin heavy (IgH) locus [10]. The most prevalent of these translocations are t(11;14)(q13;q32) and t(4;14)(p16;q32), where the former has better survival than the latter. Another chromosomal translocation in t(8;14)(q24;q32), causing MYC activation, is considered as a secondary hit. Other genetic abnormalities include mutations and copy number changes. Mutational spectrum reveals a heterogeneous landscape with few recurrently affected genes. Only three genes have been reported in more than 10% patients, including *KRAS*, *NRAS*, and *FAM46C* [11–14]. For copy number changes, the most common ones being gains are on 1q, 3p, 6p, 9p, 11q, 19p, 19q and 21q along with deletions of 1p, 4q, 16q and 22q, among which candidate oncogenes and tumor suppressors have been identified [15–18]. Thus, it is anticipated that identifying and applying molecular biomarkers to predict clinical response to drugs will help to provide more precise prognostic and predictive classifiers for a specific therapy in MM.

With the emergence of high-throughput sequencing technology, it is expected that the number of biomarkers will rise. The markers in predicting drug response will become more reproducible as well [19, 20]. However, the effect sizes of the discovered markers are usually small, which only could contribute a relatively trivial portion to the drug response since it is typically a complex trait, generally influenced by many genomic and environmental factors [19]. Thus, predictive modeling with multiple markers should be used to predict complex traits, such as drug response [19].

Distinct gene expression profiling is believed to be associated with the drug response variability of bortezomib, leading to various disease prognoses [10]. The relationships between the heterogeneity of drug response in bortezomib or its combined therapy with the genomic background of multiple myeloma patients have been investigated [2, 10]. Mulligan et al. [10] generated gene expression data from a national and international phase 2 and 3 clinical trials of bortezomib to develop a genomic classifier for prediction of drug response in relapsed MM. Terragna et al. [2] analyzed the gene expression from MM patients to explore predictors of bortezomib-thalidomide-dexamethasone (VTD) first-line therapy. According to European Group for Bone

Marrow Transplantation criteria, drug responses in MM were classified as achieving complete response (CR), partial response (PR), minimal response (MR), no change (NC) and Progressive Disease (PD) [21]. However, in Mulligan et al. [10], the five-level ordinal drug response was categorized as a binary response where only responsive and non-responsive groups were considered. Terragna et al. [2] focused on CR vs non-CR groups by converting the ordinal outcome to a binary outcome. To provide more informative prediction and more efficiently identify biomarkers, it is desirable to analyze multi-level drug response, rather than combining the response to two groups. To address this shortcoming in the previous analyses, we here present a novel approach to identify significantly associated biomarkers and develop genomic classifier using hierarchical ordinal logistic regression. We apply our approach to two public available datasets [2, 10]. Our results show that our hierarchical ordinal regression approach is able to identify genes associated with the multi-level response and to generate predictive models for predicting the multi-level response.

Methods

Datasets acquisition for ordinal response prediction

The gene expression datasets analyzed to predict the drug response were acquired from two independent clinical trials. The two datasets are publically available from GEO under accession number [GEO: GSE9782] and [GEO: GSE68871]. They were originally published in Mulligan et al. [10] and Terragna et al. [2]. Mulligan et al. [10] recruited patients ($n = 169$) with relapsed myeloma enrolled in phase 2 and phase 3 clinical trials of bortezomib, whose pretreated tumor samples were further analyzed for genomic profiling with consent. Myeloma samples were collected prior to enrollment in clinical trials of bortezomib and samples were subject to replicate gene expression profiling using the Affymetrix 133A microarray (22,283 probes). Terragna et al. [2] focused primarily on treating the new MM patients ($n = 118$) with the induction therapy of VTD. The gene expression profiling (54,677 probes) was carried out in the Affymetrix Human Genome U133 Plus 2.0 Array. We standardized all the probes for statistical analysis in both datasets.

Outcome definitions

The original datasets contained drug response and overall survival as two outcomes. We here focused on developing genomic classifier for the drug response. In Mulligan et al. [10], patients were classified as achieving complete response (CR), partial response (PR), minimal response (MR), no change (NC) and Progressive Disease (PD) according to European Group for Bone Marrow Transplantation criteria [21]. In Terragna et al. [2], patients' drug responses were classified as five

categories: complete response (CR), near complete response (nCR), very good partial response (VGPR), partial response (PR) and stable disease (SD). For both datasets, the five-level ordinal drug response was used in our analysis. In the meantime, we combined the five-level drug response to a new three-level drug response in both datasets to avoid low frequencies in certain levels. For Mulligan et al. [10], we combined PD and NC to a new level, and PR and MR as another new level. For Terragna et al. [2], we combined SD and PR as a new level, and VGPR and nCR as another new level. Both the original five-level and the new three-level outcomes were separately analyzed for these two datasets.

Ordinal drug response prediction modeling

Let y_i be the ordinal outcome for which there exists a clear ordering of the response categories, and X_{ij} the gene expression profile for the i th individual and j th probe in the data of sample size n with a total number J of probes. For notational convenience, we code the ordinal outcome as the integers 1, 2, ..., K , with K being the number of categories.

Univariate ordinal logistic regression to rank top probes

It will not be efficient to include all the genes in a predictive model, due to the large number of genes. We thus first use the univariate ordinal logistic regression to filter top q associated probes of the gene expression profile with the ordinal outcome. For the j -th gene expression X_{ij} , the univariate ordinal logistic regression is expressed as:

$$\Pr(y_i = k) = \begin{cases} 1 - \text{logit}^{-1}(X_{ij}\alpha - c_{1j}) & \text{if } k = 1 \\ \frac{\text{logit}^{-1}(X_{ij}\alpha - c_{(k-1)j}) - \text{logit}^{-1}(X_{ij}\alpha - c_{kj})}{\text{logit}^{-1}(X_{ij}\alpha - c_{(k-1)j})} & \text{if } 1 < k < K \\ \text{logit}^{-1}(X_{ij}\alpha - c_{(k-1)j}) & \text{if } k = K \end{cases}$$

where c_{kj} denoted cut-points or thresholds, are constrained to increase, $c_{1j} < \dots < c_{(K-1)j}$. We then select the top q associated probes based on the p -value for testing the hypothesis $H_0: \alpha = 0$.

Predictive modeling for genomic classifiers

We use all the q selected probes to build a multivariable ordinal model for predicting the multi-level response, i. e.

$$\Pr(y_i = k) = \begin{cases} 1 - \text{logit}^{-1}(X_i\beta - c_1) & \text{if } k = 1 \\ \frac{\text{logit}^{-1}(X_i\beta - c_{k-1}) - \text{logit}^{-1}(X_i\beta - c_k)}{\text{logit}^{-1}(X_i\beta - c_{k-1})} & \text{if } 1 < k < K \\ \text{logit}^{-1}(X_i\beta - c_{k-1}) & \text{if } k = K \end{cases}$$

where the vector X_i includes the expression measures of the q genes, and $\beta = (\beta_1, \dots, \beta_q)^T$ is a vector of the effects. With hundreds or tens of correlated top associated probes, however, the standard ordinal logistic regression may fail, due to the non-identifiability and overfitting. To overcome the problems, we use an

appropriate prior distribution to constrain the coefficients to lie in reasonable ranges and allow the model to be reliably fitted and to identify important predictors [22, 23]. We employ the commonly used Cauchy prior distribution on the coefficients in the ordered logistic regression:

$$p(\beta_j) = \text{Cauchy}(0, s) = \frac{1}{\pi s} \frac{1}{1 + \beta_j^2/s^2}$$

The scale parameter s controls the amount of shrinkage in the coefficient estimate; smaller s induces stronger shrinkage and forces more coefficients towards zero. For the cut-points parameters, we use a uniform prior. We have developed a quasi-Newton algorithm (BFGS) for fitting the hierarchical ordinal model by finding the posterior mode of the parameters (β, c) , i.e., estimating the parameters by maximizing the posterior density.

Our algorithm has been implemented in our R package BhGLM, which is freely available from the website <http://www.ssg.uab.edu/bhglm/> and the public GitHub repository <https://github.com/abbyyan3/BhGLM> that includes R codes for examples.

Assessing the performance of a fitted hierarchical ordinal logistic regression

After fitting a hierarchical ordinal model, we obtain the estimate $(\hat{\beta}, \hat{c})$ and can estimate the probabilities: $p_{ik} = \Pr(y_i = k | X_i\hat{\beta}, \hat{c}), i = 1, \dots, n; k = 1, \dots, K$. Denote $y_{ik} = I(y_i = k)$ as the binary indicator response for the k -th category. We can evaluate the performance using several measures:

- (1) *Deviance*: $d = -2\sum_{i=1}^n \log p_{ik}$. Deviance measures the overall quality of a fitted model;
- (2) *AUC* (area under the ROC curve). We can calculate AUC for the k -th category using $\{y_{ik}, p_{ik}; i = 1, \dots, n\}$ as usual. Then the AUC for all the categories is defined as $\frac{1}{K} \sum_{k=1}^K \text{AUC}_k$.
- (3) *MSE* (mean squared error). MSE is defined as: $MSE = \frac{1}{K} \sum_{k=1}^K [\frac{1}{n} \sum_{i=1}^n (y_{ik} - p_{ik})^2]$.
- (4) *Misclassification*. The misclassification is defined as: $MIS = \frac{1}{K} \sum_{k=1}^K [\frac{1}{n} \sum_{i=1}^n I(|y_{ik} - p_{ik}| > 0.5)]$, where $I(|y_{ik} - p_{ik}| > 0.5) = 1$ if $|y_{ik} - p_{ik}| > 0.5$, and $I(|y_{ik} - p_{ik}| > 0.5) = 0$ if $|y_{ik} - p_{ik}| \leq 0.5$;

To evaluate the predictive performance of the model, we use the pre-validation method, a variant of cross-validation [24, 25], by randomly splitting the data to H subsets of roughly the same size, and using $(H - 1)$ subsets to fit a model. We then calculate the measures described above with h th subset and cycle through all H subsets to get the averaged measurements to evaluate the predictive performance. To get stable results, we can

run H -fold cross-validation multiple times and use the average of the measure over the repeats to assess the predictive performance. We also can use leave-one-out cross-validation (i.e., $H = n$) to obtain unique result. In this study, 10-fold cross-validation with 10 repeats and leave-one-out cross-validation were both utilized. Deviance AUC, MSE and misclassification rate were all reported.

Selecting optimal scale values

The performance of the hierarchical ordinal model can depend on the scale parameter in the Cauchy prior. We fit a sequence of models with different scales ranging from 0.01 to 1 by 0.01, from which we can choose an optimal one based on the criteria described above.

Selecting the optimal number of q probes

To select an optimal number of top q probes, we fit a sequence of models with a different number of probes with the options from 30 and 50 to 500 by 50. The number q will be determined based on the predictive performance of the hierarchical ordinal logistic regression by evaluating the deviance of the models. The chosen top q probes will be identified as associated significant biomarkers and present in heatmaps for visual examination.

Results

Data summary

There were 169 samples analyzed in the dataset, with a total of 22,283 gene expression probes in Mulligan et al. [10]. In Terragna et al. [2], we analyzed 118 samples with a total of 54,677 gene expression probes. The details of both studies and the frequencies of the five-level ordinal drug response outcomes were summarized in Table 1. To avoid low frequencies in some levels for both datasets, we combined the five-level drug response to a new three-level drug response. By combining drug response in Mulligan et al. [10] as the new three-level ordinal outcome, there were 73 patients having a response as PD or NC, 55 patients having a response as MR or PR and 41 patients having a response as CR. By combining drug response in Terragna et al. [2] as the new three-level ordinal outcome, there were 49 patients having a response as SD or PR, 54 patients having a response as VGPR or nCR and 15 patients having a response as CR.

Predictive genomic classifiers analysis

For the original ordinal drug response outcome in Mulligan et al. [10], we first filtered probes based on all the options from top 30 and then 50 probes to top 500 probes by 50 probes. Based on the predictive performance of all the 11 models (Table 2) evaluated with 10-fold cross-validation with 10 repeats, it showed that the best predictive model would include all 450 top probes for smallest deviance.

Table 1 Summary of studies and frequency table for original ordinal outcome in both studies

Study	Mulligan et al. [10]	Terragna et al. [2]
Treatment	Bortezomib	VTD
Number of Samples	169	118
Number of Probes	22,283	54,677
Patients Population	Relapsed MM	New-Diagnosis
	Progressive Disease (PD)	Stable Disease (SD)
	13 (7.70%)	7 (5.93%)
	No Change (NC)	Partial Response (PR)
	60 (35.50%)	42 (35.59%)
	Minimal Response (MR)	Very good partial response (VGPR)
	12 (7.10%)	40 (33.90%)
	Partial Response (PR)	near Complete Response (nCR)
43 (25.44%)	14 (11.87%)	
Complete Response (CR)	Complete Response (CR)	
41 (24.26%)	15 (12.71%)	

However, for the final predictive model, we included top 50 probes as predictors with a balance between a reasonable decrease in deviance and simplicity in predictive model for clinical application. The prior scale in the final model was chosen at 0.14. Deviance of the final model was 441.919 and AUC was 0.632; while MSE was 0.136 and misclassification rate was 0.189. For the new combined ordinal drug response outcome in Mulligan et al. [10], we filtered probes based on all the options from top 30 and then 50 probes to top 500 probes by 50 probes. Based on the predictive performance of all the 11 models (Table S1 [See Additional file 1]) evaluated by 10-fold cross-validation with 10 repeats, it showed that the best predictive model included all 400 top probes for smallest deviance. However, for the final predictive model, we included top 50 probes as predictors with a balance between a reasonable decrease in deviance and simplicity in predictive model for clinical application. The prior scale in the final model was chosen at 0.16. For the final model, deviance was 316.118, AUC was 0.696, MSE was 0.190 and misclassification rate was 0.273. Comparing the predictive modeling performance between five-level ordinal drug response and the new combined ordinal drug response, AUC increased by 0.060 and deviance decreased by 125.801 for the combined ordinal outcome with a trade-off in MSE increasing by 0.054 and misclassification rate increasing by 0.084.

For the original ordinal drug response outcome in Terragna et al. [2], we also filtered probes based on all the options from top 30 and then 50 probes to top 500 probes by 50 probes. Based on the predictive

Table 2 Summary of predictive performance using different number of top probes for drug response prediction (five levels) in two studies

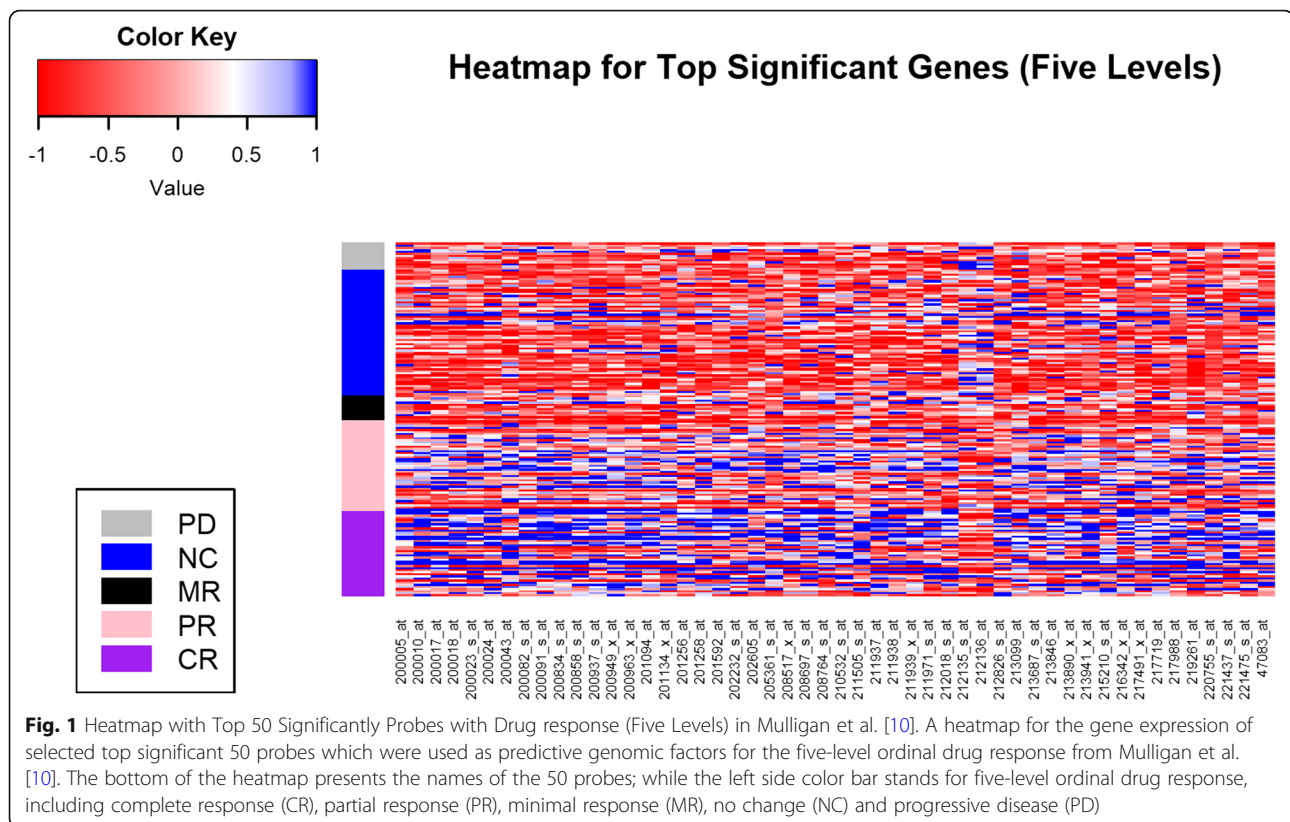
Number of top genes	Prior Scale	10 fold with 10 repeats cross-validation				Leave one out cross-validation			
		Deviance	AUC	MSE	misclassification	Deviance	AUC	MSE	misclassification
Mulligan et al. [10] (Five Level Drug response Outcome)									
30	0.4	452.868	0.616	0.140	0.204	452.805	0.605	0.140	0.201
50	0.14	441.919	0.632	0.136	0.189	442.235	0.618	0.137	0.194
100	0.14	439.464	0.634	0.136	0.190	438.839	0.625	0.136	0.188
150	0.15	431.669	0.657	0.133	0.185	434.309	0.645	0.134	0.186
200	0.15	425.471	0.673	0.132	0.186	427.950	0.662	0.133	0.189
250	0.15	428.840	0.669	0.133	0.187	431.478	0.661	0.133	0.189
300	0.15	437.913	0.660	0.135	0.189	443.893	0.642	0.137	0.194
350	0.15	425.215	0.690	0.133	0.185	430.448	0.679	0.135	0.186
400	0.09	411.556	0.692	0.130	0.181	407.755	0.690	0.129	0.178
450	0.09	410.313	0.700	0.129	0.179	402.852	0.704	0.127	0.179
500	0.14	426.196	0.705	0.134	0.194	434.344	0.691	0.138	0.209
Terragna et al. [2] (Five Level drug response Outcome)									
30	0.95	270.440	0.776	0.126	0.188	266.318	0.780	0.126	0.186
50	0.17	270.285	0.755	0.128	0.195	267.102	0.764	0.127	0.192
100	0.17	267.890	0.757	0.126	0.185	264.300	0.764	0.124	0.183
150	0.26	277.956	0.766	0.129	0.190	276.060	0.770	0.130	0.197
200	0.26	277.353	0.767	0.130	0.187	275.197	0.765	0.130	0.188
250	0.26	274.008	0.775	0.128	0.187	265.904	0.778	0.127	0.169
300	0.23	273.854	0.779	0.128	0.185	268.504	0.779	0.126	0.180
350	0.23	272.087	0.776	0.126	0.174	271.819	0.779	0.124	0.169
400	0.12	264.856	0.769	0.126	0.182	258.087	0.774	0.121	0.173
450	0.16	263.406	0.779	0.124	0.176	258.174	0.785	0.121	0.173
500	0.16	259.613	0.789	0.122	0.166	252.812	0.796	0.118	0.159

performance of all the 11 models (Table 2), it showed that the best predictive model included all 500 top probes for smallest deviance. However, for the final predictive model, we included top 30 probes as predictors with a balance between a comparable low deviance and simplicity in predictive model for clinical application. The prior scale in the final model was chosen at 0.95. For the final model, deviance was 270.440, AUC was 0.776, MSE was 0.126 and misclassification rate was 0.188. For the new combined ordinal drug response outcome in Terragna et al. [2], we filtered probes based on all the options from top 30 and then 50 probes to top 500 probes by 50 probes. Based on the predictive performance of all the 11 models (Additional file 1: Table S1), it showed that the best predictive model included all 500 top probes for smallest deviance. However, for the final predictive model, we included top 50 probes as predictors with a balance between a comparable low deviance and simplicity in predictive model for clinical application. The prior scale in the final model was chosen at 0.

26. Deviance of the final model was 167.130 and AUC was 0.800; while MSE was 0.152 and misclassification rate was 0.233. We compared the predictive performance of all models using 10-fold cross-validation with 10 repeats and leave one out cross-validation, which lead to similar results. The results were shown in Table 2 and Additional file 1: Table S1. Comparing the predictive modeling performance between five-level ordinal drug response and the new combined ordinal drug response, AUC increased by 0.024 and deviance decreased by 103.310 for the combined ordinal outcome with a trade-off in MSE increasing by 0.026 and misclassification rate increasing by 0.045.

Genes identification

To visualize the selected significant probes and its relationship with the clinical outcome in Mulligan et al. [10], a heatmap was presented in Fig. 1 with the top 50 significant probes which were used as predictive genomic factors for the five-level ordinal drug response.



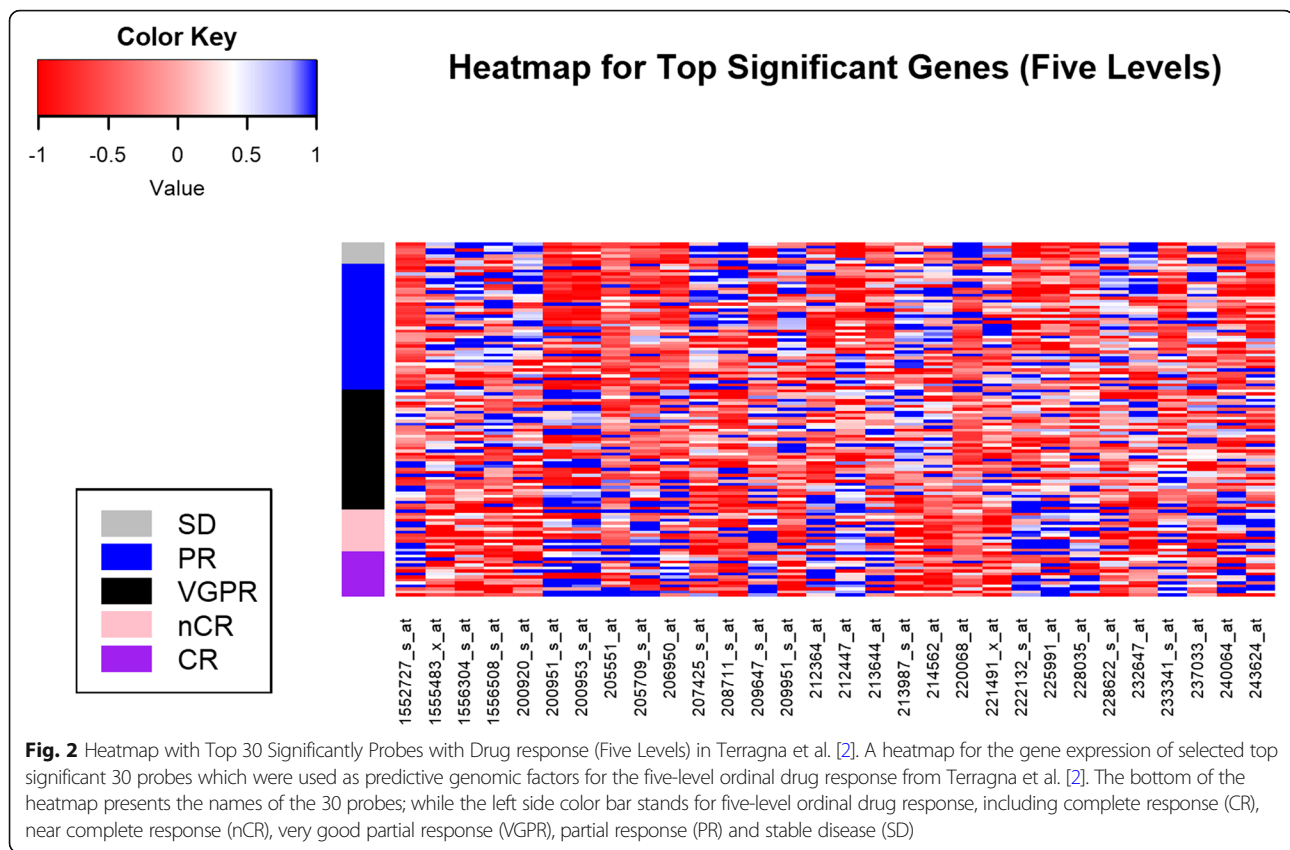
The top 50 significant probes represent genes of known function. Most of the probes are overexpressed in patients with PR or CR, which covers various functions including ribosomal protein (RPL11, RPL15, RPS7, RPS13), mitochondrial (COX7C), eukaryotic translation initiation factors (EIF3D, EIF3E, EIF3F, EIF3H) genes. Two of the probes are under-expressed in patients achieving PR or CR, which represent the gene function as ATPase plasma membrane Ca²⁺ transporting 4 (ATP2B4). For the three-level ordinal drug response, a heatmap was presented in Figure S1 [See Additional file 2] with the top 50 significant probes which were used as predictive genomic factors in our final predictive model. The top 50 probes for three-level drug response overlapped with most of the top 50 probes for five-level drug response. Only a few probes represent different genes of functions, including eukaryotic translation elongation factor (EEF2), chloride voltage-gated channel (CLCN3), abhydrolase domain containing (ABHD14A).

To visualize the selected significant probes and its relationship with the clinical outcome in Terragna et al. [2], a heatmap was presented in Fig. 2 with the top 30 significant probes which were used as predictive genomic factors for the five-level ordinal drug response. The top 30 probes in this dataset differentiated with similar chance to over-express or down-express, which also cover various functions including BTG anti-

proliferation factor 1 (BTG1), CDP-diacylglycerol synthase 1 (CDS1), RNA polymerase I subunit B (POLR1B), acylglycerol kinase (AGK), cyclin D1 (CCND1), cyclin D2 (CCND2), major histocompatibility complex, class II, DQ beta 1 (HLA-DQB1), mitogen-activated protein kinase 7 (MAP2K7) and suppressor of cytokine signaling 5 (SOCS5) genes. For the three-level ordinal drug response, a heatmap was presented in Figure S2 [See Additional file 3] with the top 50 significant probes which were used as predictive genomic factors in our final predictive model. The top 30 probes for three-level drug response overlapped with most of the top 50 probes for five-level drug response. Only a few probes represent different genes of functions, including SRC proto-oncogene, non-receptor tyrosine kinase (SRC), TNF receptor superfamily member 13C (TNFRSF13C) and checkpoint kinase 1 (CHEK1).

Discussion

It is highly important to identify genetic biomarkers to predict drug response with a narrow therapeutic index [19, 26]. Chemotherapeutic agents are medications in that category, since the response is variable with potentially lethal side effects [19, 26]. Many studies have been conducted and a large number of biomarkers have been reported [19]. However, a complex outcome, such as drug response, is generally affected by many genomic



and environmental factors [19]. Thus, it is desirable that a predictive procedure will possess the capability to consider mutual effects of various biomarkers for drug response [19].

Another potential issue is that the multi-level ordinal drug response is usually recoded in clinical practice. For analytic simplicity, however, such multi-level ordinal outcome is usually combined as just two levels, as in the original papers [2, 10] that we reanalyzed in this study. However, this strategy not only risks both the loss of information in the data and arbitrary to select the recode strategy, but also cannot provide informative prediction [27].

We here utilized a more efficient approach by combining the standard ordinal logistic regression and the hierarchical modeling. Our method can jointly analyze numerous variables for detecting important predictors and for predicting multi-level drug response. We applied our method to reanalyze two publicly available clinical trial datasets, which assessed response to bortezomib in relapsed MM patients [10] and VTD in newly diagnosed MM patients [2]. The original studies both treated the five-level ordinal drug responses as binary responses. To address the drawback of the potential loss of information from recoding, we reanalyzed the datasets by using the original ordinal drug responses. To avoid low frequencies in several levels of the five-level drug responses, we

redefined the five-level drug response as a three-level ordinal drug response in both datasets. The results reveal that the predictive performance from VTD in new MM patients is more powerful than treating relapsed MM patients with bortezomib alone. Meanwhile, comparing the analysis results between five-level ordinal drug response and reduced three-level ordinal drug response, AUC increased and deviance decreased for the combined ordinal outcome with a trade-off in MSE and misclassification rate. Our analyses show that the combining ordinal outcome could result in higher MSE and misclassification rate, thus, potential loss of information and misleading interpretation. Although we only compared the five-level ordinal outcome with three-level ordinal outcome, it is anticipated that similar differences will exist if compared with binary outcome. It also implies that the original approach to analyze ordinal outcome as binary outcome will possibly lead to information loss.

Furthermore, we identified probes that represent genes of known function. In Mulligan et al. [10], for both five level and three level ordinal drug response, most of the top significant probes are overexpressed in patients with PR or CR, including ribosomal protein (RPL11, RPL15, RPS7, RPS13), mitochondrial (COX7C), eukaryotic translation initiation factors (EIF3D, EIF3E, EIF3F, EIF3H) genes. Among them, ribosomal protein has been

investigated by multiple studies to show that mutations in ribosomal protein genes have been found in endometrial cancer (RPL22), T-cell acute lymphoblastic leukemia (RPL10, RPL5 and RPL11), chronic lymphocytic leukemia (RPS15), colorectal cancer (RPS20), and glioma (RPL5) [28]. Moreover, it has also been discussed that eukaryotic initiation factors (EIFs) play an important role in translation initiation and protein synthesis which could alter angiogenesis, tumor development, and apoptosis in cancer progression [29]. Two of the probes are under-expressed in patients achieving PR or CR, which represent the gene ATP2B4. ATP2B4 plays a critical role in intracellular calcium homeostasis by regulating the enzymes to remove bivalent calcium ions from eukaryotic cells against very large concentration gradients [30]. In Terragna et al. [2], we carried a function enrichment analysis to identify the functional annotation of the top probes with KEGG [31] using the Bioinformatics tool DAVID [32, 33]. The top 30 probes for the five-level ordinal drug response also cover various gene functions which belong to multiple important pathways, e.g., Metabolic pathways, p53 signaling pathway, PI3K-Akt signaling pathway, AMPK signaling pathway, Wnt signaling pathway, Jak-STAT signaling pathway, Viral carcinogenesis and MAPK signaling pathway. For the three-level ordinal drug response, the top 50 significant probes cover similar functions as the top 30 probes for the five-level ordinal drug response, with several additional functions such as Cytokine-cytokine receptor interaction, NF-kappa B signaling pathway, Intestinal immune network for IgA production, HTLV-I infection and Primary immunodeficiency. This suggests that the probes we identified are correlated biologically. Based on the functional enrichment analysis results, the probes could be grouped to multiple pathways. One plausible solution is to utilize a pathway-structured model to incorporate that biological information into the predictive model to include more probe information into the prediction, which will be considered in our further work.

Although the predictive classifier and genetic biomarkers described here are promising, further research is necessary to assess the relevance of these genomic predictors with more data from other trials or other trials with novel or multi-agent therapy. Our analysis strategy is directly applicable to new data with bortezomib or other therapies in multiple myeloma for patients with newly diagnosed or relapsed cancer. This analysis will help to quickly identify the patient groups that could benefit from the proposed drug therapy or in need of other novel therapies.

Conclusions

We propose a novel method to directly analyze the multi-level drug response, rather than combining the

response to two groups. Our method employs a hierarchical ordinal logistic model with the heavy-tailed Cauchy prior on the coefficients. The proposed method allows us to jointly fit numerous correlated predictors and thus build efficient models for predicting the multi-level drug response. The predictive model for the multi-level drug response can be more informative than the previous approaches. Thus, the proposed approach provides a powerful tool for predicting multi-level drug response and has important impact on cancer studies.

Additional files

Additional file 1: Table S1. Summary of predictive performance using different number of top probes for drug response prediction (three levels) in two studies. (DOCX 21 kb)

Additional file 2: Figure S1. Heatmap with Top 50 Significantly Probes with Drug response (Three Levels) in Mulligan et al. [10]. (DOCX 77 kb)

Additional file 3: Figure S2. Heatmap with Top 50 Significantly Probes with Drug response (Three Levels) in Terragna et al. [2]. (DOCX 72 kb)

Abbreviations

AUC: Area under the ROC curve; BFGS: Quasi-Newton algorithm; CR: Complete response; EIFs: Eukaryotic initiation factors; IgH: Immunoglobulin heavy; MM: Multiple Myeloma; MR: Minimal response; MSE: Mean squared error; NC: No change; nCR: Near complete response; PD: Progressive disease; PR: Partial response; SD: Stable disease; VGPR: Very good partial response; VTD: Bortezomib-thalidomide-dexamethasone (VTD)

Acknowledgments

This work was supported in part by research grants from USA National Institutes of Health (R03-DE025646), National Natural Science Foundation of China (81673448), Natural Science Foundation of Jiangsu Province China (BK 20161218), and The Applied Basic Research Programs of Suzhou City (SYS201546).

Funding

This work was supported in part by research grants from USA National Institutes of Health (R03-DE025646), National Natural Science Foundation of China (81673448), Natural Science Foundation of Jiangsu Province China (BK 20161218), and The Applied Basic Research Programs of Suzhou City (SYS201546). The first two grants mainly support the development of statistical methods and software, and the others mainly support the design of the study and analysis, and interpretation of data and writing the manuscript.

Availability of data and materials

The datasets used and analyzed in the current study are publically available from GEO under accession number [GEO: GSE9782] and [GEO: GSE68871], (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE9782> and <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE68871>). The method has been incorporated into the freely available R package BhGLM (<http://www.ssg.uab.edu/bhglm/> and <https://github.com/abbyyan3/BhGLM>).

Authors' contributions

WZ and BL proposed the idea and identified the real data sets. NY developed the statistical method and the software. XZ performed simulation studies and real data analysis. XZ, WZ, and NY drafted the manuscript. BL, HH, SS, HX, and YH organized the real data sets, revised the manuscript, and discussed the project with WZ and NY as it progressed and commented on the drafts of the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Biostatistics, Jiann-Ping Hsu College of Public Health, Georgia Southern University, Statesboro, GA, USA. ²Department of Hematology, The Second Affiliated Hospital of Soochow University, Suzhou, China. ³Department of Cell Biology, School of Biology & Basic Medical Sciences, Soochow University, Suzhou, China. ⁴Department of Biostatistics, University of Alabama at Birmingham, Birmingham, AL 35294, USA.

Received: 29 December 2017 Accepted: 7 May 2018

Published online: 10 May 2018

References

- Kyle RA, Rajkumar SV. Multiple myeloma. *Blood*. 2008;111(6):2962–72.
- Terragna C, Remondini D, Martello M, Zamagni E, Pantani L, Patriarca F, Pezzi A, Levi G, Offidani M, Proserpio I, et al. The genetic and genomic background of multiple myeloma patients achieving complete response after induction therapy with bortezomib, thalidomide and dexamethasone (VTD). *Oncotarget*. 2016;7(9):9666–79.
- American Cancer S. *Cancer Facts & Figures 2017*. Atlanta: American Cancer Society; 2017.
- Decaux O, Lode L, Magrangeas F, Charbonnel C, Gouraud W, Jezequel P, Attal M, Harousseau JL, Moreau P, Bataille R, et al. 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 Intergrroupe francophone du Myelome. *J Clin Oncol*. 2008;26(29):4798–805.
- Broyl A, Hose D, Lokhorst H, de Knecht Y, Peeters J, Jauch A, Bertsch U, Buijs A, Stevens-Kroef M, Beverloo HB, et al. Gene expression profiling for molecular classification of multiple myeloma in newly diagnosed patients. *Blood*. 2010;116(14):2543–53.
- Hofman IJF, van Duijn M, De Bruyne E, Fancello L, Mulligan G, Geerdens E, Garelli E, Mancini C, Lemmens H, Delforge M, et al. RPL5 on 1p22.1 is recurrently deleted in multiple myeloma and its expression is linked to bortezomib response. *Leukemia*. 2017;31(8):1706–14.
- Kumar SK, Rajkumar SV, Dispenzieri A, Lacy MQ, Hayman SR, Buadi FK, Zeldenrust SR, Dingli D, Russell SJ, Lust JA, et al. Improved survival in multiple myeloma and the impact of novel therapies. *Blood*. 2008;111(5):2516–20.
- Rajkumar SV. Treatment of multiple myeloma. *Nat Rev Clin Oncol*. 2011;8(8):479–91.
- Malek E, Abdel-Malek MA, Jagannathan S, Vad N, Karns R, Jegga AG, Broyl A, van Duijn M, Sonneveld P, Cottini F, et al. Pharmacogenomics and chemical library screens reveal a novel SCFSKP2 inhibitor that overcomes bortezomib resistance in multiple myeloma. *Leukemia*. 2017;31(3):645–53.
- Mulligan G, Mitsiades C, Bryant B, Zhan F, Chng WJ, Roels S, Koenig E, Fergus A, Huang Y, Richardson P, et al. Gene expression profiling and correlation with outcome in clinical trials of the proteasome inhibitor bortezomib. *Blood*. 2007;109(8):3177–88.
- Bolli N, Avet-Loiseau H, Wedge DC, Van Loo P, Alexandrov LB, Martincorena I, Dawson KJ, Iorio F, Nik-Zainal S, Bignell GR, et al. Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nat Commun*. 2014;5:2997.
- Walker BA, Boyle EM, Wardell CP, Murison A, Begum DB, Dahir NM, Proszek PZ, Johnson DC, Kaiser MF, Melchor L, 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):3911–20.
- Lohr JG, Stojanov P, Carter SL, Cruz-Gordillo P, Lawrence MS, Auclair D, Sougnez C, Knoechel B, Gould J, Saksena G, et al. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. *Cancer Cell*. 2014;25(1):91–101.
- Walker BA, Wardell CP, Murison A, Boyle EM, Begum DB, Dahir NM, Proszek PZ, Melchor L, Pawlyn C, Kaiser MF, et al. APOBEC family mutational signatures are associated with poor prognosis translocations in multiple myeloma. *Nat Commun*. 2015;6:6997.
- Kuehl WM, Bergsagel PL. Molecular pathogenesis of multiple myeloma and its premalignant precursor. *J Clin Invest*. 2012;122(10):3456–63.
- Morgan GJ, Walker BA, Davies FE. The genetic architecture of multiple myeloma. *Nat Rev Cancer*. 2012;12(5):335–48.
- Corre J, Munshi N, Avet-Loiseau H. Genetics of multiple myeloma: another heterogeneity level? *Blood*. 2015;125(12):1870–6.
- Lopez-Corral L, Sarasquete ME, Bea S, Garcia-Sanz R, Mateos MV, Corchete LA, Sayagues JM, Garcia EM, Blade J, Oriol A, et al. SNP-based mapping arrays reveal high genomic complexity in monoclonal gammopathies, from MGUS to myeloma status. *Leukemia*. 2012;26(12):2521–9.
- Geeleher P, Cox NJ, Huang RS. Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines. *Genome Biol*. 2014;15(3):R47.
- Simon R, Roychowdhury S. Implementing personalized cancer genomics in clinical trials. *Nat Rev Drug Discov*. 2013;12(5):358–69.
- Blade J, Samson D, Reece D, Apperley J, Bjorkstrand B, Gahrton G, Gertz M, Giralt S, Jagannath S, Vesole D. Criteria for evaluating disease response and progression in patients with multiple myeloma treated by high-dose therapy and haemopoietic stem cell transplantation. Myeloma subcommittee of the EBMT. European Group for Blood and Marrow Transplant. *Br J Haematol*. 1998;102(5):1115–23.
- Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian data analysis, third edition* edn. New York: Chapman & Hall/CRC Press; 2014.
- Gelman A, Hill J. *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press; 2007.
- Tibshirani RJ, Efron B. Pre-validation and inference in microarrays. *Stat Appl Genet Mol Biol*. 2002;1:Article1.
- Hastie T, Tibshirani R, Wainwright M. *Statistical learning with sparsity - the lasso and generalization*. New York: CRC Press; 2015.
- Jiang Y, Wang M. Personalized medicine in oncology: tailoring the right drug to the right patient. *Biomark Med*. 2010;4(4):523–33.
- Warner P. Ordinal logistic regression. *J Fam Plann Reprod Health Care*. 2008;34(3):169–70.
- Goudarzi KM, Lindstrom MS. Role of ribosomal protein mutations in tumor development (review). *Int J Oncol*. 2016;48(4):1313–24.
- Sharma DK, Bressler K, Patel H, Balasingam N, Thakor N. Role of eukaryotic initiation factors during cellular stress and Cancer progression. *J Nucleic Acids*. 2016;2016:8235121.
- NCBI: ATPase plasma membrane Ca²⁺ transporting 4. 2017.
- Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30.
- Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44–57.
- Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009;37(1):1–13.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://www.biomedcentral.com/submissions)

