

## Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

---

Dulin Wang

---

Date

Gain of Chromosome 1q is Associated with Early Progression in Multiple Myeloma Patients  
Treated with Lenalidomide, Bortezomib, and Dexamethasone

By

Dulin Wang  
Master of Science in Public Health

Biostatistics and Bioinformatics

---

Zhengjia (Nelson) Chen, PhD  
Thesis Advisor

---

Yuan Liu, PhD  
Reader

Gain of Chromosome 1q is Associated with Early Progression in Multiple Myeloma Patients  
Treated with Lenalidomide, Bortezomib, and Dexamethasone

By

Dulin Wang

B.Sc.  
Zhejiang University  
2018

Thesis Advisor: Zhengjia (Nelson) Chen, PhD

An abstract of  
A thesis submitted to the Faculty of the  
Rollins School of Public Health of Emory University  
in partial fulfillment of the requirements for the degree of  
Master of Science in Public Health  
in Biostatistics and Bioinformatics  
2020

## Abstract

### Gain of Chromosome 1q is Associated with Early Progression in Multiple Myeloma Patients Treated with Lenalidomide, Bortezomib, and Dexamethasone

By Dulin Wang

**Background:** Multiple myeloma patients identified with gain of chromosome 1q usually have inferior progression-free survival (PFS) and overall survival (OS) outcomes. This study examined and determined the prognostic effects of patients with +1q compared to those without +1q. 201 patients treated with lenalidomide, bortezomib, and dexamethasone (RVD) were included in the study.

**Methods:** Descriptive analysis was used to describe the demographic and clinical variables. Univariate analysis was conducted to determine the risk factors with gain of chromosome 1q for patients. Response rate to treatment associated with +1q was plotted. Survival analysis was performed to identify the risk factors for patients' PFS and OS. After univariate analysis, hazard ratio and p-value for each potential risk factor was calculated. Backward model selection was applied to determine the multivariate Cox proportional hazard model. Proportional Hazard (PH) assumption was checked. A final extended Cox model with time-dependent variables was determined. Kaplan Meier Curves stratified by interested sub-groups were applied.

**Results:** Patients with +1q (n=94), compared to those without +1q (n=107), had shorter median progression-free survival (53.2 months vs 70.5 months,  $p=0.010$ ,  $HR=1.76$ ) and overall survival ( $p=0.003$ ,  $HR=2.68$ ). In subgroup analyses, patients with co-occurring +1q and  $t(14;16)$  or  $del(17p)$ , or patients with 4 or more copies of 1q had significantly worse PFS (22.9 months and 34.8 months,  $p=0.001$  and  $p=0.015$ , respectively), whereas patients with 3 copies and standard risk cytogenetic abnormalities had no significant difference in PFS.

**Conclusion:** Patients with +1q and treated with RVD induction be considered at high risk for early progression and death in multiple myeloma when 4 or more 1q copies are detected or in the context of  $t(14;16)/del(17p)$  cytogenetic abnormalities.

**Key Word:** Survival analysis; extended Cox model; PH assumption; myeloma; cytogenetics abnormalities

Gain of Chromosome 1q is Associated with Early Progression in Multiple Myeloma Patients  
Treated with Lenalidomide, Bortezomib, and Dexamethasone

By

Dulin Wang

B.Sc.  
Zhejiang University  
2018

Thesis Advisor: Zhengjia (Nelson) Chen, PhD

A thesis submitted to the Faculty of the  
Rollins School of Public Health of Emory University  
in partial fulfillment of the requirements for the degree of  
Master of Science in Public Health  
in Biostatistics and Bioinformatics  
2020

## Table of Contents

1.	Introduction.....	1
2.	Method .....	3
	2.1 Patients and Demographics .....	3
	2.2 Outcomes .....	4
	2.3 Statistical Analysis Method .....	5
	2.3.1 <i>Descriptive analysis</i> .....	5
	2.3.2 <i>Cox Proportional Hazard model formulation</i> .....	5
	2.3.3 <i>Extended Cox Proportional Hazard model for time-dependent variables</i> .....	6
	2.3.4 <i>Univariate and multivariate survival analysis</i> .....	7
	2.4 Evaluate assumption .....	8
3.	Result.....	9
	3.1 Descriptive analysis .....	10
	3.2 Survival analysis.....	12
	3.2.1 <i>Response rates</i> .....	12
	3.2.2 <i>Survival outcomes</i> .....	13
	3.2.3 <i>Impact of co-occurring cytogenetic abnormalities</i> .....	17
	3.2.4 <i>Impact of copy number and detection threshold</i> .....	19
4.	Discussion .....	22
5.	Reference .....	24
6.	Appendix.....	27
	6.1 Supplementary Table 1 Patient characteristics .....	27
	6.2 Supplementary Table 2 Univariate survival analysis.....	29



## 1. Introduction

Multiple myeloma (MM) is a blood-related cancer and is characterized by uncontrolled proliferation of malignant plasma cells in the bone marrow, leading to the accumulation of monoclonal protein in the blood or urine, and associated organ dysfunction (Palumbo & Anderson, 2011). The prognosis of MM is based on the presence or absence of specific cellular genetic abnormalities detected by conventional karyotype analysis and/or fluorescence in situ hybridization (FISH) (Sawyer, 2011). Through this FISH analysis, it has been determined that the chromosomal abnormalities of t(4;14), t(14;16), t(14;20), gain(1q), del(1p), and amp(1p) are associated with disease progression and shorter overall survival. These chromosomal abnormalities have been combined into The Revised International Staging System (RISS) that creates a unified prognosis index and helps predict survival time for newly diagnosed patients (Palumbo et al., 2015; Rajkumar, 2018).

The gain in Chromosome 1q (+1q) is the most crucial cytogenetic abnormalities in MM that portends a worse diagnosis and up to 40% are recognized in patients (Sonneveld et al., 2016). +1q means amplifying the bad genes in myeloma. This abnormality generally occurs in the form of balanced translocation, amplification, and jumping translocation that leads to the increase of the copy number (Marzin). MM has two basic precursor conditions, Monoclonal gammopathy of undetermined significance (MGUS) and smoldering MM (SMM). The +1q detection rates in SMM patients are higher than in MGUS patients (Hanamura et al., 2006). Furthermore, 4 or more increases in the copy number of chromosomes 1q, which is defined as amplification (amp(1q)), may lead to

worse negative effects on survival. The deletions in Chromosome are also strongly associated with 1q gains (Chang et al., 2010). Shaughnessy's group identified the overexpression of CKS1B is the potential driver gene of 1q gain (Shaughnessy, 2005). Recent studies suggested that the gene of ADAR1 and MCL1 also impact the gain in 1q (Samo et al., 2018; Teoh et al., 2018). Previous studies in MM considered +1q an independent negative prognostic parameter (Avet-Loiseau et al., 2009; Hanamura et al., 2006; Nemeč et al., 2010), but its significance remains disputed. Other studies indicated +1q is a non-independent predictor on disease progression when incorporated with other high-risk factors such as proliferation marker and t(4:14) in multivariate analysis (Fonseca et al., 2006).

The induction therapy with lenalidomide, bortezomib, and dexamethasone (RVD) has significant progression-free and overall survival benefits (Smetana et al., 2013). Although the bortezomib-based regimens were proved to overcome the negative prognostic effects of +1q (Smetana et al., 2013), studies on patients treated with RVD indicated +1q still a prognostic parameter (Dimopoulos et al., 2010; Shah et al., 2017). Previous studies were rarely conducted with this novel induction therapy; thus, the conclusion was limited by the sample size and patient's selection bias.

In this paper, we performed a survival analysis based on a large cohort of Winship Cancer Institute of Emory University patients with Multiple Myeloma who were treated with RVD induction therapy between February 2010 and April 2015. Our analysis compared overall survival (OS) and progression free survival (PFS) between patients with +1q abnormalities and those without +1q, all of whom received RVD induction therapy. In addition, we analyzed distinct clinical features of co-occurring cytogenetic

abnormalities and high-risk cytogenetic abnormalities among the +1q patients compared to those patients without +1q abnormalities across the copy numbers. Findings from this study aim to provide evidence for clinicians and patients to predict early progression of MM and better inform their treatment options and decision-making process.

## **2. Method**

### **2.1 Patients and Demographics**

The data for analysis came from a retrospective study on MM patients conducted at Winship Cancer Institute of Emory University, and the use of data was approved by Emory Institutional Review Board. There were some criterions to select patients from Emory University database for evaluation in the final analysis: 1) Patients not tested for +1q or had insufficient diagnostic material from bone marrow biopsy were excluded ; 2) Patients tested for +1q only at relapse were excluded; 3) Among patients were not enriched for CD138 cells in diagnostic of bone marrow, we excluded patients if +1q was not detected by FISH; 4) Patients were excluded if +1q was detected only by karyotype and not by FISH.

Baseline demographic (age, sex, race, ethnicity), clinical features, and laboratory characteristics (isotype, hemoglobin, platelets, creatinine, calcium, albumin, lactate dehydrogenase (LDH), beta-2-microglobulin (B2M), M-spike, and serum free light chains) were obtained from the patients' electronic medical records. Patients were classified by International Staging System (ISS) stage and Revised International Staging System (R-ISS) stage at diagnosis. Fish were performed to determine the presence of

t(11;14), t(4;14), t(14;16), del(17p), del(13q), hyperdiploidy, +1q, and del(1p) at diagnosis. Patients were also categorized as whether having a complex karyotype, which is defined by the presence of more than or equal to 3 chromosomal abnormalities. All these features will go through descriptive analysis and be incorporated into patient characteristic table (Table 1).

The following data were also collected for each patient: date of diagnosis, treatment initiation, ASCT, maintenance therapy, best response to induction therapy and transplantation, and dates of first progression and death.

## **2.2 Outcomes**

Primary outcomes of interest included best response to RVD induction therapy, median PFS, and OS for patients with +1q and those without +1q. In this paper, treatment response was assessed according to the criteria established by International Myeloma Working Group (IMWG). PFS was defined as the time from treatment to relapse or death by any cause or censored at last follow-up that half of the patients were still alive. OS was defined as the time from treatment to death by any cause or censored at last follow-up that half of the patients were still alive.

## **2.3 Statistical Analysis Method**

### *2.3.1 Descriptive analysis*

The descriptive table for patients' characteristics were conducted. For continuous variables, the median was summarized and two-category with a specific cut-off for each continuous variable were shown on descriptive table in the result. For binary or categorical variables, the frequencies and percentage were presented.

The univariate analysis was performed, the crude odds ratio and 95% confidence interval were calculated for each risk factor to present a general idea of the association between outcome and a single independent variable.

The univariate analysis for the patients with 1q compared to the patients without 1q were also summarized. For categorical covariates, a contingency table along with the Chi-square test (parametric p-value) or Fisher's exact test (non-parametric p-value) were produced. For numerical covariates, the sample size, mean and median along with ANOVA test (parametric p-value) or Kruskal-Wallis test (non-parametric p-value) were produced.

### *2.3.2 Cox Proportional Hazard model formulation*

For survival analysis, estimated Kaplan Meier survival curves were plotted for each level of risk factors to have a general perspective about overall survival of the patients with different characteristics and to serve as a check of the proportional hazard assumption.

Cox proportional hazard regression model was constructed for estimating survival curves when assessed several explanatory variables simultaneously. The Cox proportional hazard model can be written as:

$$h(t|\mathbf{X}_i) = h_0(t) \exp(\beta_1 X_{i1} + \dots + \beta_p X_{ip}) = h_0(t) \exp(\boldsymbol{\beta}_i \mathbf{X}_i')$$

$h_0(t)$  is an arbitrary and unspecified baseline hazard function.  $\mathbf{X}_i$  is the vector of the explanatory variables for the individual  $i$ .  $\boldsymbol{\beta}$  is the vector of unknown regression parameters associated with the explanatory variables. The hazard ratio for a specific variable is:

$$\widehat{HR} = \frac{\widehat{h}(t, \mathbf{X}^*)}{\widehat{h}(t, \mathbf{X})} = \exp \left[ \sum_{i=1}^p \beta_i (\mathbf{X}_i^* - \mathbf{X}_i) \right]$$

### 2.3.3 Extended Cox Proportional Hazard model for time-dependent variables

An important feature of Cox model, which concerns the proportional hazards (PH) assumption, is that the baseline hazard is a function of  $t$  but does not involve the  $\mathbf{X}_i$ , whereas the exponential expression involves the  $\mathbf{X}_i$  but does not involve  $t$ . The  $\mathbf{X}_i$  here are time-independent.

However, the  $\mathbf{X}_i$  can involve the  $t$  and be time-dependent. If time-dependent variables are considered, the Cox model form may still be used, but such a model no longer satisfies the PH assumption and is called the extended Cox model. We resort to formulating the extended Cox model to incorporate time-dependent variables:

$$\begin{aligned}
h(t, \mathbf{X}(t)) &= h_0(t) \exp(\beta_1 X_{i1} + \dots + \beta_{p1} X_{ip1} + \delta_1 X_{j1}(t) + \dots + \delta_{p2} X_{jp2}(t)) \\
&= h_0(t) \exp(\boldsymbol{\beta}_i \mathbf{X}'_i + \boldsymbol{\delta}_j \mathbf{X}'_j(t))
\end{aligned}$$

The extended model contains a baseline hazard function  $h_0(t)$  which is multiplied by an exponential function. The exponential part contains both time-independent predictors, as denoted by the  $\mathbf{X}_i$ , and time-dependent predictors, as denoted by the  $\mathbf{X}_j(t)$ . The entire collection of predictors at time  $t$  is denoted by  $\mathbf{X}(t)$ . The hazard ratio for a specific variable is:

$$\begin{aligned}
\widehat{HR} &= \frac{\widehat{h}(t, \mathbf{X}^*(t))}{\widehat{h}(t, \mathbf{X}(t))} = \frac{h_0(t) \exp(\sum_{i=1}^{p1} \widehat{\boldsymbol{\beta}}_i \mathbf{X}_i^* + \sum_{j=1}^{p2} \widehat{\boldsymbol{\delta}}_j \mathbf{X}_j^*(t))}{h_0(t) \exp(\sum_{i=1}^{p1} \widehat{\boldsymbol{\beta}}_i \mathbf{X}_i + \sum_{j=1}^{p2} \widehat{\boldsymbol{\delta}}_j \mathbf{X}_j(t))} \\
&= \exp \left[ \sum_{i=1}^{p1} \widehat{\boldsymbol{\beta}}_i (\mathbf{X}_i^* - \mathbf{X}_i) + \sum_{j=1}^{p2} \widehat{\boldsymbol{\delta}}_j (\mathbf{X}_j^*(t) - \mathbf{X}_j(t)) \right]
\end{aligned}$$

#### 2.3.4 Univariate and multivariate survival analysis

For univariate survival analysis, the hazard ratio with 95% confidence interval for each level of risk factors were summarized. All categorical variables were reference cell coded. The hazard ratio is a measure of the magnitude of the difference between the two curves in the Kaplan–Meier plot. Local Wald test (generated HR p-values) was performed to see if there any significant difference between difference levels of covariates in KM curves. The log rank test (generated log-rank p-value) was used to determine whether KM curves for two or more groups are statistically equivalent.

The multivariate analysis was also performed. Backward selection with an alpha level of 0.1 was conducted to select the best Cox regression model to predict the outcome. The general backward selection procedure is as following:

Step1: There are N significant variables in univariate survival analysis are added in the model and will be included in the model for the following selection process.

Step2: Remove the variable with the largest p-value. That is the variable with the least statistically significant. Refit the new model with (N-1) variables.

Step3: Repeat step 2 until all remaining variables have a significant p-value under the significance threshold. The model with all the significant variables is the final model selected by backward selection.

Adjusted hazard ratio was calculated for each risk factor in multivariate model. Local Wald tests were conducted for each variable in the final model and p-value were output.

## **2.4 Evaluate assumption**

In fitting the Cox PH models, we assumed independence of survival times between the distinct individual in the sample and the ratio of the hazards for any two individuals is constant over time. We also assumed censoring was non-informative, that is we assumed each subject's censoring time was independent of their survival time and the censoring was not related to the patients' physical condition. Harrel & Lee test based on Schoenfeld residuals for proportional hazard assumption and p-values were reported. if the PH

assumption holds for a covariate, then the Schoenfeld residuals for that covariate will not be related to survival time. The test procedure is as following:

Step1: Run a Cox PH model and obtain Schoenfeld residuals for each predictor;

Step2: Create a variable that ranks the order of failures. The subject who has the earliest event gets a value of 1, the next gets a value of 2, and so on;

Step3: Test the correlation between the variables created in the first and second steps.

The null hypothesis is that the correlation between the Schoenfeld residuals and ranked failure time is 0. Rejection of the null hypothesis leads to a conclusion that the PH assumption is violated.

All the analyses were performed using SAS 9.4. The significance level was set to 0.05 if not mentioned.

### **3. Result**

We identified 553 MM patients who were treated with RVD induction at Winship Cancer Institute of Emory University. A total number of 201 patients were identified in the final analysis after excluding patients who were not tested for +1q at diagnosis or conditions could not confirmed. There were 94 (46.7%) patients having at least one extra copy of chromosome 1q by FISH. Median follow up was 48 months among all patients.

### 3.1 Descriptive analysis

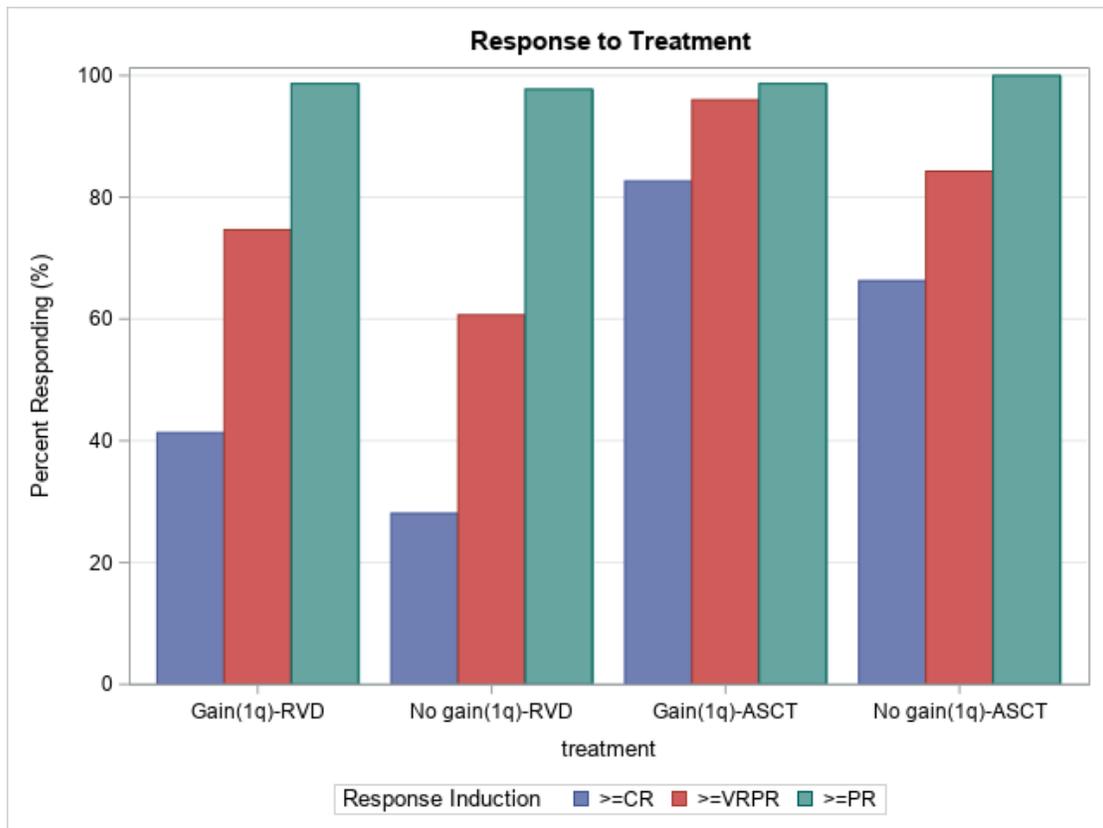
**Table 1 Patient characteristics**

Covariate	Level	No +1q (n=107)	+1q (n=94)	Total	P-value
Age	Median	63	64	64	
	<65	61 (57.01)	52 (55.32)		0.810
	>=65	46 (42.99)	42 (44.68)		
Hemoglobin (g/dL)	Median	11.07	10.09	10.50	
	<10	26 (26.8)	39 (45.88)	65 (35.7)	<b>0.007</b>
	>=10	71 (73.2)	46 (54.12)	117 (64.3)	
Platelets (x10 <sup>3</sup> /μL)	Median	226.63	300.66	207.00	
	<150	11 (12.64)	23 (30.26)	34 (20.9)	<b>0.006</b>
	>=150	76 (87.36)	53 (69.74)	129 (79.1)	
Calcium (mg/dL)	Median	9.37	9.67	9.20	
	>10.5	3 (3.57)	13 (16.25)	16 (9.8)	<b>0.006</b>
	<=10.5	81 (96.43)	67 (83.75)	148 (90.2)	
β2M (mg/L)	Median	3.31	4.39	2.95	
	>5.5	9 (10.23)	17 (23.61)	26 (16.3)	<b>0.022</b>
	<=5.5	79 (89.77)	55 (76.39)	134 (83.8)	
ISS stage	1	39 (45.88)	20 (28.57)	59 (38.1)	<b>0.024</b>
	2	37 (43.53)	33 (47.14)	70 (45.2)	
	3	9 (10.59)	17 (24.29)	26 (16.8)	
R-ISS Stage	1	34 (38.64)	14 (19.44)	48 (30.0)	<b>0.003</b>
	2	53 (60.23)	51 (70.83)	104 (65.0)	
	3	1 (1.14)	7 (9.72)	8 (5.0)	
T(4,14)	No	103 (99.04)	81 (91.01)	184 (95.3)	<b>0.008</b>
	Yes	1 (0.96)	8 (8.99)	9 (4.7)	
T(14,16)	No	103 (99.04)	79 (90.8)	182 (95.3)	<b>0.007</b>
	Yes	1 (0.96)	8 (9.2)	9 (4.7)	
Del(13q)	No	79 (74.53)	33 (36.67)	112 (57.1)	<b>&lt;.001</b>
	Yes	27 (25.47)	57 (63.33)	84 (42.9)	
del(1p)	No	99 (94.29)	73 (82.95)	172 (89.1)	<b>0.012</b>
	Yes	6 (5.71)	15 (17.05)	21 (10.9)	
Complex Karyotype	Yes	15 (14.02)	29 (31.52)	44 (22.1)	<b>0.003</b>
	No	92 (85.98)	63 (68.48)	155 (77.9)	

The univariate analysis of patient characteristics is summarized in **Table 1**. From the table, the median age of the patients was 64 and there is no significant difference between patients who gained +1q abnormalities (n=94) at diagnosis and who did not (n=107, p=0.810). Patients with +1q were significantly associated with lower hemoglobin and platelets (10.1 vs. 11.1, 226.6 vs. 300.7 in median) compared to no +1q patients, which indicated anemia and/or thrombocytopenia. Patients with +1q were significantly associated with higher Calcium, beta-2-microglobulin and a higher ISS and R-ISS stage. Additionally, patients with +1q were significantly associated with presence of t(4;14), t(14;16), del(13q), del(1p), and complex karyotype at diagnosis (1 vs. 8, 1 vs. 8, 27 vs. 57, 6 vs. 15) compared to no +1q patients. There was no significant association between patients with +1q and patients without +1q in the frequency of whether received upfront transplant or whether maintenance therapy was prescribed. The complete univariate analysis is in **Supplementary Table 1**.

## 3.2 Survival analysis

### 3.2.1 Response rates

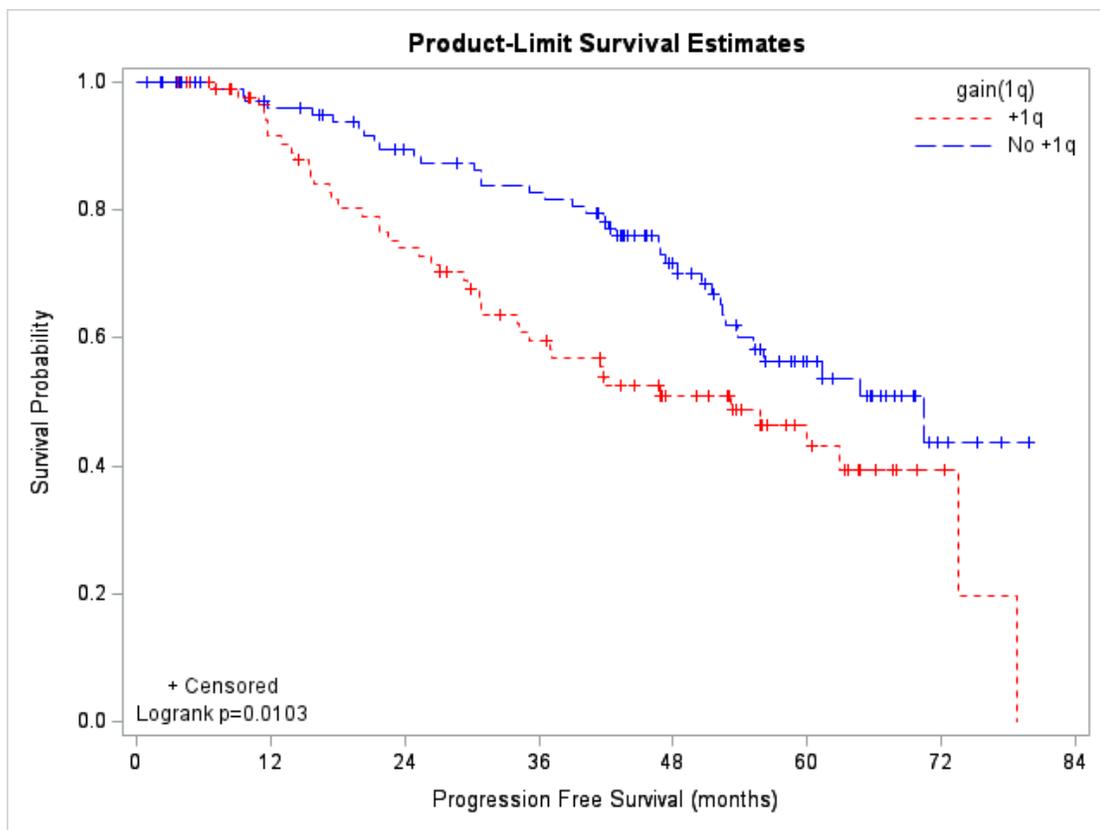


**Figure 1 Best response to RVD and ASCT**

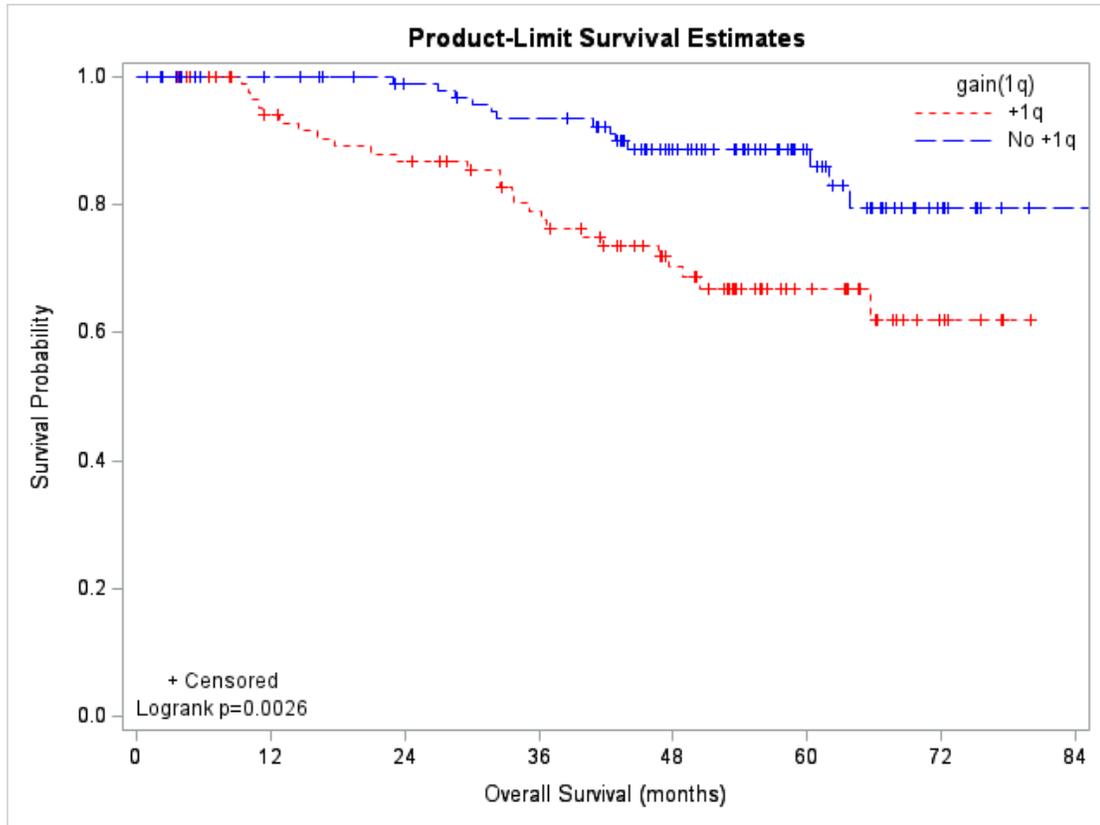
Best response to the RVD induction and ASCT therapy was categorized as complete response (CR), Very Good Partial Response (VGPR), Partial Response (PR). Best response rate, which was defined as the percentage of patients achieving at least CR, VFPR and PR to RVD induction and ASCT therapy are shown on Fig. 1. Among the patients who received RVD induction treatment, overall response rate was similar for patients with +1q and patients without +1q (98.9% vs. 98.1%). Patients with +1q were more likely to achieve a VGPR response level or better conditions compared to patients without +1q (74.7% vs. 60.7%,  $p=0.015$ ).

There are 76 out of 92 (82.6%) patients with +1q and 89 out of 107 (83.2%) patients without +1q underwent ASCT therapy after receiving RVD induction. The overall response rate was also similar for patients with +1q and patients without +1q (98.7% vs. 100%). Patients with +1q were more likely to achieve a VGPR response level or better conditions compared to patients without +1q (96.0% vs. 84.3%,  $p=0.007$ ).

### 3.2.2 Survival outcomes



**Figure 2 Progression-free survival**



**Figure 3 Overall survival**

Estimated Kaplan Meier survival curves (Fig. 1, 2) for PFS and OS were plotted for patients with +1q and patients without +1q. The Kaplan Meier curves visualized the difference of a survival outcome between two groups. Median PFS for patients with +1q was 53.2 months (95% CI 34.2-78.7 mo) compared 70.5 months (95% CI 52.7 mo-undefined) for patients without +1q ( $p=0.010$ ,  $HR=1.76$ ). Patients with +1q had a 5-year PFS rate of 43.1% (95% CI 30.2-55.3%) compared to 56.3% (95% CI 43.9-67.0%) for patients without +1q. Median OS was undefined in either group. The Kaplan-Meier curves shows clear early separation and the patients with +1q had significantly worse OS rates compared to those without +1q ( $p=0.003$ ,  $HR=2.68$ ). Five-year OS rate of patients with +1q was 66.9% (95% CI 54.8-76.4%) compared to 88.7% (95% CI 79.9-93.7%) for patients without +1q.

**Table 3 multivariate survival analysis**

Covariate	Level	N	Hazard Ratio (95% CI)	HR P-value
Calcium	>10.5	111	2.37 (1.01-5.54)	<b>0.047</b>
	<=10.5	65	-	-
Maintenance	No	26	2.22 (1.06-4.67)	<b>0.035</b>
	Yes	70	-	-
T(14,16)	Yes	157	2.12 (0.99-4.57)	0.054
	No	40	-	-
Del(17p)	Yes	155	4.55 (1.82-11.11)	<b>0.001</b>
	No	35	-	-
gain(1q)	+1q	112	1.59 (0.91-2.77)	0.105
	No +1q	69	-	-

The results of univariate survival analysis are in **Supplementary Table 2**. The group with large count was coded as reference group. The risk factors significantly associated with worse PFS diagnosed from the univariate analysis were calcium > 10.5 mg/dL, lack of maintenance therapy, t(14;16), del(17p), complex karyotype, and +1q.

Backward selection with an alpha level of removal of 0.1 was used and multivariate analysis results are shown in **Table 3**. Adjusted hazard ratios and p-values were calculated. The Cox model took the following form:

$$h(t|\mathbf{X}) = h_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5)$$

Where  $h_0(t)$  was the baseline hazard function;  $\beta_1$  was coefficient for calcium that larger than 10.5 mg/dL, and  $Z_1$  was covariate for calcium;  $\beta_2$  was coefficient for receiving maintenance therapy, and  $Z_2$  was covariate for maintenance therapy;  $\beta_3$  was coefficient for t(14;16) existed, and  $Z_3$  was covariate for t(14;16);  $\beta_4$  was coefficient for del(17p)

existed, and  $Z_4$  was covariate for del(17p);  $\beta_5$  was coefficient for +1q, and  $Z_5$  was covariate for gain of 1q.

**Table 4 Test for proportional hazard assumption based on the Schoenfeld residuals from the Cox model (see table 3)**

Covariate	HR P-value
Calcium	0.690
Maintenance	<b>0.002</b>
T(14,16)	<b>0.016</b>
Del(17p)	0.664
gain(1q)	0.163

**Table 5 Extended Cox proportional hazard model summary**

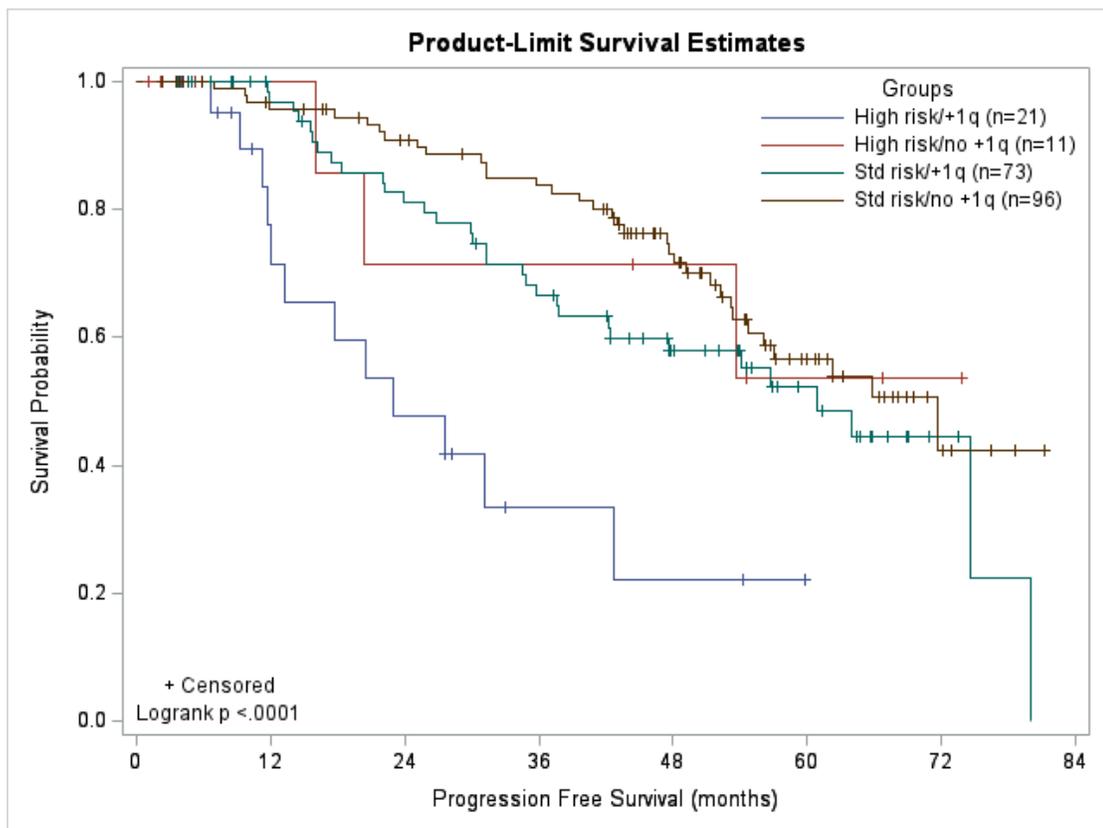
Covariate	level	Parameter Estimate	Hazard Ratio	HR P-value
Calcium	>10.5	0.96	2.62	0.026
Maintenance	No	8.71	Not constant	<0.001
T(14,16)	Yes	7.35	Not constant	0.008
Del(17p)	Yes	0.74	2.10	0.548
gain(1q)	+1q	0.45	1.58	0.110
Maintenance*g(t)		-2.51	Not constant	0.003
T(14;16)*g(t)		-2.02	Not constant	0.037

After proportional hazard assumption assessment, the correlation coefficients for calcium, del(17p), and gain(1q) are not significant ( $p=0.690$ ,  $0.664$ ,  $0.163$ , respectively), suggesting all predictors satisfy the proportional hazard assumption. However, the p-value for maintenance and t(14;16) ( $p= 0.002$ ,  $0.016$ , respectively) are significant. This result suggests that maintenance and t(14;16) does not satisfy proportional hazard assumption. Thus, we defined two interaction terms and use  $g(t) = \log(t)$  as function of time:

$$h(t, \mathbf{X}(t)) = h_0(t) \exp\{\beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \delta_1 X_2 \cdot g(t) + \delta_2 X_3 \cdot g(t)\}$$

The Cox model summary are demonstrated in **Table 5**. The hazard ratio for maintenance and t(14;16) are changing over time. The association between gain 1q and overall survival has a hazard ratio of 1.76 (p=0.011) for +1q compare to no +1q in the univariate analysis, and after controlling for the other covariates of interest, the hazard ratio is reduced to 1.58 (p=0.110) in the multivariable model. It indicates that patients without +1q may have prolonged survival than those have +1q abnormalities.

### 3.2.3 Impact of co-occurring cytogenetic abnormalities



**Figure 4 PFS of patients stratified by different risk**

Thirty-two of 201 patients in this cohort were identified as high-risk population (defined by the co-occurring of t(14;16) and/or del(17p) by FISH) and the rest patients were identified as standard-risk population (“Std”, defined by no occurring of t(14;16) and/or del(17p) by FISH). There were 21 patients with +1q among high-risk patients. Kaplan-Meier curves for PFS of patients stratified by different risk groups are shown in Fig. 4. Median PFS of high-risk patients with +1q was 22.9 months (95% CI 12.0-42.7mo), which was not significantly worse than high-risk patients without +1q (n=11, median PFS was undefined, p=0.115). Patients with +1q who did not have high-risk cytogenetic abnormalities (n=73) had a median PFS of 60.9 months (95% CI 42.2-80.0mo), which was significantly better than high-risk patients with +1q (p=0.001). Median PFS of patients without +1q and high-risk abnormalities (n=96) was 71.6 months (95%CI 54.7-undefined), which was similar to the patients with +1q but without high-risk cytogenetic abnormalities (p=0.794). Patients with standard-risk cytogenetic abnormalities had no significant difference in PFS between those with +1q and without +1q (p=0.118).

### 3.2.4 Impact of copy number and detection threshold

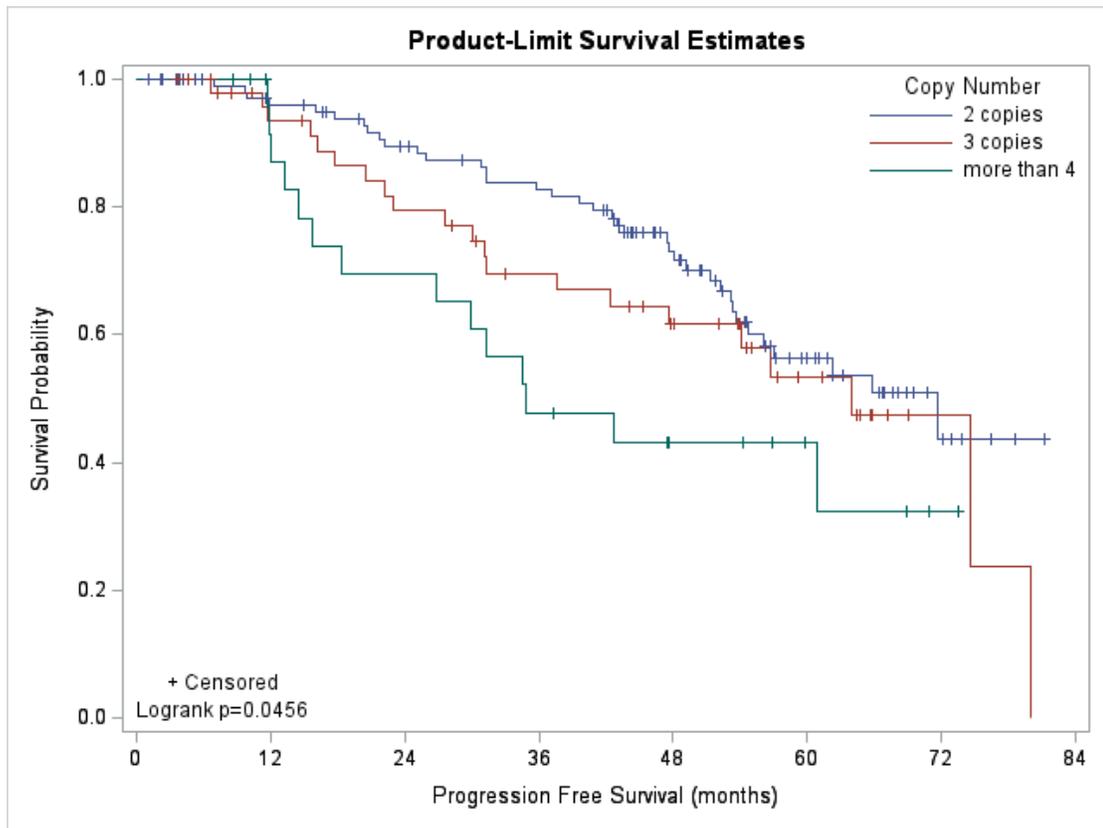


Figure 5 PFS of patients stratified by different copy number

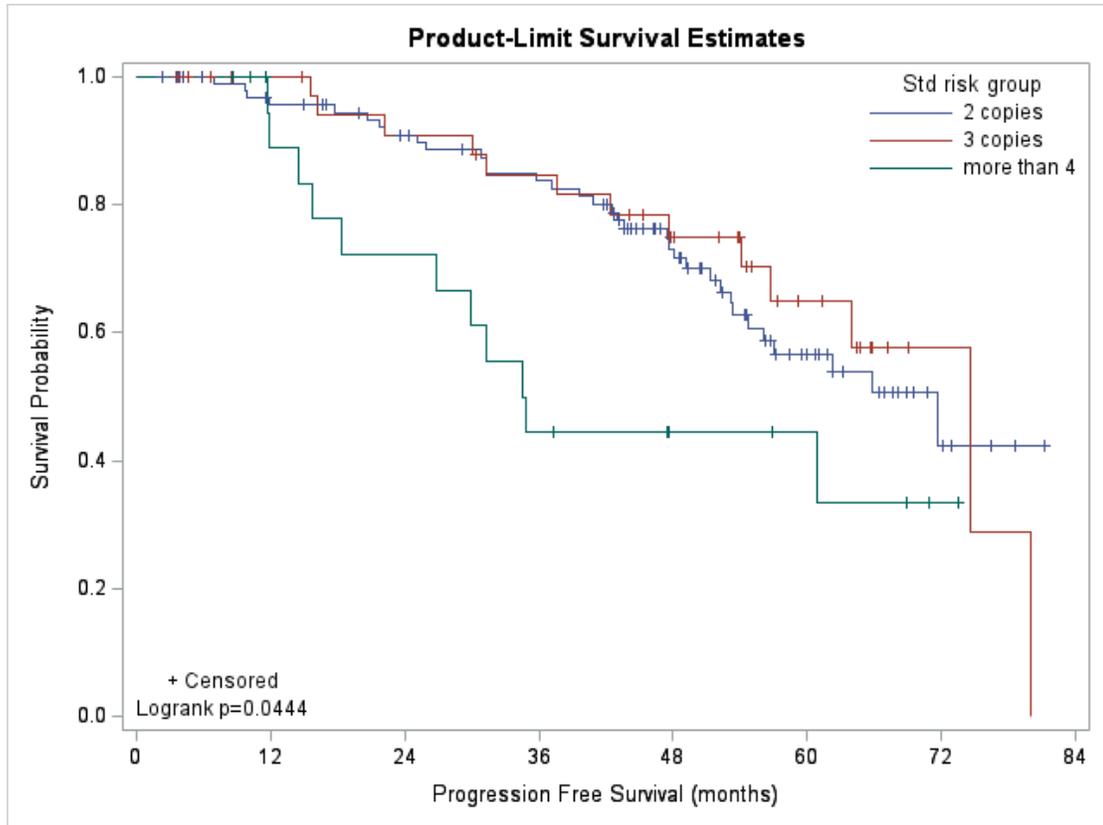
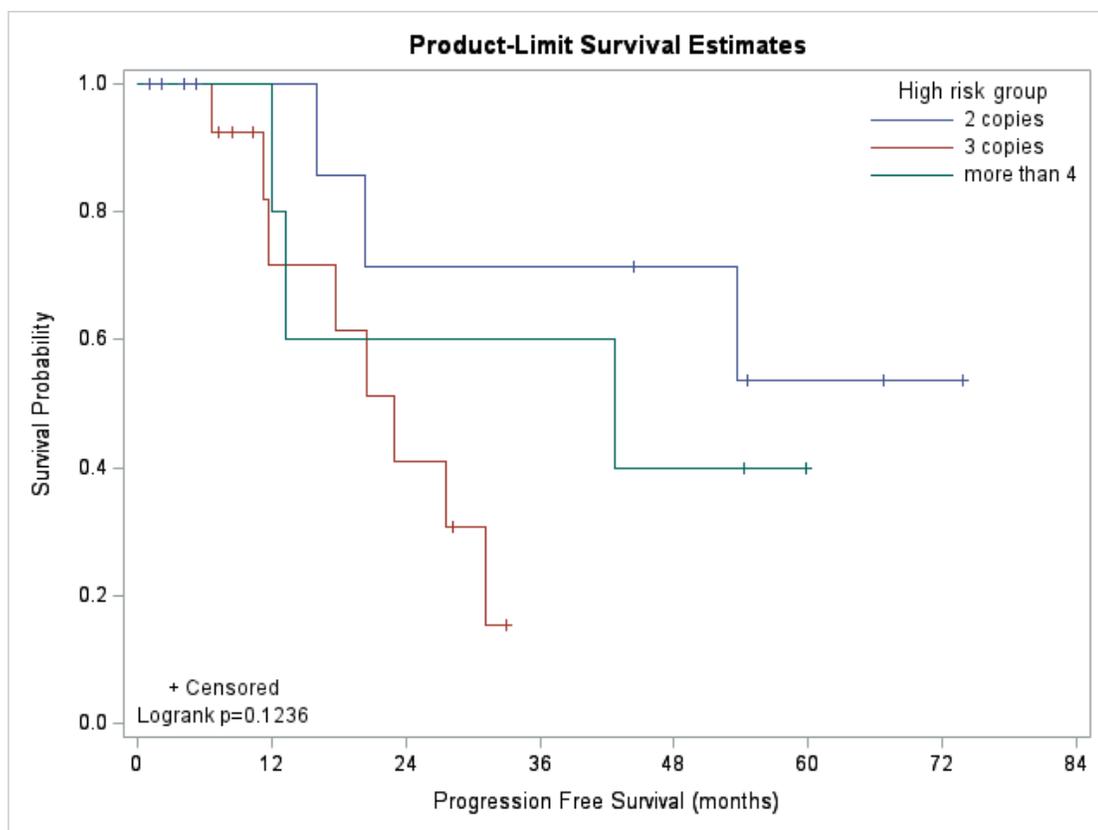


Figure 6 PFS of standard-risk patients stratified by different copy number



**Figure 7 PFS of high-risk patients stratified by different copy number**

Gain(1q) means there two more copies of chromosome 1q in MM patients. Among +1q patients, whether 1q copy number, percentage of cells with +1q by FISH, and co-occurrence of del(1p) impact survival outcomes were determined. The findings of copy number impact are shown in Fig. 5. 78 out of 94 (83.0%) patients with +1q had more than 2 copies quantified in the FISH. In 78 patients with quantified copies, 52 (66.7%) patients had only one additional chromosome 1q copy (three copies) and 26 (33.3%) had two or more extra copies (four or more copies/amp(1q)) at the time of diagnosis. Median survival time of patients with two copies was 71.63 months (95% CI 53.6mo-undefined), which is significantly better than 34.8 months (95% CI 18.4-undefined) for patients with

four or more copies ( $p=0.015$ ). However, there is no significant association between patients with two copies and three copies ( $p=0.270$ ).

Fig. 6 and Fig. 7 show the impact of 1q copy number by detected cytogenetic risk.

Among patients with standard cytogenetic risk, those with 2, 3 and 4 or more copies had a median PFS of 71.6 months (95% CI 54.67mo-undefined), 74.7 months (56.7-80.0mo), and 34.67 months (18.4mo-undefined), respectively. Among patients with high cytogenetic risk and 2 copies, a median PFS was not defined. Patients with 3 and 4 or more copies had a median PFS of 22.9 months (95% CI 11.2-31.2mo) and 42.7months (12.0mo-undefined), respectively. There is no significant association between copy number among high-risk patients ( $p=0.124$ ), but among standard-risk patients, those with normal copy number had significantly better PFS compared to patients with amp(1q) ( $p=0.044$ ).

There was no significant difference in PFS for +1q patients that had 20% or more evaluated positive cells compared to those had less than 20% positive cells ( $p=0.364$ ).

There was also no significant difference in PFS for patients with +1q and del(1p) compared to patients with +1q only ( $p=0.985$ ). Of 196 patients with +1q who were also tested for del(13q), 84 patients also had del(13q) and there was significant difference in PFS between these patients and those patients with +1q alone ( $p=0.054$ ).

#### **4. Discussion**

We identified that 55.7% of newly diagnosed MM patients had +1q abnormalities tested by FISH. Patients with +1q at diagnosis were more likely to have higher calcium

concentration and develop anemia or thrombocytopenia in terms of lower hemoglobin and platelets; have high disease burden in terms of beta-2-microglobulin and a higher R-ISS stage. Additionally, patients with +1q were more likely to have co-occurrence of t(4;14), t(14;16), del(13q), del(1p), and complex karyotype at diagnosis but no significant association with del(17p).

In our univariate survival analysis, calcium concentration, maintenance therapy, +1q, t(14;16), del(17p), and complex karyotype were significantly associated with PFS respectively. However, in the Cox PH model, only variables of calcium, maintenance, t(14;16), del(17p) and gain(1q) were shown to be prognostic factors of the patient's PFS. Patients with lower calcium concentration, maintenance therapy tended to have better PFS outcomes. Patients with extra 1q copies and co-occurrence of t(14;16), del(17p), and/or complex karyotype had worse PFS outcomes. Based on the results, we defined the patients with co-occurrence of t(14;16), del(17), and/or complex karyotype were high-risk patients.

We found that among patients treated with RVD therapy, those with +1q had a significantly higher chance to achieve a CR or VGPR. This impact on prognosis remained significant among patients underwent ASCT after RVD therapy and had improved response rate. Despite the greater response rate, patients with +1q still had significantly shorter PFS and OS compared to those without +1q. The results from fig. 6 and fig. 7 show that there was no significant difference in PFS between standard-risk patients with or without one extra 1q copy, but all patients with amp(1q) and high-risk patients with extra 1q copies had a substantial reduction in PFS. These findings suggest

that whether 1q chromosome abnormalities a driver of these inferior outcomes is still unclear.

In conclusion, patients with four or more copies of 1q and/or co-occurrence of high-risk cytogenetic abnormalities should be considered to have a high risk of early progression or death and be treated with aggressive therapies and/or clinical trials early in their courses. The limitation of our analysis is that this is a retrospective study and the data were collected in a single center. Patients selection bias occurred due to the patients that referral to this center were healthy enough to accept transplantation. A lot of missing data regarding key abnormalities also leads to the inaccurate of the final model and results.

## 5. Reference

- Avet-Loiseau, H., Li, C., Magrangeas, F., Gouraud, W., Charbonnel, C., Harousseau, J.-L., . . . Minvielle, S. (2009). Prognostic Significance of Copy-Number Alterations in Multiple Myeloma. *Journal of Clinical Oncology*, 27(27), 4585-4590.  
doi:10.1200/jco.2008.20.6136
- Chang, H., Qi, X., Jiang, A., Xu, W., Young, T., & Reece, D. (2010). 1p21 deletions are strongly associated with 1q21 gains and are an independent adverse prognostic factor for the outcome of high-dose chemotherapy in patients with multiple myeloma. 45(1), 117-121.  
doi:10.1038/bmt.2009.107
- Dimopoulos, M. A., Kastritis, E., Christoulas, D., Migkou, M., Gavriatopoulou, M., Gkotzamanidou, M., . . . Terpos, E. (2010). Treatment of patients with relapsed/refractory multiple myeloma with lenalidomide and dexamethasone with or without bortezomib:

- prospective evaluation of the impact of cytogenetic abnormalities and of previous therapies. *Leukemia*, 24(10), 1769-1778. doi:10.1038/leu.2010.175
- Fonseca, R., Van Wier, S. A., Chng, W. J., Ketterling, R., Lacy, M. Q., Dispenzieri, A., . . . Gertz, M. A. (2006). Prognostic value of chromosome 1q21 gain by fluorescent in situ hybridization and increase CKS1B expression in myeloma. *Leukemia*, 20(11), 2034-2040. doi:10.1038/sj.leu.2404403
- Hanamura, I., Stewart, J. P., Huang, Y., Zhan, F., Santra, M., Sawyer, J. R., . . . Shaughnessy, J. D. (2006). Frequent gain of chromosome band 1q21 in plasma-cell dyscrasias detected by fluorescence in situ hybridization: incidence increases from MGUS to relapsed myeloma and is related to prognosis and disease progression following tandem stem-cell transplantatio. *Blood*, 108(5), 1724-1732. doi:10.1182/blood-2006-03-009910
- Marzin, Y. Chromosome 1 Abnormalities in Multiple Myeloma.
- Nemec, P., Zemanova, Z., Greslikova, H., Michalova, K., Filkova, H., Tajtlova, J., . . . Hajek, R. (2010). Gain of 1q21 Is an Unfavorable Genetic Prognostic Factor for Multiple Myeloma Patients Treated with High-Dose Chemotherapy. *Biology of Blood and Marrow Transplantation*, 16(4), 548-554. doi:10.1016/j.bbmt.2009.11.025
- Palumbo, A., & Anderson, K. (2011). Multiple Myeloma. *New England Journal of Medicine*, 364(11), 1046-1060. doi:10.1056/nejmra1011442
- Palumbo, A., Avet-Loiseau, H., Oliva, S., Lokhorst, H. M., Goldschmidt, H., Rosinol, L., . . . Moreau, P. (2015). Revised International Staging System for Multiple Myeloma: A Report From International Myeloma Working Group. *Journal of Clinical Oncology*, 33(26), 2863-2869. doi:10.1200/jco.2015.61.2267
- Rajkumar, S. V. (2018). Multiple myeloma: 2018 update on diagnosis, risk-stratification, and management. *American Journal of Hematology*, 93(8), 1091-1110. doi:10.1002/ajh.25117

- Samo, A. A., Li, J., Zhou, M., Sun, Y., Yang, Y., Zhang, Y., . . . Fan, X. (2018). MCL1 gene co-expression module stratifies multiple myeloma and predicts response to proteasome inhibitor-based therapy. *Genes, Chromosomes and Cancer*, *57*(8), 420-429. doi:10.1002/gcc.2
- Sawyer, J. R. (2011). The prognostic significance of cytogenetics and molecular profiling in multiple myeloma. *Cancer Genetics*, *204*(1), 3-12. doi:10.1016/j.cancergencyto.2010.11.002
- Shah, G. L., Landau, H., Londono, D., Devlin, S. M., Kosuri, S., Lesokhin, A. M., . . . Giralt, S. A. (2017). Gain of chromosome 1q portends worse prognosis in multiple myeloma despite novel agent-based induction regimens and autologous transplantation. *Leukemia & Lymphoma*, *58*(8), 1823-1831. doi:10.1080/10428194.2016.1260126
- Shaughnessy, J. (2005). Amplification and overexpression of CKS1B at chromosome band 1q21 is associated with reduced levels of p27 Kip1 and an aggressive clinical course in multiple myeloma. *10*(Supplement-1), 117-126. doi:10.1080/10245330512331390140
- Smetana, J., Berankova, K., Zaoralova, R., Nemecek, P., Greslikova, H., Kupska, R., . . . Kuglik, P. (2013). Gain(1)(q21) is an Unfavorable Genetic Prognostic Factor for Patients With Relapsed Multiple Myeloma Treated With Thalidomide but Not for Those Treated With Bortezomib. *13*(2), 123-130. doi:10.1016/j.clml.2012.11.012
- Sonneveld, P., Avet-Loiseau, H., Lonial, S., Usmani, S., Siegel, D., Anderson, K. C., . . . Orłowski, R. (2016). Treatment of multiple myeloma with high-risk cytogenetics: a consensus of the International Myeloma Working Group. *Blood*, *127*(24), 2955-2962. doi:10.1182/blood-2016-01-631200
- Teoh, P. J., An, O., Chung, T.-H., Chooi, J. Y., Toh, S. H. M., Fan, S., . . . Chng, W. J. (2018). Aberrant hyperediting of the myeloma transcriptome by ADAR1 confers oncogenicity

and is a marker of poor prognosis. *Blood*, 132(12), 1304-1317. doi:10.1182/blood-2018-02-832576

## 6. Appendix

### 6.1 Supplementary Table 1 Patient characteristics

**Table 1 Patient characteristics**

Covariate	Level	No +1q (n=107)	+1q (n=94)	Total	P-value
Age	Median	63	64	64	
	<65	61 (57.01)	52 (55.32)		0.810
	>=65	46 (42.99)	42 (44.68)		
Sex	Male	64 (59.81)	50 (53.19)	113 (56.2)	
	Female	43 (40.19)	44 (46.81)	88 (43.8)	
Race	Unknown/Other	10 (9.35)	6 (6.38)	16 (8.0)	0.186
	African-American	34 (31.78)	21 (22.34)	55 (27.4)	
	Caucasian	63 (58.88)	67 (71.28)	130 (64.7)	
Isotype code	IgG	66 (61.68)	45 (47.87)	16 (8.0)	0.256
	IgA	19 (17.76)	27 (28.72)	55 (27.4)	
	IgD	1 (0.93)	0 (0)	130 (64.7)	
	FLC	20 (18.69)	20 (21.28)	16 (8.0)	
	Nonsecretory	1 (0.93)	1 (1.06)	55 (27.4)	
	Oligosecretory	0 (0)	1 (1.06)	130 (64.7)	
Hemoglobin (g/dL)	Median	11.07	10.09	10.50	
	<10	26 (26.8)	39 (45.88)	65 (35.7)	<b>0.007</b>
	>=10	71 (73.2)	46 (54.12)	117 (64.3)	
Platelets (x10 <sup>3</sup> /μL)	Median	226.63	300.66	207.00	
	<150	11 (12.64)	23 (30.26)	34 (20.9)	<b>0.006</b>
	>=150	76 (87.36)	53 (69.74)	129 (79.1)	
Creatinine (mg/dL)	Median	1.33	1.36	1.02	
	>2.0	8 (8.6)	11 (13.58)	19 (10.9)	0.294
	<=2.0	85 (91.4)	70 (86.42)	155 (89.1)	

**Table 1 Patient characteristics**

Covariate	Level	No +1q (n=107)	+1q (n=94)	Total	P-value
Calcium (mg/dL)	Median	9.37	9.67	9.20	
	>10.5	3 (3.57)	13 (16.25)	16 (9.8)	<b>0.006</b>
	<=10.5	81 (96.43)	67 (83.75)	148 (90.2)	
Albumin (g/dL)	Median	3.6	3.53	3.6	
	<3.5	37 (39.36)	35 (46.67)	72 (42.6)	0.340
	>=3.5	57 (60.64)	40 (53.33)	97 (57.4)	
LDH (units/L)	Median	158.25	174.78	150.00	
	>=ULN	56 (94.92)	58 (92.06)	114 (93.4)	0.525
	<ULN	3 (5.08)	5 (7.94)	8 (6.6)	
$\beta$ 2M (mg/L)	Median	3.31	4.39	2.95	
	>5.5	9 (10.23)	17 (23.61)	26 (16.3)	<b>0.022</b>
	<=5.5	79 (89.77)	55 (76.39)	134 (83.8)	
M-spike (g/dL)	Median	2.51	2.87	2.24	
	>3.0	34 (37.78)	37 (45.12)	71 (41.3)	0.329
	<=3.0	56 (62.22)	45 (54.88)	101 (58.7)	
K/L abnormal	>100 or <0.001	42 (48.84)	46 (60.53)	88 (54.3)	0.136
	0.001 < k/l < 100	44 (51.16)	30 (39.47)	74 (45.7)	
ISS stage	1	39 (45.88)	20 (28.57)	59 (38.1)	<b>0.024</b>
	2	37 (43.53)	33 (47.14)	70 (45.2)	
	3	9 (10.59)	17 (24.29)	26 (16.8)	
R-ISS Stage	1	34 (38.64)	14 (19.44)	48 (30.0)	<b>0.003</b>
	2	53 (60.23)	51 (70.83)	104 (65.0)	
	3	1 (1.14)	7 (9.72)	8 (5.0)	
Upfront Transplant	Yes	88 (82.24)	69 (73.4)	157 (78.1)	0.131
	2	19 (17.76)	25 (26.6)	44 (21.9)	
Maintenance	Yes	83 (79.05)	72 (80)	155 (79.5)	0.870
	2	22 (20.95)	18 (20)	40 (20.5)	
T(11,14)	No	81 (77.88)	76 (86.36)	157 (81.8)	0.129
	Yes	23 (22.12)	12 (13.64)	35 (18.2)	
T(4,14)	No	103 (99.04)	81 (91.01)	184 (95.3)	<b>0.008</b>
	Yes	1 (0.96)	8 (8.99)	9 (4.7)	
T(14,16)	No	103 (99.04)	79 (90.8)	182 (95.3)	<b>0.007</b>
	Yes	1 (0.96)	8 (9.2)	9 (4.7)	

**Table 1 Patient characteristics**

Covariate	Level	No +1q (n=107)	+1q (n=94)	Total	P-value
Del(17p)	No	97 (90.65)	76 (84.44)	173 (87.8)	0.184
	Yes	10 (9.35)	14 (15.56)	24 (12.2)	
Del(13q)	No	79 (74.53)	33 (36.67)	112 (57.1)	<b>&lt;.001</b>
	Yes	27 (25.47)	57 (63.33)	84 (42.9)	
Hyperdiploidy	No	35 (33.65)	34 (39.53)	69 (36.3)	0.401
	Yes	69 (66.35)	52 (60.47)	121 (63.7)	
del(1p)	No	99 (94.29)	73 (82.95)	172 (89.1)	<b>0.012</b>
	Yes	6 (5.71)	15 (17.05)	21 (10.9)	
Complex Karyotype	Yes	15 (14.02)	29 (31.52)	44 (22.1)	<b>0.003</b>
	No	92 (85.98)	63 (68.48)	155 (77.9)	

## 6.2 Supplementary Table 2 Univariate survival analysis

**Table 2 Univariate survival analysis**

Covariate	Level	N	Hazard Ratio (95% CI)	HR P-value	Log-rank P-value
Age	<65	113	0.90 (0.58-1.39)	0.622	0.622
	>=65	88	Ref	-	
Sex	Male	114	1.19 (0.76-1.85)	0.449	0.449
	Female	87	Ref	-	
Race	Unknown/Other	16	0.47 (0.15-1.49)	0.199	0.390
	African-American	55	0.87 (0.52-1.46)	0.595	
	Caucasian	130	Ref	-	
Isotype	IgA	113	1.05 (0.61-1.81)	0.851	0.127
	FLC	88	0.92 (0.51-1.65)	0.768	
	Others	114	3.66 (1.13-11.92)	<b>0.031</b>	
	IgG	87	Ref	-	
Hemoglobin	<10	16	1.17 (0.72-1.90)	0.522	0.522
	>=10	55	Ref	-	
Platelets	<150	130	1.50 (0.84-2.70)	0.174	0.171
	>=150	46	Ref	-	
Creatinine	<=2.0	40	1.81 (0.92-3.56)	0.085	0.080
	>2.0	4	-	-	

**Table 2 Univariate survival analysis**

Covariate	Level	N	Hazard Ratio (95% CI)	HR P- value	Log-rank P-value
Calcium	>10.5	111	3.20 (1.57-6.53)	<b>0.001</b>	<b>&lt;.001</b>
	<=10.5	65	Ref	-	
Albumin	<3.5	117	0.69 (0.41-1.16)	0.162	0.160
	>=3.5	34	Ref	-	
LDH	>=ULN	129	1.95 (0.27-14.18)	0.510	0.502
	<ULN	19	Ref	-	
β2M	>5.5	155	1.01 (0.54-1.90)	0.976	0.976
	<=5.5	16	Ref	-	
M-spike	>3.0	148	0.87 (0.53-1.42)	0.573	0.573
	<=3.0	72	Ref	-	
FLC Ratio	0.001< k/l <100	97	0.61 (0.36-1.03)	0.065	0.062
	>100 or <0.001	114	Ref	-	
ISS	1	8	1.00 (0.57-1.74)	0.990	0.989
	3	26	1.05 (0.53-2.08)	0.893	
	2	134	Ref	-	
R-ISS	1	71	0.79 (0.45-1.38)	0.400	0.681
	3	101	1.08 (0.33-3.47)	0.902	
	2	74	Ref	-	
Upfront Transplant	No	88	1.44 (0.81-2.57)	0.217	0.215
	Yes	59	Ref	-	
Maintenance	No	26	1.97 (1.12-3.47)	<b>0.018</b>	<b>0.016</b>
	Yes	70	Ref	-	
T(11,14)	Yes	48	0.88 (0.48-1.60)	0.666	0.665
	No	8	Ref	-	
T(4,14)	Yes	104	1.76 (0.64-4.81)	0.274	0.268
	No	44	Ref	-	
T(14,16)	Yes	157	3.10 (1.24-7.72)	<b>0.015</b>	<b>0.011</b>
	No	40	Ref	-	
Del(17p)	Yes	155	2.26 (1.19-4.29)	<b>0.013</b>	<b>0.010</b>
	No	35	Ref	-	
Del(13q)	Yes	157	1.53 (0.99-2.37)	0.056	0.054
	No	9	Ref	-	
Hyperdiploidy	No	184	1.49 (0.94-2.37)	0.092	0.090
	Yes	9	Ref	-	

**Table 2 Univariate survival analysis**

<b>Covariate</b>	<b>Level</b>	<b>N</b>	<b>Hazard Ratio (95% CI)</b>	<b>HR P- value</b>	<b>Log-rank P-value</b>
Del(1p)	Yes	182	1.23 (0.61-2.46)	0.564	0.563
	No	24	Ref	-	
Complex Karyotype	Yes	173	2.01 (1.24-3.25)	<b>0.004</b>	<b>0.004</b>
	No	84	Ref	-	
·1q gain	+1q	112	1.76 (1.14-2.73)	<b>0.011</b>	<b>0.010</b>
	No +1q	69	Ref	-	