Biomarker analysis report

2017-08-29

Contents

1 The dataset

The dataset have 368 entries. In which 309 are in biomarker evaluable population (BEP).

- Endpoint of interest: Response
- Biomarker: KRAS.exprs
- Biomarker type: numeric

2 Representativeness: Selection Bias of Biomarker Population

In this section, we are trying to answer the question: *Are biomarker evaluable population representative of the full population population?*

Key baseline demographics and prognostic characteristics (including stratification variables and any variables with known prognostic effect) and efficacy outcomes should be summarized by treatment groups and compared between biomarker evaluable population (BEP) and the full population. These analyses are conducted to investigate any potential selection bias associated with the availability of the biomarker (e.g. we may not get enough tissue for patients whose tumor size is small. Therefore they may be exlcuded from BEP).

The key clinical biomarkers considered are:

```
clinical.vars.class
```

```
##           Sex           Age
## "categorical"     "numeric"
```

2.1 Check selection bias in terms of key clinical variables, between full population and BEP

```
kable(SummaryVars(data=input, trt=trt, subgroup=BEP, subgroup.indicator = BEP.indicator,
                  var=clinical.vars, var.class=clinical.vars.class))
```

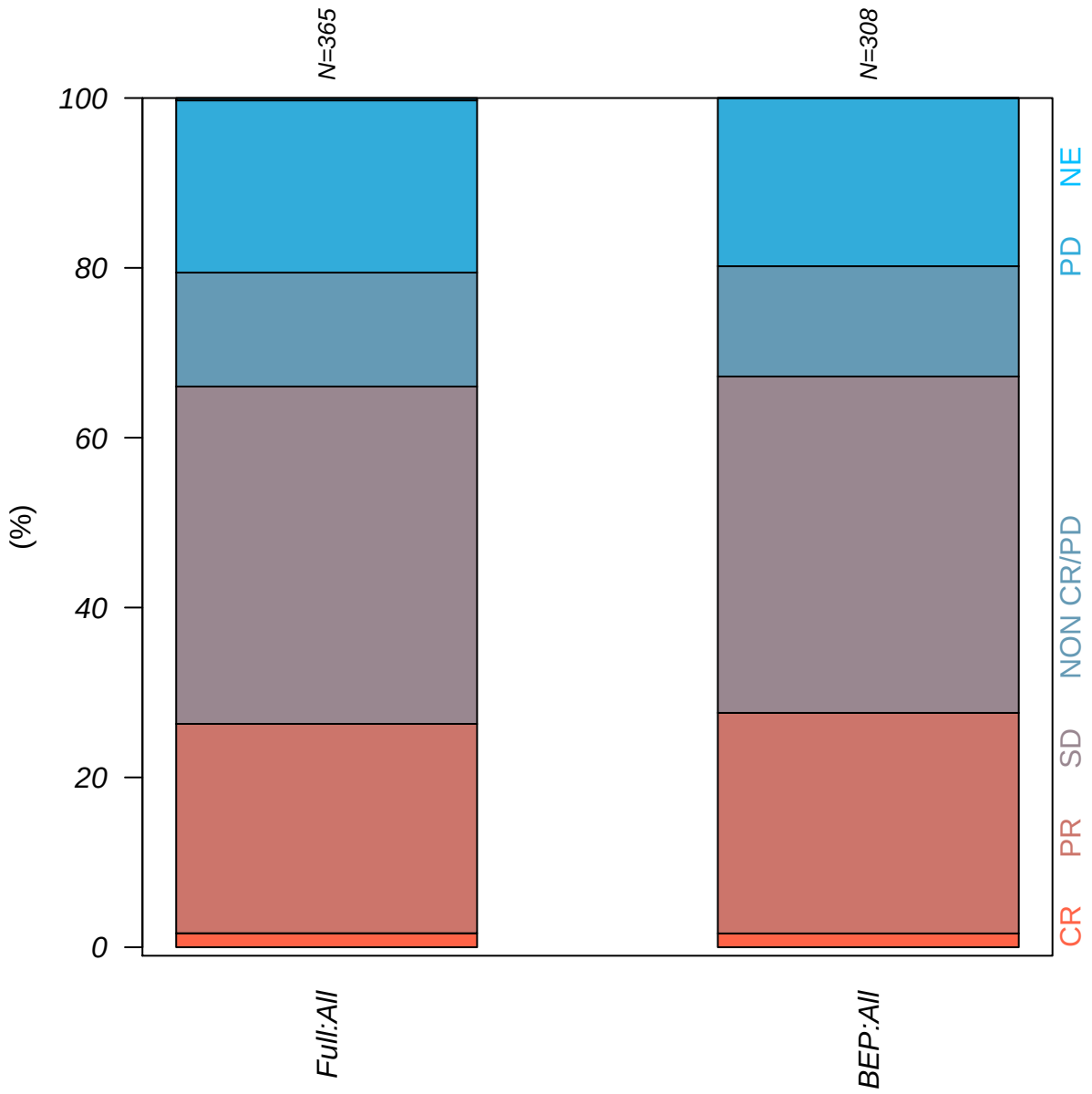
	All	BEP
Sex		
Total	368	309
NA's	0	0
F	184 (50%)	147 (47.57%)
M	184 (50%)	162 (52.43%)
Age		
N	368	309
Mean	54.03	54.29
Median	54	54
Min-Max	30...89	30...89
NA's	0	0

2.2 Check whether the clinical outcome in BEP is comparable to the full population

The following plot compares response outcome in BEP vs. the full population. The response category distribution is plotted for each arm. The BEP bars are expected to be comparable to those in full population.

```
Rsp.out <- PlotRspBar(data=input, outcome.var=outcome.var, binary=FALSE,
rsp.levels=rsp.levels, trt=trt, bep.indicator=BEP.indicator,
compare.bep.itt=TRUE, bep = BEP)
```

```
## entries with missing outcome.var are removed! 365 entries left
```




```
kable(Rsp.out$count,caption="count")
```

Table 2: count

	CR	PR	SD	NON CR/PD	PD	NE
Full:All	6	90	145	49	74	1
BEP:All	5	80	122	40	61	0

```
kable(round(Rsp.out$perc,2), caption="percentage")
```

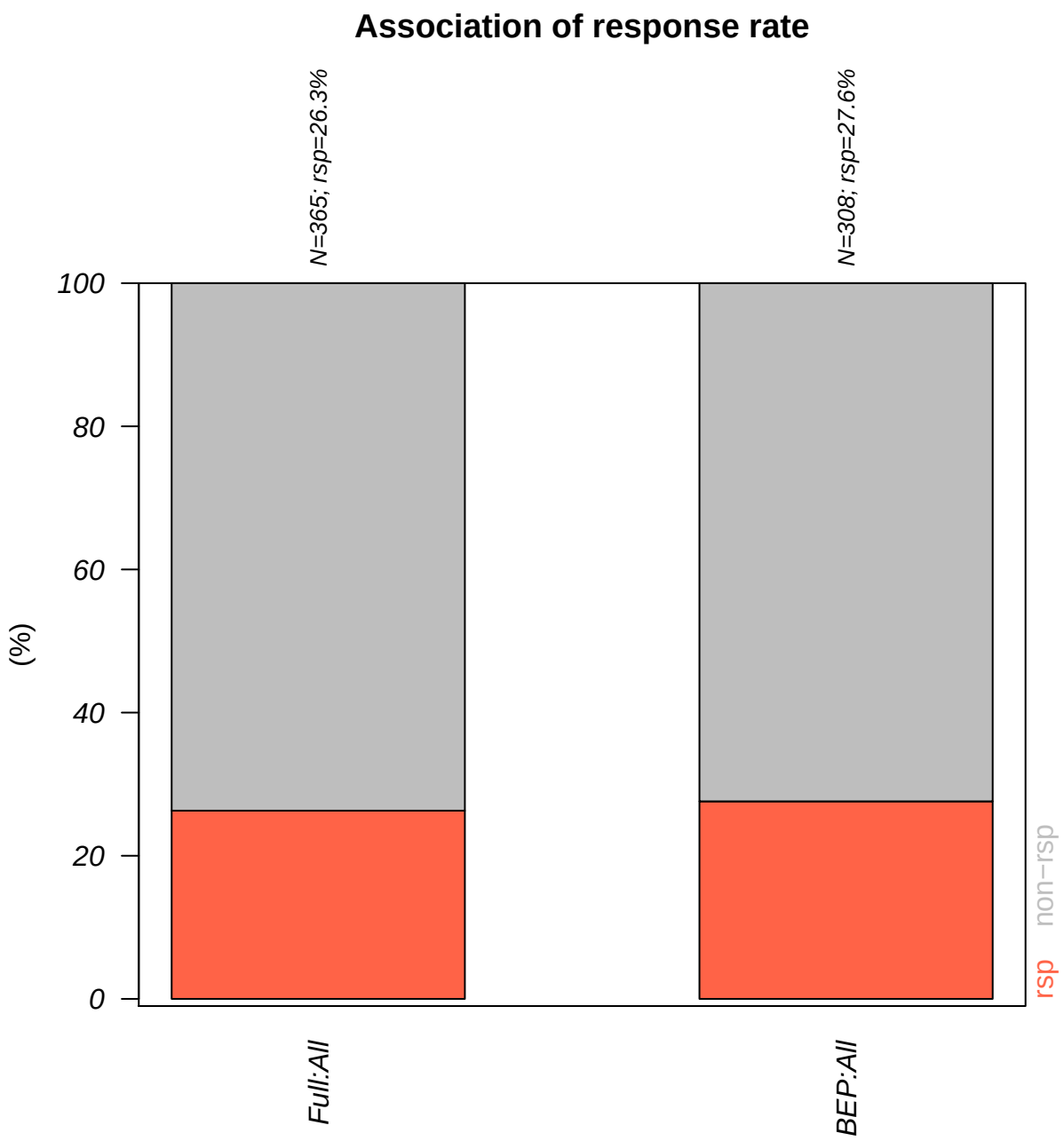
Table 3: percentage

	CR	PR	SD	NON CR/PD	PD	NE
Full:All	0.02	0.25	0.4	0.13	0.2	0
BEP:All	0.02	0.26	0.4	0.13	0.2	0

Summarize the response in two classes - rsp vs. non-rsp:

```
Rsp.out.2 <- PlotRspBar(data=input, outcome.var=outcome.var, binary=TRUE,
rsp.response = rsp.response,
rsp.nonresponse = rsp.nonresponse,rsp.name=rsp.name, trt=trt, bep.indicator=BEP.indicator,
compare.bep.itt=TRUE, bep = BEP)
```

```
## entries with missing outcome.var are removed! 365 entries left
```




```
kable(Rsp.out.2$count,caption="count")
```

Table 4: count

	rsp	non-rsp
Full:All	96	269
BEP:All	85	223

```
kable(round(Rsp.out.2$perc,2), caption="percentage")
```

Table 5: percentage

	rsp	non-rsp
Full:All	0.26	0.74
BEP:All	0.28	0.72

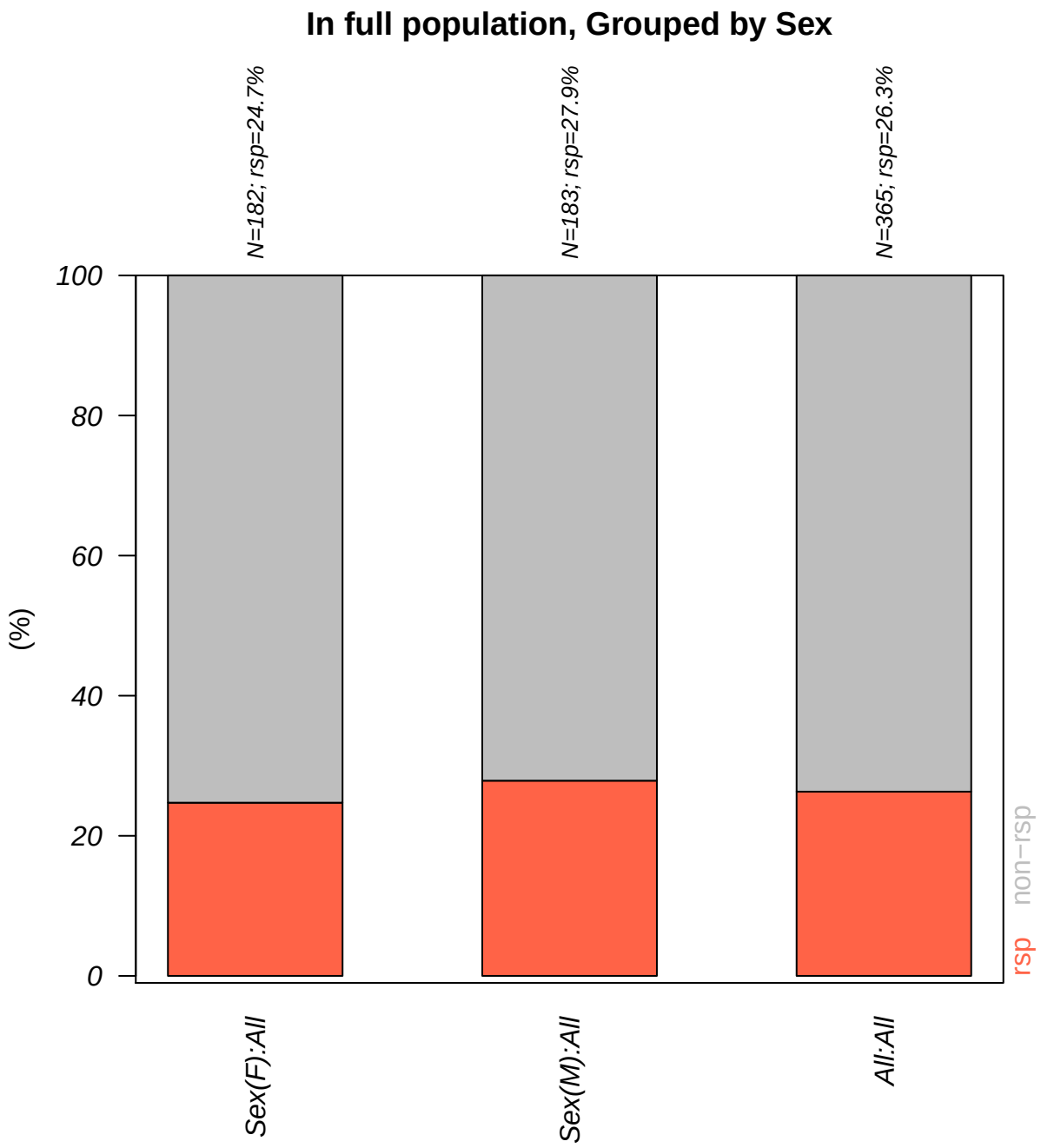
2.3 Examine whether the prognostic/null trend of key clinical variables holds in BEP

The following forest plot can be used to examine whether any of the key prognostic clinical variables still show prognostic trend in BEP:

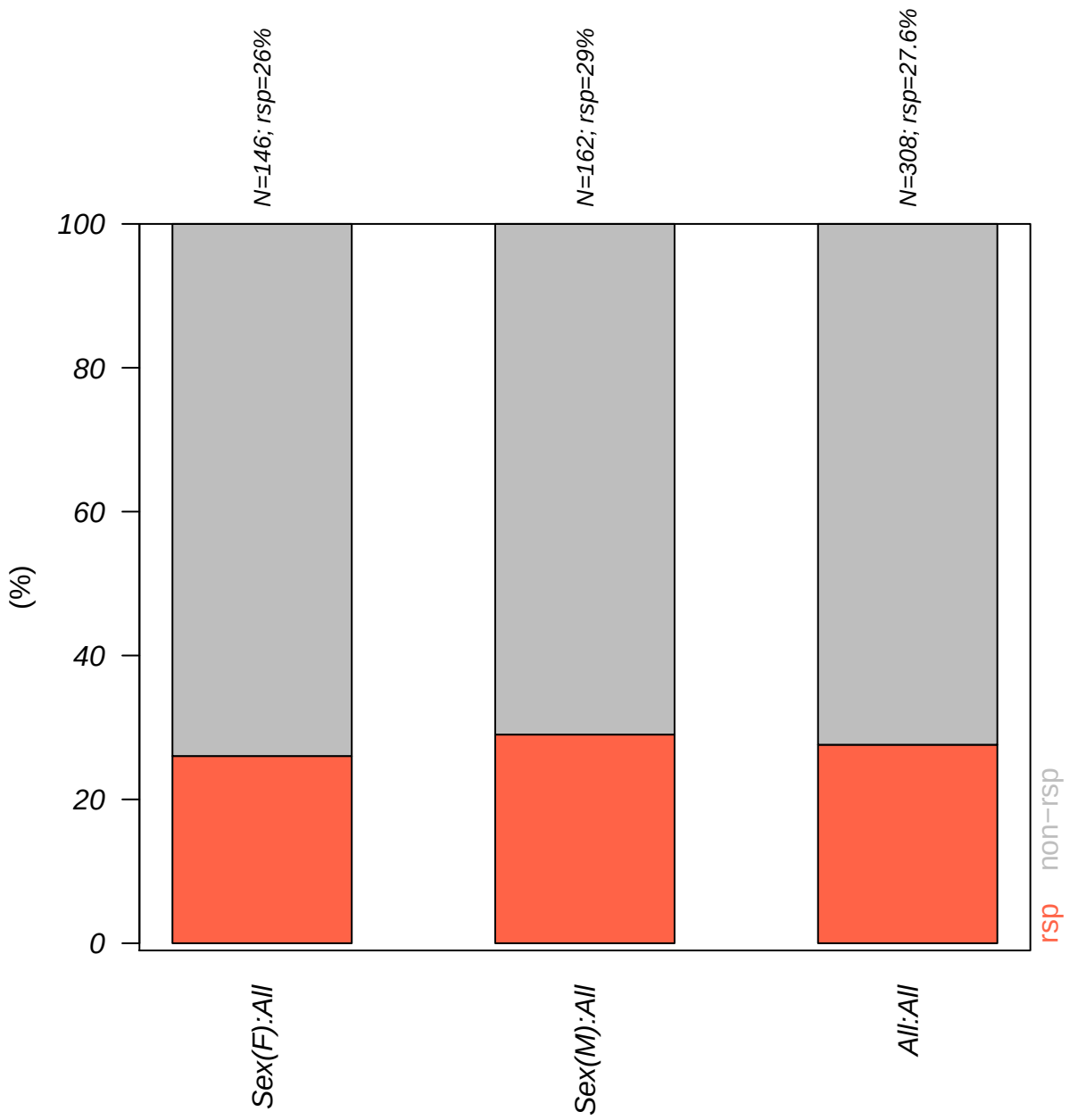
```
for(vv in clinical.vars.2){

  Rsp.out.tmp <- PlotRspBar(data=input, outcome.var=outcome.var, binary=TRUE,
  rsp.response = rsp.response,
  rsp.nonresponse = rsp.nonresponse,rsp.name=rsp.name, trt=trt,
  compare.var =TRUE,  var=vv, main=paste("In full population, Grouped by", vv))

  Rsp.out.tmp <- PlotRspBar(data=input.bep,
                            outcome.var=outcome.var, binary=TRUE,
  rsp.response = rsp.response,
  rsp.nonresponse = rsp.nonresponse,rsp.name=rsp.name, trt=trt,
  compare.var =TRUE, var=vv, main=paste("In BEP, Grouped by", vv))
}
```

```
## entries with missing outcome.var are removed! 365 entries left
```

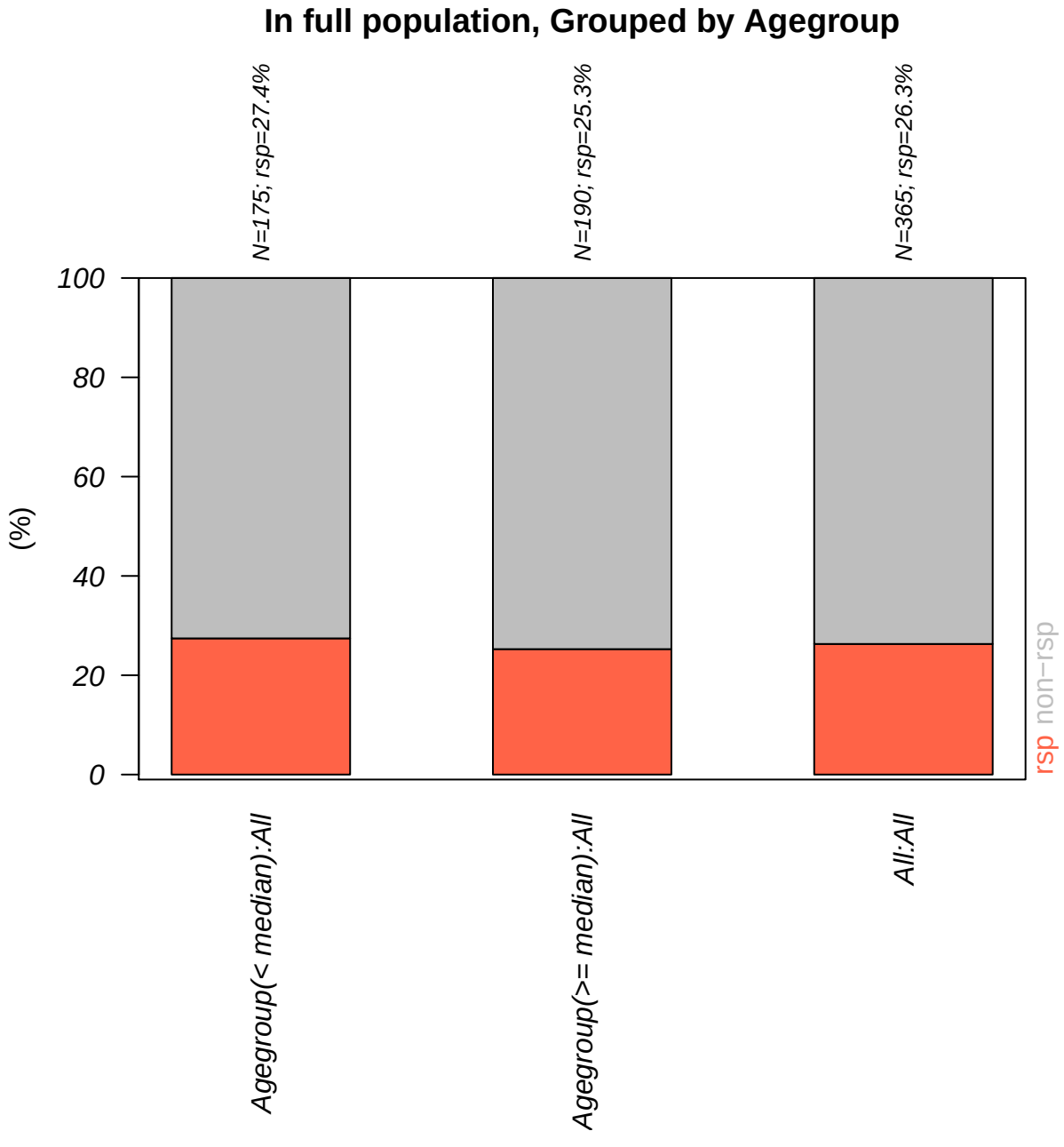
```
## entries with missing outcome.var are removed! 308 entries left
```

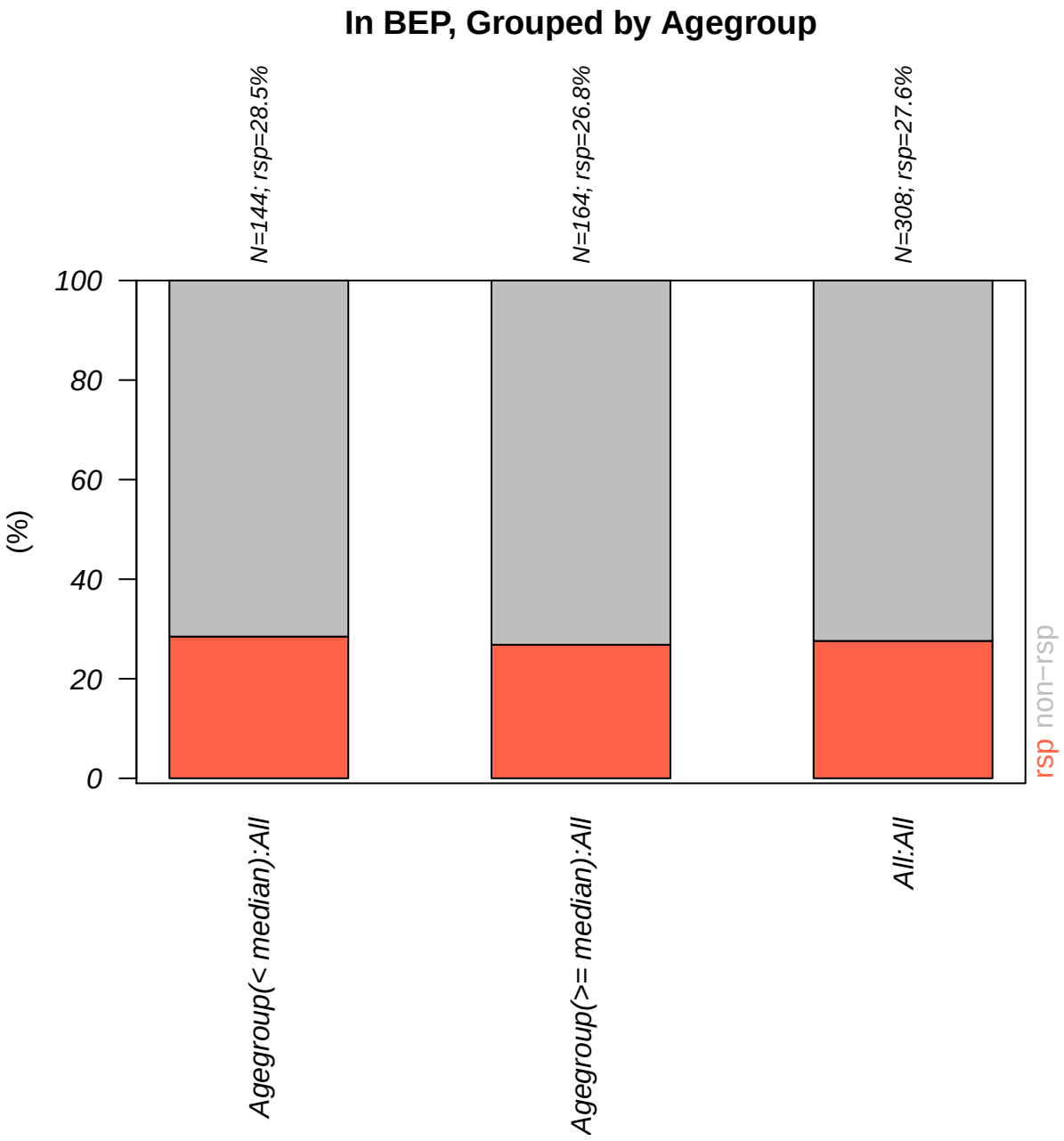



```
## entries with missing outcome.var are removed! 365 entries left
```




```
## entries with missing outcome.var are removed! 308 entries left
```


In full population, Grouped by Agegroup
N=175; rsp=27.4%
N=190; rsp=25.3%
N=365; rsp=26.3%
100
80
60
40
20
0
(%)
Agegroup(< median):All
Agegroup(>= median):All
All:All
rsp
non−rsp



3 Biomarker property and its association to clinical variables

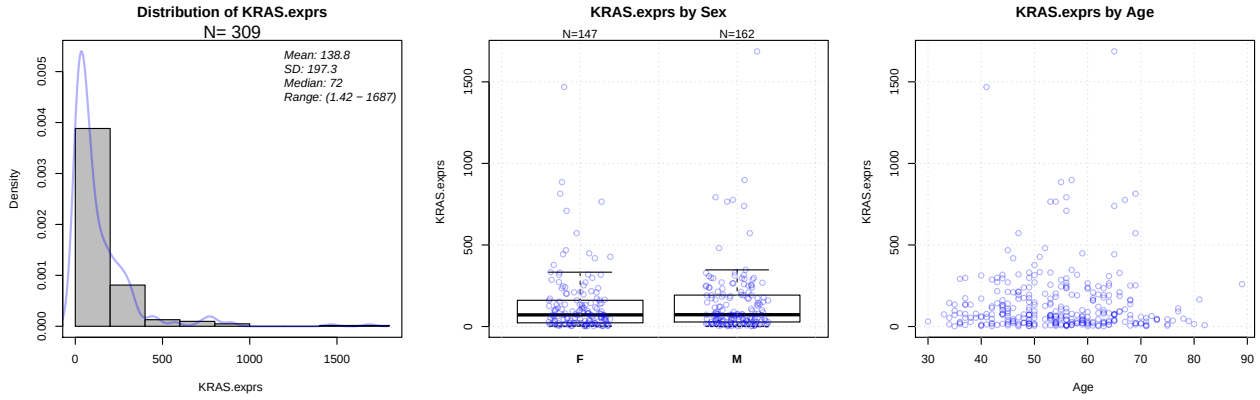
Before performing cutoff exploraotry analysis, it is important to check a biomarker's property. For example, whether this biomarker has a bi-modal or multi modal distribution - if so, this biomarker may has natural cutoff.

Relationship between the biomarker and key demographic and prognostic variables should also be investigated using bivariate plots. Prognostic property of the biomarker should also be assessed, by estimates of the clinical efficacy in the control arm.

3.1 Biomarker property and relationship to clinical variable

The following results show single variate plot for biomaker and bi-variate plots to investigate biomarker-clinical variable relationship.

```
PlotProperty(data=input, biomarker.var=bm, biomarker.class=bm.class,
             var=clinical.vars,
             var.class=clinical.vars.class,
             log2=FALSE, par.param = list(mfrow=c(2,3)))
```



4 Biomarker cutoff exploration/selection

Results in this section could be used to examine multiple candidate cutoffs for a continuous biomarker. The need for cut-off determination should be rooted in the development strategy. In general, an exhaustive search looking at all possible cut-off values is not recommended for decision making. Over-optimized cutoff using one set of clinical data may lead to hard-to-reproduce results. When determining a cutoff, biomarker property should be considered - e.g. cut at a low-dense point may be more robust to population shift. The cutoff selection should also fit the program's stratigitic considerations. There is always a prevalence-effect size trade-off, inputs from multiple functions are needed - for example whether the team is willing to take more risk in PTS (high prevalence, weaker signal) or the team is willing to target at smaller population (lower prevalence, stronger signal)

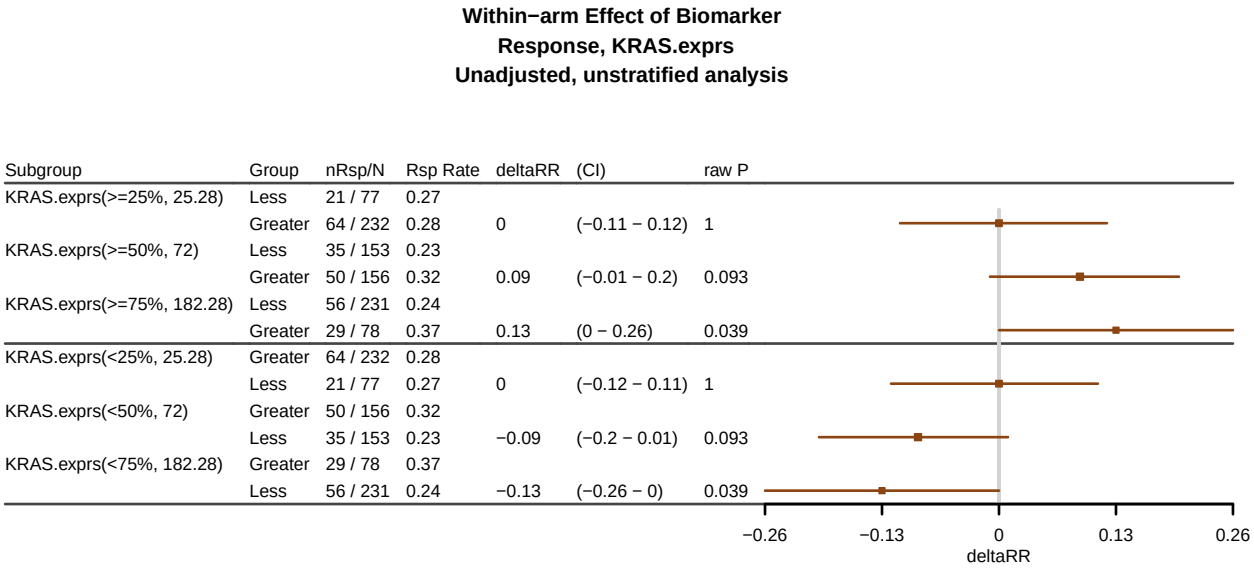
4.1 Try different cutoffs - look for consistent trend

The following figures investigate whether the biomarker shows a within-arm effect (e.g. patients with higher biomarker value tend to have better clinical outcome):

```
res.multicut <- PlotTabForestBiomarker(data=input,
                                       outcome.class=outcome.class,
                                       outcome.var=outcome.var,
                                       trt=trt,
                                       var=bm,
                                       var.class=bm.class,
                                       percentile.cutoff=percentile.trycut,
                                       numerical.cutoff = numerical.trycut,
                                       rsp.response = rsp.response,
                                       rsp.nonresponse = rsp.nonresponse,
                                       greater=TRUE, less=TRUE,
```

```
                                    show.itt=TRUE, show.bep=TRUE,
                                    covariate=covariate, strata=strata)
```

```
## Covariate adjustment and stratification are not supported for binary outcome
## Some NAs in var column, will define the non NA entries as BEP
## only 1 arm; show.itt is set to FALSE
## only 1 arm; show.bep is set to FALSE
```



The forest plots above show within-arm response rate difference (delta) between biomarker high (>= cutoff) vs. low (< cutoff) group. For a given arm, if the delta is not all around 0, it indicates that within this arm the biomarker has an association to the clinical outcome. For example, within treatment arm, suppose all high vs. low delta are greater than 0 and the delta is larger when cutting at a higher value. This indicates that among patients who received treatment, patients who have higher biomarker value tends to have better clinical outcome.

5 Biomarker subgroup analysis (using selected cutoff)

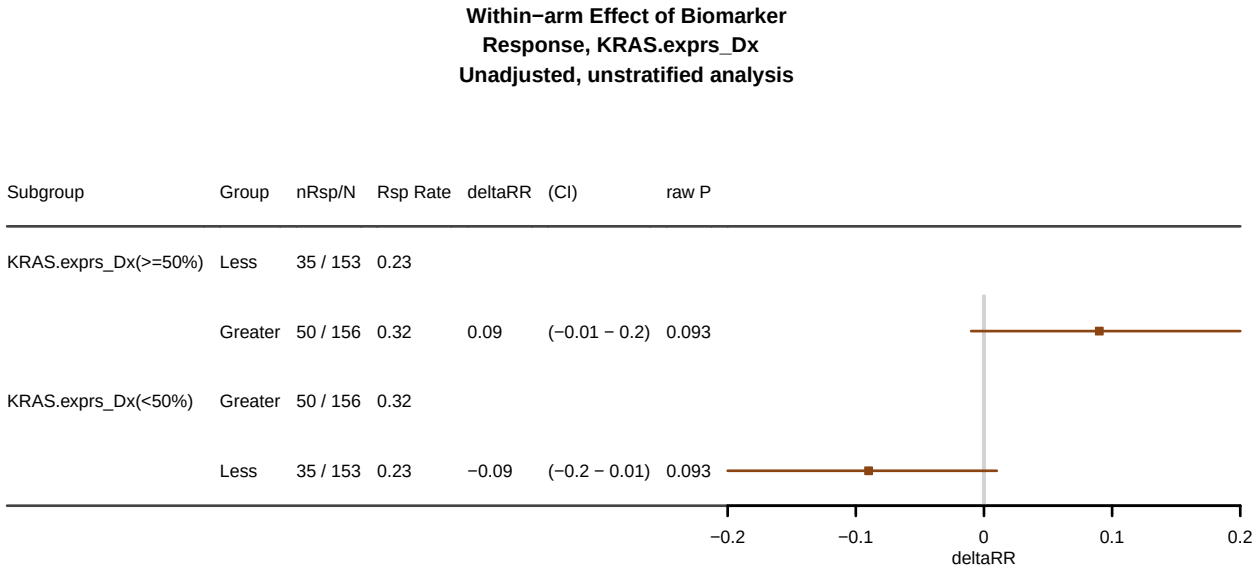
5.1 Estimations within each subgroup

The following figure shows estimate of treatment effect in biomarker subgroups, defined by the selected cutoff.

```
if(bm.class=="numeric"){
if(!is.null(numerical.finalcut)) levs <- paste0(c(">=","<"),numerical.finalcut)
if(is.null(numerical.finalcut)) {
  nm <- quantile(input.bep[[bm]],percentile.finalcut, 2) # default quantile type in forest plot functio
  numerical.finalcut <- round(nm,2) # default rounding decimal in forest plots
  levs <- paste0(c(">=","<"),percentile.finalcut*100,"%")
}

bm2 <- paste0(bm,"_Dx")
input[[bm2]] <- ifelse(input[[bm]] >= numerical.finalcut, levs[1],levs[2])
input[[bm2]]<- factor(input[[bm2]], levels=levs) # ">=" as Dx+
}
```

```
if(bm.class=="categorical") {
  bm2 <- bm
  levs <- unique(input[[bm2]])
}

res.2group <- PlotTabForestBiomarker(data=input,
                                     outcome.class=outcome.class,
                                     outcome.var=outcome.var,
                                     trt=trt,
                                     var=bm2,
                                     var.class="categorical",
                                     greater=TRUE, less=TRUE,
                                     show.itt=TRUE, show.bep=TRUE,
                                     covariate=covariate, strata=strata)
```

```
## Covariate adjustment and stratification are not supported for binary outcome

## Some NAs in var column, will define the non NA entries as BEP

## only 1 arm; show.itt is set to FALSE

## only 1 arm; show.bep is set to FALSE
```

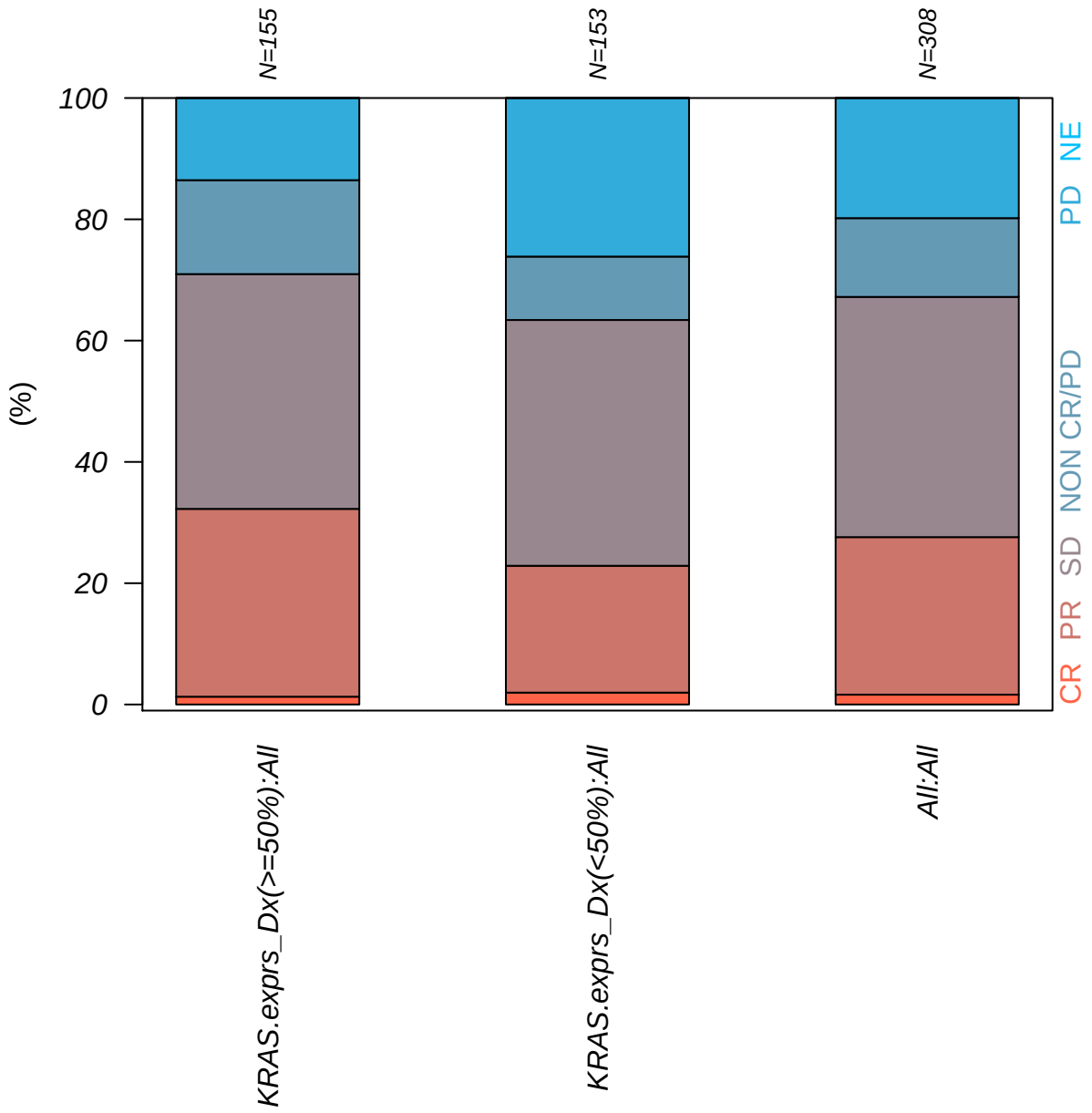


5.2 Subgroup analysis

The following figure show response category distributions of the biomarker subgroups, based on selected cutoff:

```
Rsp.out <- PlotRspBar(data=input, outcome.var=outcome.var, binary=FALSE,
rsp.levels=rsp.levels, trt=trt, bep.indicator=BEP.indicator, var=bm2,
compare.var=TRUE, bep = BEP)
```

```
## entries with missing outcome.var are removed! 365 entries left
```




```
kable(Rsp.out$count,caption="count")
```

Table 6: count

	CR	PR	SD	NON CR/PD	PD	NE
KRAS.exprs_Dx(>=50%):All	2	48	60	24	21	0
KRAS.exprs_Dx(<50%):All	3	32	62	16	40	0
All:All	5	80	122	40	61	0

```
kable(round(Rsp.out$perc,2), caption="percentage")
```

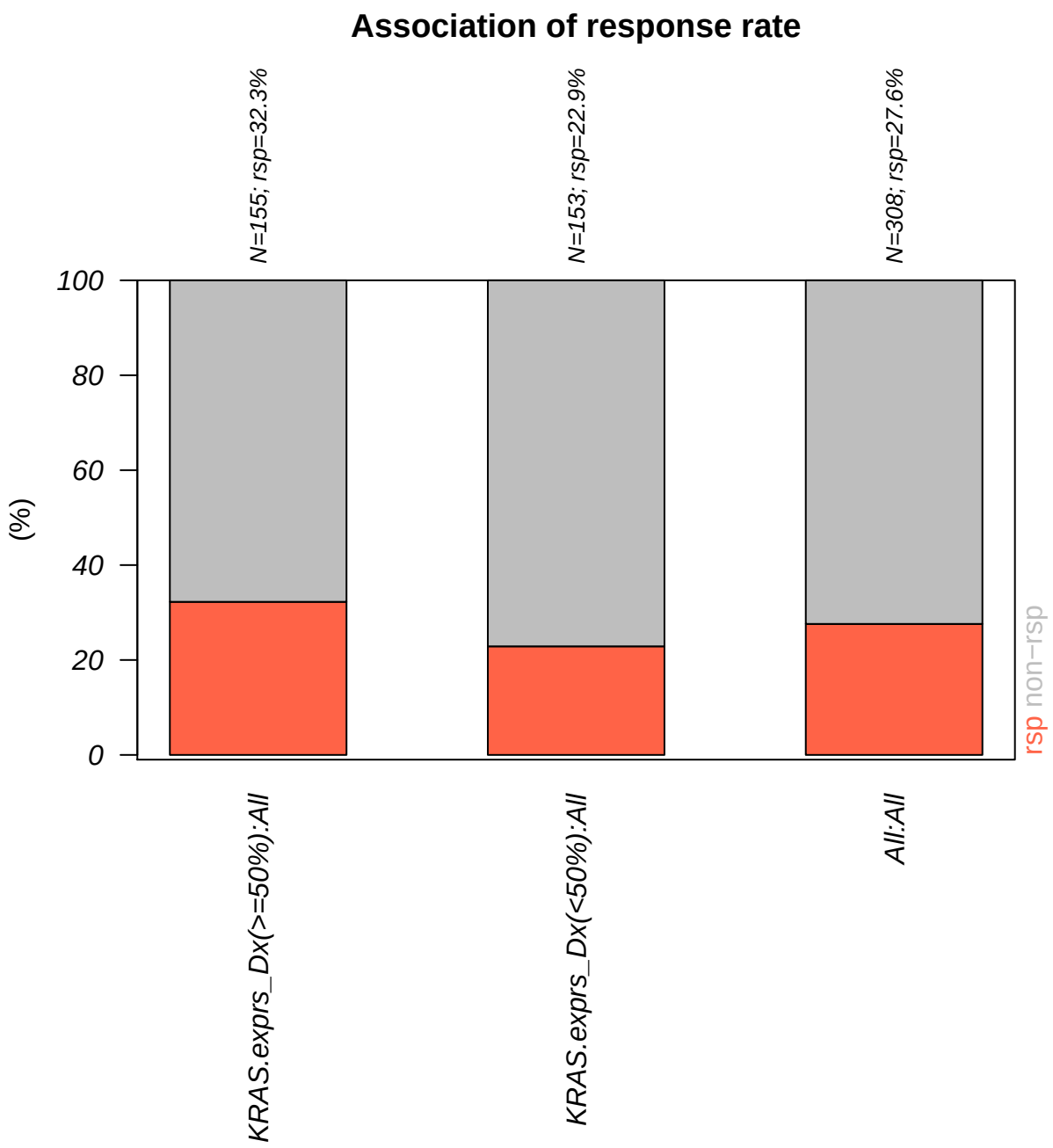
Table 7: percentage

	CR	PR	SD	NON CR/PD	PD	NE
KRAS.exprs_Dx(>=50%):All	0.01	0.31	0.39	0.15	0.14	0
KRAS.exprs_Dx(<50%):All	0.02	0.21	0.41	0.10	0.26	0
All:All	0.02	0.26	0.40	0.13	0.20	0

Summarize the response in two classes - rsp vs. non-rsp:

```
Rsp.out.2 <- PlotRspBar(data=input, outcome.var=outcome.var, binary=TRUE,
rsp.response = rsp.response,
rsp.nonresponse = rsp.nonresponse,rsp.name=rsp.name, trt=trt, bep.indicator=BEP.indicator,
var=bm2, compare.var=TRUE, bep = BEP)
```

```
## entries with missing outcome.var are removed! 365 entries left
```




```
kable(Rsp.out.2$count,caption="count")
```

Table 8: count

	rsp	non-rsp
KRAS.exprs_Dx(>=50%):All	50	105
KRAS.exprs_Dx(<50%):All	35	118
All:All	85	223

```
kable(round(Rsp.out.2$perc,2), caption="percentage")
```

Table 9: percentage

	rsp	non-rsp
KRAS.exprs_Dx(>=50%):All	0.32	0.68
KRAS.exprs_Dx(<50%):All	0.23	0.77
All:All	0.28	0.72

5.3 Check whether biomarker subgroup is confounded with key clinical variables

The following table checks whether the biomarker subgroup is confounded with clinical variables. Here we check whether the clinical variable distribution is comparable in biomarker high and low group.

```
input.bep <- input[which(input[[BEP]]==BEP.indicator),]


kable(
SummaryVars(data=input.bep,trt=trt, subgroup=bm2, var=clinical.vars,
       var.class=clinical.vars.class, subgroup.indicator=levs[1],compare.subgroup=TRUE)
)
```

	KRAS.exprs_Dx_>=50%	KRAS.exprs_Dx_<50%
Sex		
Total	156	153
NA's	0	0
F	72 (46.15%)	75 (49.02%)
M	84 (53.85%)	78 (50.98%)
Age		
N	156	153
Mean	53.34	55.25
Median	53	55
Min-Max	33...89	30...82
NA's	0	0