Biomarker analysis report

2017-08-29

Contents

1 The dataset **1**

2 Representativeness: Selection Bias of Biomarker Population **1**
2.1 Check selection bias in terms of key clinical variables, between full population and BEP . . . 2
2.2 Check whether the clinical outcome in BEP is comparable to the full population 2
2.3 Examine whether the prognostic/predictive/null trend of key clinical variables holds in BEP . 3

3 Biomarker property and its association to clinical variables **6**
3.1 Biomarker property and relationship to clinical variable 6
3.2 Whether the biomarker shows within-arm effect 7

4 Biomarker cutoff exploration/selection **9**
4.1 Try different cutoffs - look for consistent trend 9
4.2 Estimations within non-overlapping bins 10
4.3 Estimations within overlapped sliding windows - STEPP plot 11

5 Biomarker subgroup analysis (using selected cutoff) **12**
5.1 Estimations within each subgroup 12
5.2 Subgroup analysis 13
5.3 Check whether biomarker subgroup is confounded with key clinical variables 14

1 The dataset

The dataset have 550 entries. In which 461 are in biomarker evaluable population (BEP).

- Endpoint of interest: Lab_ontrt
- 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

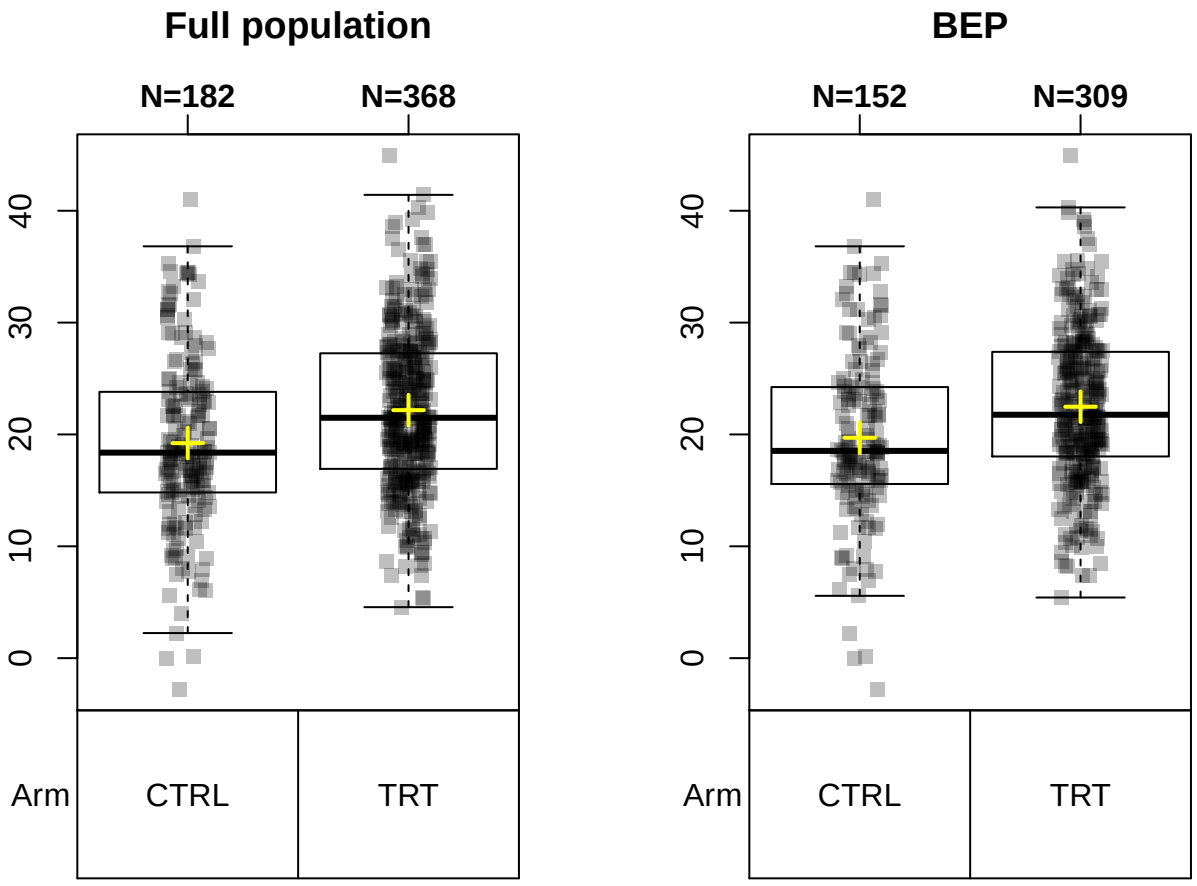
```
kable(SummaryVars(data=input, trt=trt, subgroup=BEP, subgroup.indicator = BEP.indicator,
                  var=clinical.vars, var.class=clinical.vars.class))
```

	All(CTRL)	BEP(CTRL)	All(TRT)	BEP(TRT)
Sex				
Total	182	152	368	309
NA's	0	0	0	0
F	89 (48.9%)	73 (48.03%)	184 (50%)	147 (47.57%)
M	93 (51.1%)	79 (51.97%)	184 (50%)	162 (52.43%)
Age				
N	182	152	368	309
Mean	52.54	52.35	54.03	54.29
Median	51.5	51	54	54
Min-Max	27...85	27...85	30...89	30...89
NA's	0	0	0	0

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

The following plot compares continuous 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.

```
par(mfrow=c(1,2))
BoxPlot(obj=input, form=as.formula(paste(outcome.var,"~",trt)),
        XaxisTab=list(font=2, col="darkblue", cex=1.25),
        Title=list(main="Full population"), mar=c(5,5,5,1))
BoxPlot(obj=input.bep, form=as.formula(paste(outcome.var,"~",trt)),
        XaxisTab=list(font=2, col="darkblue", cex=1.25),
        Title=list(main="BEP"), mar=c(5,5,5,1))
```



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

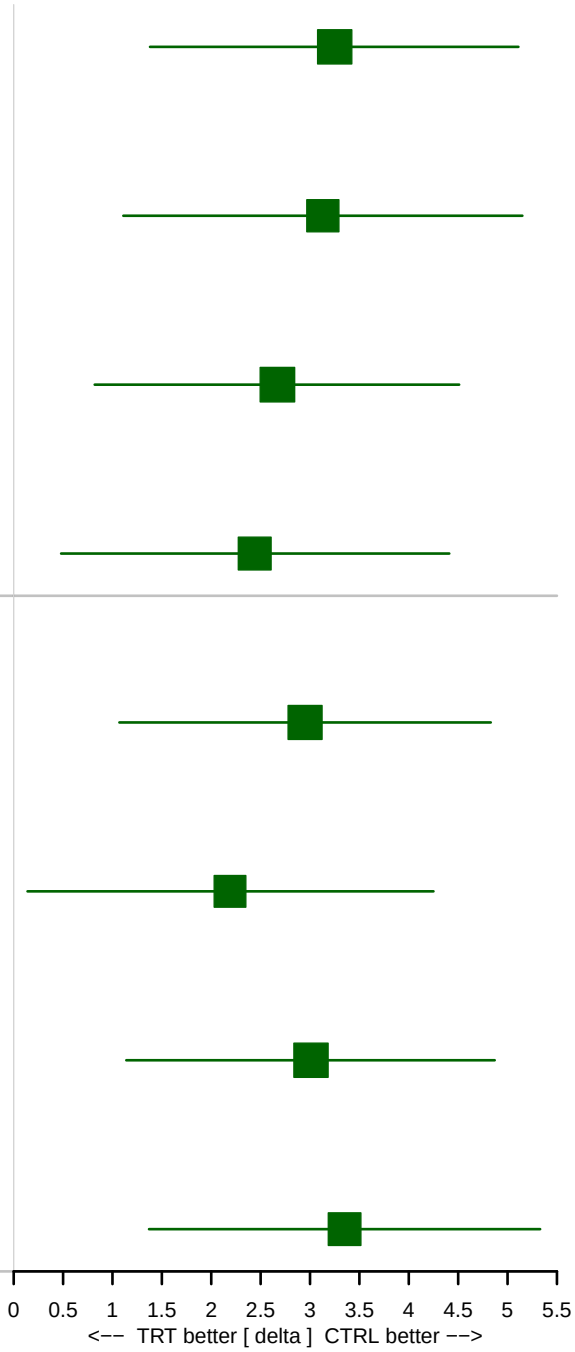
The following forest plot can be used to examine whether any of the key prognostic/predictive clinical variables still show prognostic/predictive trend in BEP:

```
forest.bep <- PlotTabForestMulti(data=input,
                  outcome.class=outcome.class,
                  outcome.var=outcome.var,
                  trt=trt,
                  var=clinical.vars,
                  var.class=clinical.vars.class,
                  bep=BEP,bep.indicator=BEP.indicator,
                  compare.bep.itt=TRUE
                   )
```

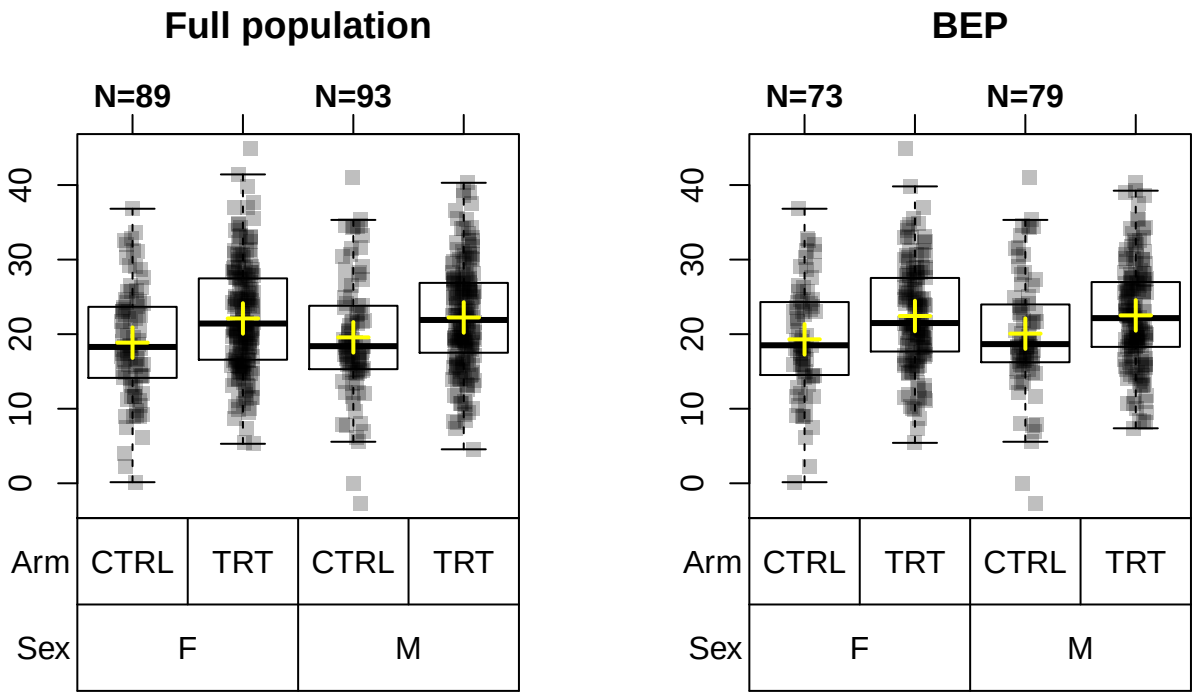
```
## Stratification is not supported for continuous outcome
## Stratification is not supported for continuous outcome
## Stratification is not supported for continuous outcome
## Stratification is not supported for continuous outcome
```

Across arm, Compare BEP vs. All
Lab_ontrt
Unadjusted, unstratified analysis

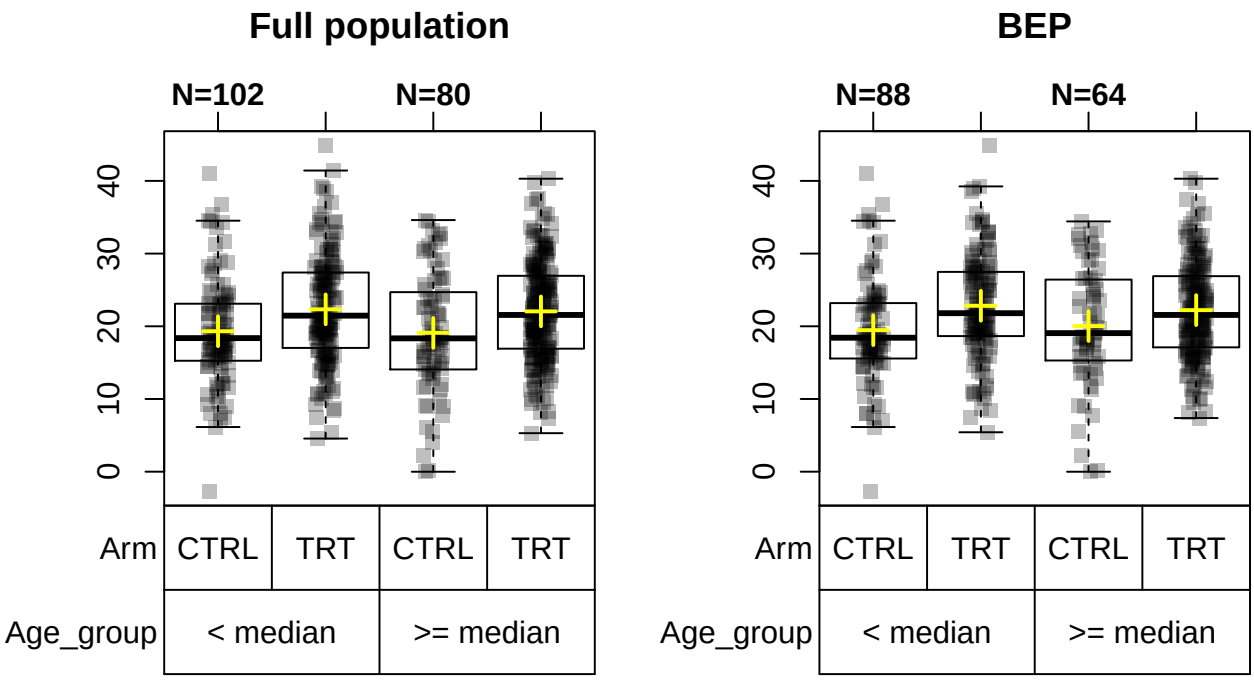
Subgroup	Group	Mean	delta	(CI)	raw P
ITT Sex(F)	CTRL	18.86			
	TRT	22.11	3.25	(1.38 − 5.11)	0.00073
BEP Sex(F)	CTRL	19.31			
	TRT	22.43	3.13	(1.11 − 5.15)	0.0025
ITT Sex(M)	CTRL	19.57			
	TRT	22.24	2.67	(0.82 − 4.51)	0.0048
BEP Sex(M)	CTRL	20.07			
	TRT	22.52	2.44	(0.48 − 4.41)	0.015
ITT Age(>=50%, 53)	CTRL	19.1			
	TRT	22.05	2.95	(1.07 − 4.83)	0.0022
BEP Age(>=50%, 53)	CTRL	20.03			
	TRT	22.22	2.19	(0.14 − 4.25)	0.036
ITT Age(<50%, 53)	CTRL	19.32			
	TRT	22.33	3.01	(1.14 − 4.87)	0.0017
BEP Age(<50%, 53)	CTRL	19.47			
	TRT	22.82	3.35	(1.37 − 5.33)	0.00099



```
for(i in clinical.vars.2){
par(mfrow=c(1,2))
BoxPlot(obj=input, form=as.formula(paste(outcome.var,"~",trt,":",i)),
        XaxisTab=list(font=2, col="darkblue", cex=1.25),
        Title=list(main="Full population"), mar=c(5,5,5,1))
BoxPlot(obj=input.bep, form=as.formula(paste(outcome.var,"~",trt,":",i)),
        XaxisTab=list(font=2, col="darkblue", cex=1.25),
        Title=list(main="BEP"), mar=c(5,5,5,1))
print("")
}
```




```
## [1] ""
```




```
## [1] ""
```

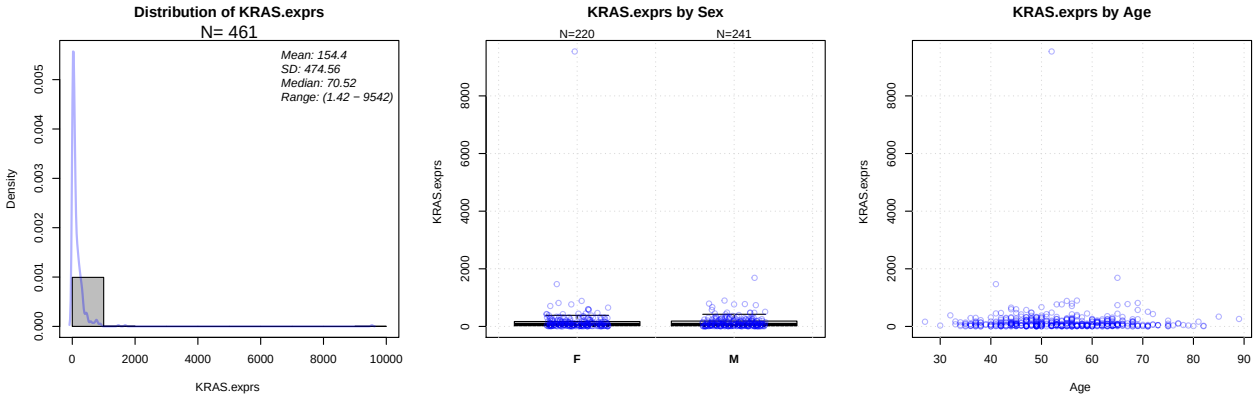
3 Biomarker property and its association to clinical variables

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)))
```



3.2 Whether the biomarker shows within-arm effect

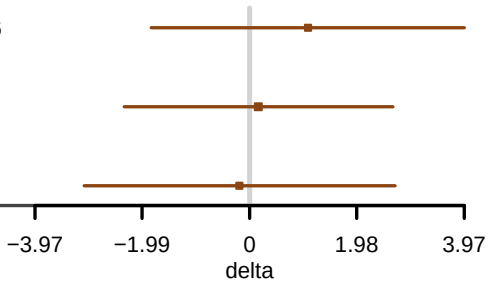
The following figures investigate whether the biomarker shows a within-arm effect (e.g. patients with higher biomarker value tend to have better clinical outcome):

```
input.ctrl <- subset(input, Arm==placebo.code) ## Data with only ctrl samples
res.multicut.ctrl <- PlotTabForestBiomarker(data=input.ctrl,
                                    outcome.class=outcome.class,
                                    outcome.var=outcome.var,
                                    var=bm,
                                    var.class=bm.class,
                                    percentile.cutoff=percentile.trycut,
                                    numerical.cutoff = numerical.trycut,
                                    rsp.response = rsp.response,
                                    rsp.nonresponse = rsp.nonresponse,
                                    main.prefix=placebo.code,
                                    greater=TRUE, less=FALSE,
                                    covariate=covariate, strata=strata)
```

```
## Stratification is not supported for continuous 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
```

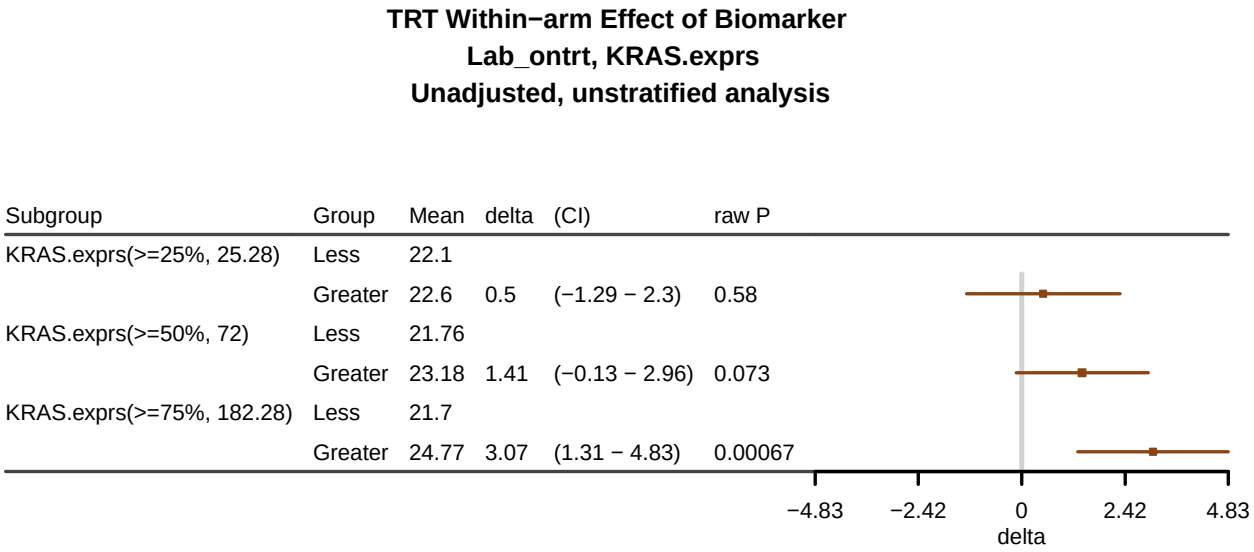
CTRL Within−arm Effect of Biomarker
Lab_ontrt, KRAS.exprs
Unadjusted, unstratified analysis

Subgroup	Group	Mean	delta	(CI)	raw P
KRAS.exprs(>=25%, 22.63)	Less	18.89			
	Greater	19.97	1.08	(−1.82 − 3.97)	0.46
KRAS.exprs(>=50%, 64)	Less	19.62			
	Greater	19.79	0.16	(−2.32 − 2.65)	0.9
KRAS.exprs(>=75%, 171.86)	Less	19.75			
	Greater	19.56	−0.19	(−3.06 − 2.69)	0.9




```
input.trt <- subset(input, Arm==active.code) ## Data with only ctrl samples
res.multicut.trt <- PlotTabForestBiomarker(data=input.trt,
                                    outcome.class=outcome.class,
                                    outcome.var=outcome.var,
                                    var=bm,
                                    var.class=bm.class,
                                    percentile.cutoff=percentile.trycut,
                                    numerical.cutoff = numerical.trycut,
                                    rsp.response = rsp.response,
                                    rsp.nonresponse = rsp.nonresponse,
                                    main.prefix=active.code,
                                    greater=TRUE, less=FALSE,
                                    covariate=covariate, strata=strata)
```

```
## Stratification is not supported for continuous 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
## Warning in par(old.par): calling par(new=TRUE) with no plot
```



The forest plots above show within-arm mean difference (delta) comparing 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 greater clinical outcome.

If similar trend is seen in both arms, it indicates that the biomarker may have a prognostic effect (the biomarker is able to identify patients with better/worse clinical outcome, regardless of treatment).

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

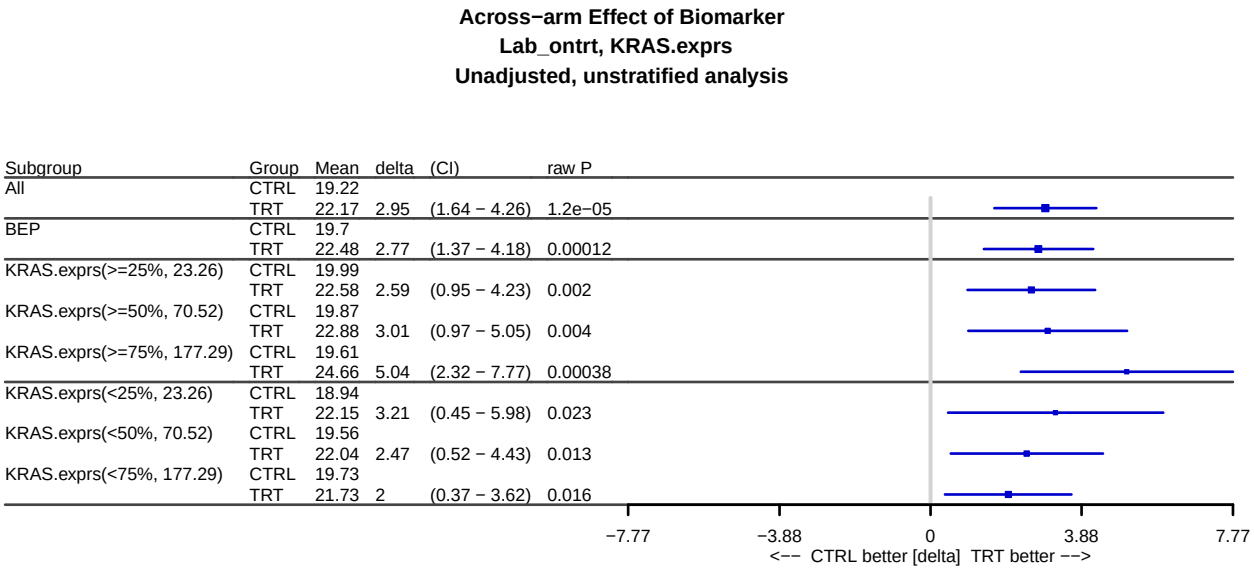
The following plots investigate whether the biomarker is predictive across arm. To perform cross-arm analysis. TRT-CTRL mean difference within biomarker high or low group are calculated. The high/low groups are defined by trying different cutoffs.

```
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)
```

```
## Stratification is not supported for continuous outcome
```

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



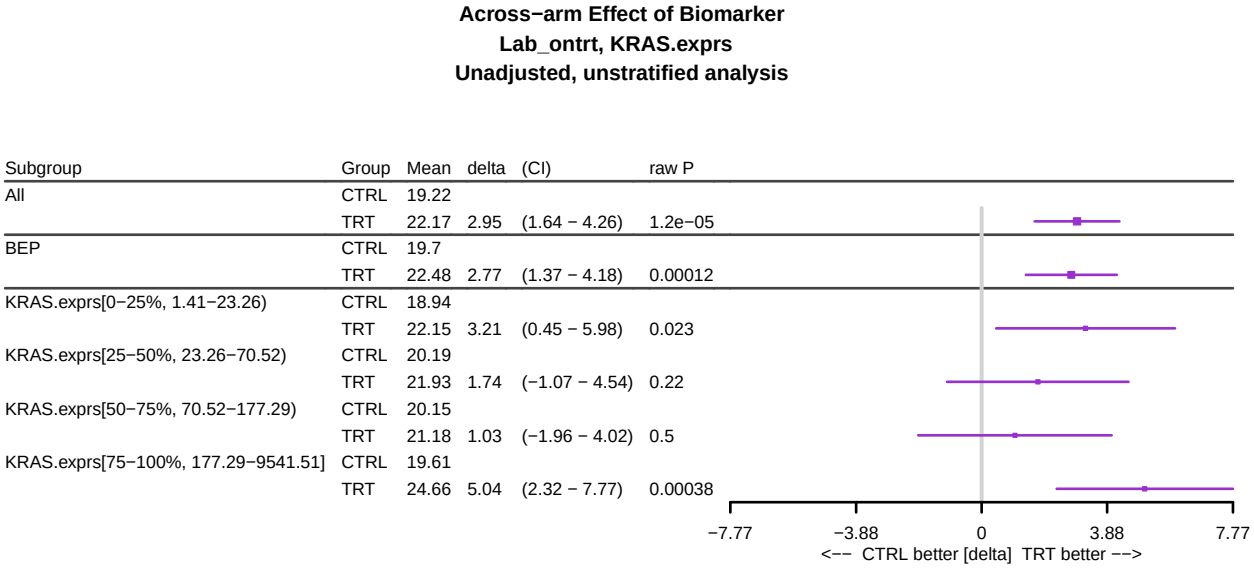
4.2 Estimations within non-overlapping bins

The following figure show TRT-CTRL mean difference within non-overlapping bins defined by those exploratory cutoffs.

```
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,
                                    greater=FALSE, less=FALSE, within.bin=TRUE,
                                    show.itt=TRUE, show.bep=TRUE,
                                    covariate=covariate, strata=strata)
```

```
## Stratification is not supported for continuous outcome
```

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



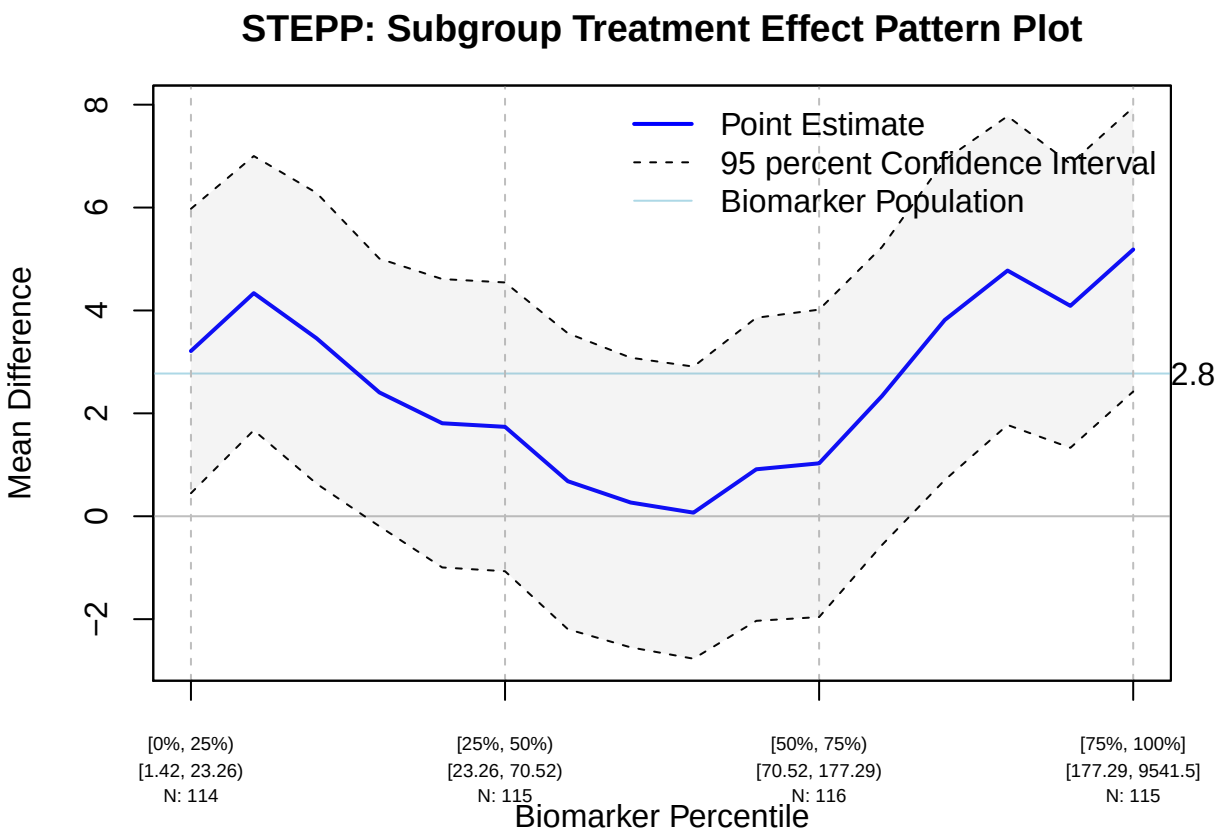
4.3 Estimations within overlapped sliding windows - STEPP plot

STEPP refers to Subgroup Treatment Effect Pattern Plot and it investigates relationship between biomarker and treatment effect. Only continuous biomarkers are suitable for STEPP analysis. STEPP performs treatment effect estimation on overlapping subsets of patients defined according to the biomarker level. The default setting of run.STEPP slides subgroup windows by 5% for each step and the subgroup size is 25% of the whole population.

A monotone trend is expected to be seen for an ideal biomarker.

```
stepp.out <- PlotSTEPP(data = input,
                                  outcome.class=outcome.class,
                                  outcome.var=outcome.var,
                                  trt=trt,
                                  var=bm,
                                  placebo.code = placebo.code,
                                  active.code = active.code
)
```

```
## some NA in var column, will ignore NA entries
```



5 Biomarker subgroup analysis (using selected cutoff)

5.1 Estimations within each subgroup

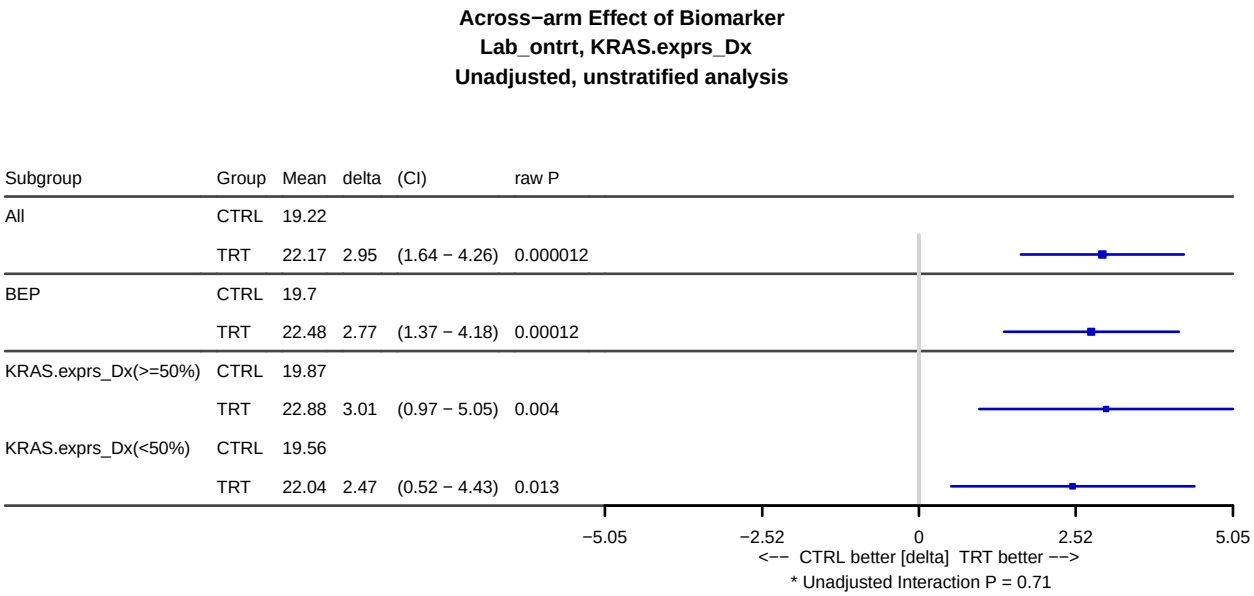
The following figure shows estimate of treatment effect in biomarker subgroups, defined by the selected cutoff.

```
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)
```

```
## Stratification is not supported for continuous outcome
```

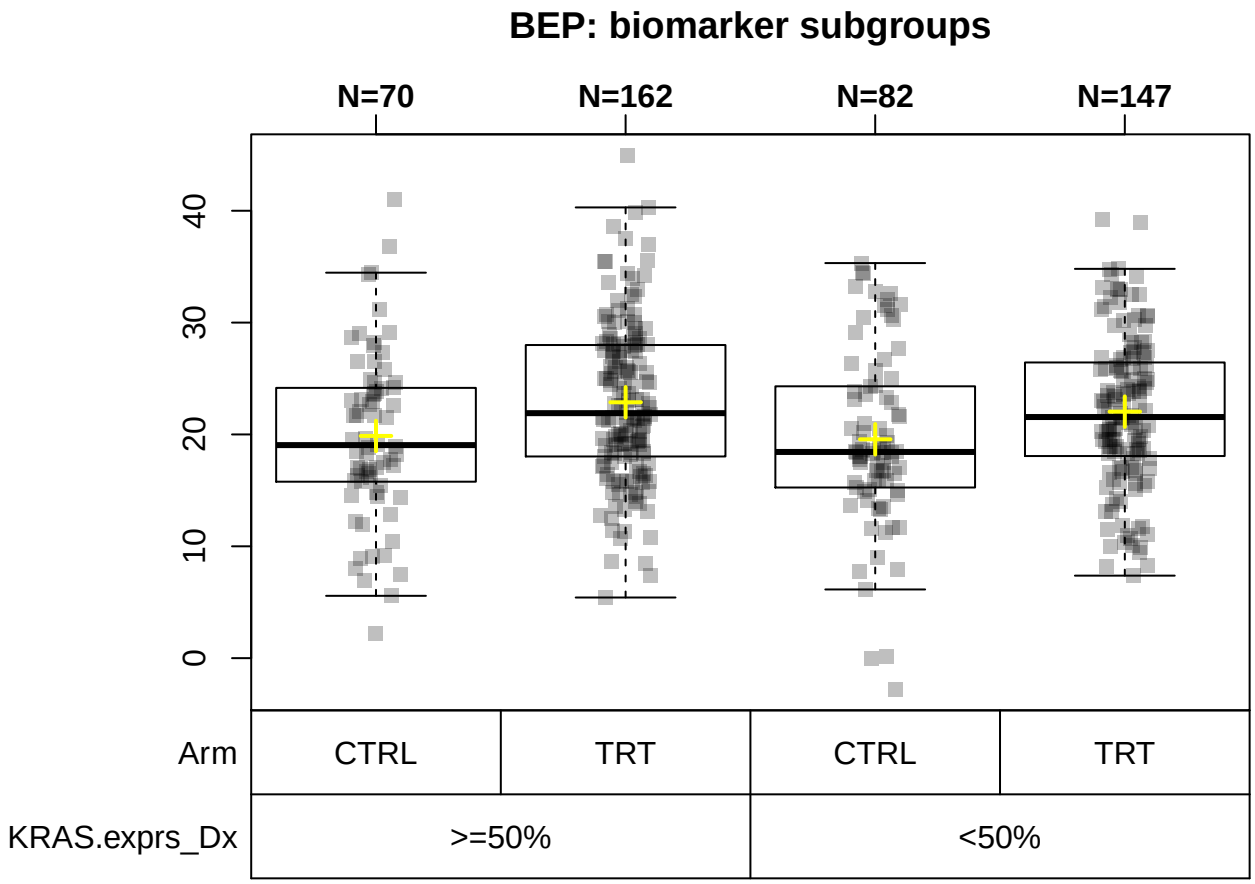
```
## Some NAs in var column, will define the non NA entries as BEP
```



5.2 Subgroup analysis

The following figure show distribution of the continuous endpoint within the biomarker subgroups, based on selected cutoff:

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


BoxPlot(obj=input.bep, form=as.formula(paste(outcome.var,"~",trt,":",bm2)),
        XaxisTab=list(font=2, col="darkblue", cex=1.25),
        Title=list(main="BEP: biomarker subgroups"),
        mar=c(5,8,5,1))
```



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.

```
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%(CTRL)	KRAS.exprs_Dx_<50%(CTRL)	KRAS.exprs_Dx_>=50%(TRT)	KRAS.exprs_Dx_<50%(TRT)
Sex				
Total	70	82	162	147
NA's	0	0	0	0
F	35 (50%)	38 (46.34%)	77 (47.53%)	70 (47.62%)
M	35 (50%)	44 (53.66%)	85 (52.47%)	77 (52.38%)
Age				
N	70	82	162	147
Mean	51.5	53.07	53.36	55.31
Median	50	52	53	55
Min-Max	27...85	35...82	33...89	30...82
NA's	0	0	0	0