

iCARE(Individualized Coherent Absolute Risk Estimators) Package

April 24, 2017

```
> library(iCARE)
```

Example 1.A

Load the breast cancer data.

```
> data("bc_data", package="iCARE")
```

In this example, we will estimate the risk of breast cancer in ages 50-80. A SNP-only model is fit, with no specific genotypes supplied for estimation. The population disease rates are from SEER.

```
> res_snps_miss = computeAbsoluteRisk(model.snp.info = bc_15_snps,
+                                     model.disease.incidence.rates = bc_inc,
+                                     model.competing.incidence.rates = mort_inc,
+                                     apply.age.start = 50,
+                                     apply.age.interval.length = 30,
+                                     return.refs.risk=TRUE)
```

Note: You did not provide apply.snp.profile. Will impute SNPs for 10000 people.

If require more, please provide apply.snp.profile input.

```
[1] "Note: As specified, the model does not adjust SNP imputations for family history."
      user system elapsed
10.45    0.22    10.67
```

Compute a summary of the risks and visualize the results

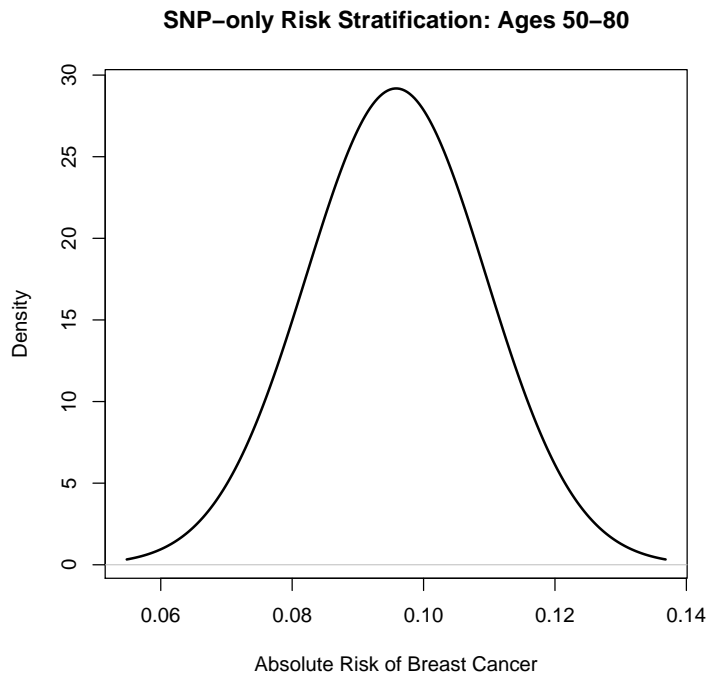
```
> summary(res_snps_miss$risk)
```

```
Risk_Estimate
Min.      :0.09583
1st Qu.:0.09583
Median :0.09583
Mean     :0.09583
3rd Qu.:0.09583
Max.     :0.09583
```

```
> summary(res_snps_miss$refs.risk)
```

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.07476	0.09190	0.09574	0.09583	0.09957	0.11773

```
> plot(density(res_snps_miss$risk), lwd=2,
+      main="SNP-only Risk Stratification: Ages 50-80",
+      xlab="Absolute Risk of Breast Cancer")
```



Example 1.B

In this example, we will again estimate the risk of breast cancer in ages 50-80. This time however, three specific genotypes are supplied for estimation (with some missing data). The argument `return.refs.risk = TRUE`, includes the referent dataset risks be included in results.

```
> res_snps_dat = computeAbsoluteRisk(model.snp.info = bc_15_snps,
+                                   model.disease.incidence.rates = bc_inc,
+                                   model.competing.incidence.rates = mort_inc,
+                                   apply.age.start = 50,
+                                   apply.age.interval.length = 30,
+                                   apply.snp.profile = new_snp_prof,
+                                   return.refs.risk = TRUE)
```

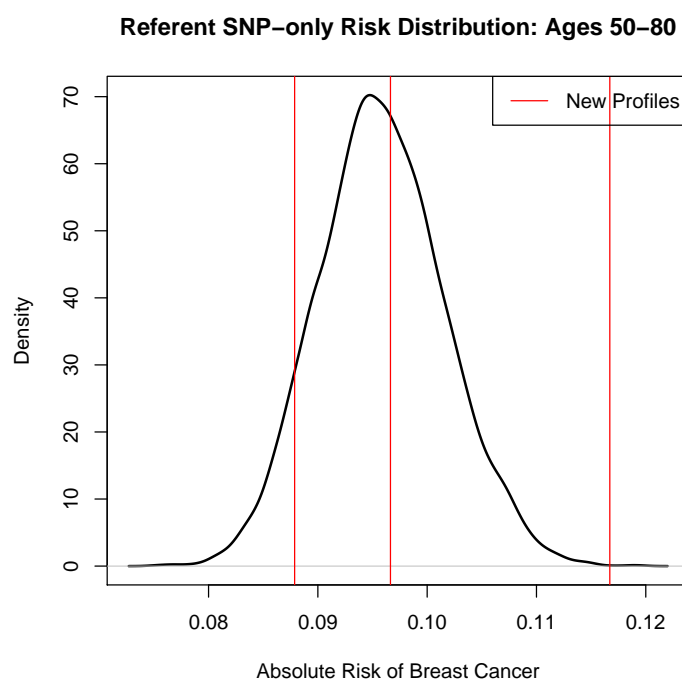
```
[1] "Note: As specified, the model does not adjust SNP imputations for family history."
      user system elapsed
      0.40    0.06    0.47
```

```
> names(res_snps_dat)
```

```
[1] "risk"      "details"   "beta.used" "refs.risk"
```

Visualize the Results

```
> plot(density(res_snps_dat$refs.risk), lwd=2,
+      main="Referent SNP-only Risk Distribution: Ages 50-80",
+      xlab="Absolute Risk of Breast Cancer")
> abline(v=res_snps_dat$risk, col="red")
> legend("topright", legend="New Profiles", col="red", lwd=1)
```



Example 2

In this example, we will estimate the risk of breast cancer in ages 50-80 by fitting a model with two risk factors and 15 SNPs, with three specific covariate profiles supplied for estimation (with some missing data). First, the `model.cov.info` argument is created.

```
> v1 <- list(name="famhist", type="continuous")
> v2 <- list(name="parity", type="factor", levels=0:4, ref=0)
> bc_model_cov_info <- list(v1, v2)
```

Now fit the model

```
> res_covs_snps = computeAbsoluteRisk(model.formula=caco~famhist+as.factor(parity),
+                                     model.cov.info=bc_model_cov_info,
+                                     model.snp.info=bc_15_snps,
+                                     model.log.RR=bc_model_log_or,
```

```

+                                     model.ref.dataset=ref_cov_dat,
+                                     model.disease.incidence.rates=bc_inc,
+                                     model.competing.incidence.rates=mort_inc,
+                                     model.bin.fh.name="famhist",
+                                     apply.age.start=50,
+                                     apply.age.interval.length=30,
+                                     apply.cov.profile=new_cov_prof,
+                                     apply.snp.profile=new_snp_prof,
+                                     return.refs.risk=TRUE)

```

```

user  system elapsed
0.06   0.00   0.06

```

Display details of the fit

```
> print(res_covs_snps$details)
```

	Int_Start	Int_End	Risk_Estimate	rs12405132	rs12048493	rs72755295		
193502	50	80	0.08045200	NA	NA	0		
126252	50	80	0.07714544	0	0	0		
15756	50	80	0.12962868	0	1	0		
	rs6796502	rs13162653	rs2012709	rs7707921	rs9257408	rs4593472	rs13365225	
193502	0	1	1	0	1	1	1	
126252	0	2	1	0	1	1	0	
15756	1	0	1	0	1	1	0	
	rs13267382	rs11627032	chr17:29230520:D	rs745570	rs6507583	famhist	parity	
193502	0	0		1	2	0	0	2
126252	1	1		0	1	0	0	4
15756	1	0		0	0	0	1	4

Session Information

```
> sessionInfo()
```

```

R version 3.4.0 (2017-04-21)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows Server 2012 R2 x64 (build 9600)

```

Matrix products: default

locale:

```

[1] LC_COLLATE=C
[2] LC_CTYPE=English_United States.1252
[3] LC_MONETARY=English_United States.1252
[4] LC_NUMERIC=C
[5] LC_TIME=English_United States.1252

```

attached base packages:

```
[1] stats      graphics  grDevices  utils      datasets  methods   base
```

other attached packages:

[1] iCARE_1.4.0

loaded via a namespace (and not attached):

[1] compiler_3.4.0 tools_3.4.0