

iCARE(Individualized Coherent Absolute Risk Estimators) Package

May 1, 2024

```
> 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_72_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
11.201  0.088  11.290
```

Compute a summary of the risks and visualize the results

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

```
Risk_Estimate
Min.   :0.096
1st Qu.:0.096
Median :0.096
Mean   :0.096
3rd Qu.:0.096
Max.   :0.096
```

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

```

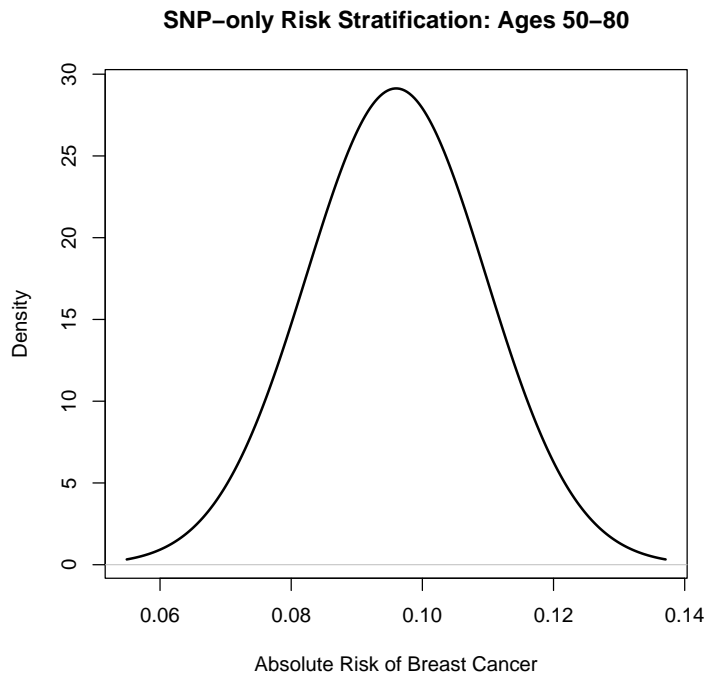
      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.05900 0.08660 0.09491 0.09600 0.10437 0.16712

```

```

> 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_72_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
0.357  0.000  0.356

```

```

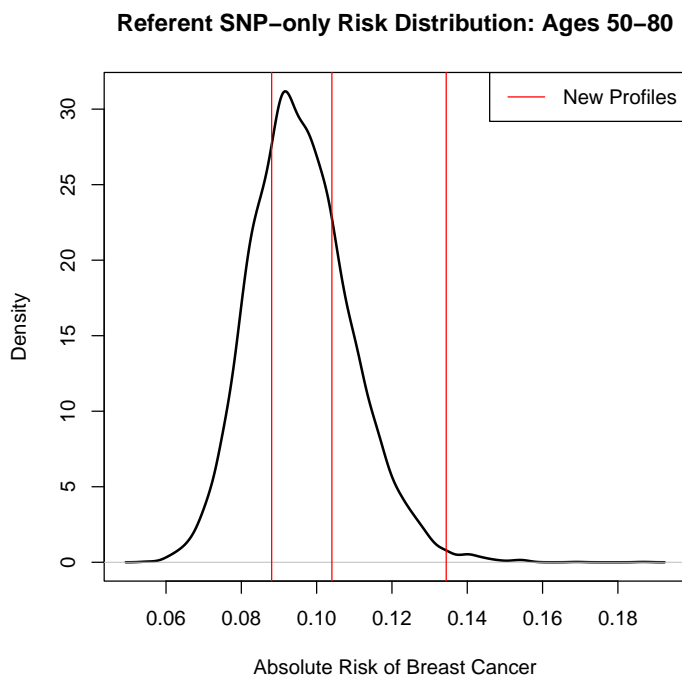
> names(res_snps_dat)

```

```
[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 13 risk factors and 72 SNPs.

```
> res_covs_snps = computeAbsoluteRisk(model.formula=bc_model_formula,  
+                                   model.cov.info=bc_model_cov_info,  
+                                   model.snp.info=bc_72_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  
1.017  0.288  1.305
```

Display details of the fit

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

	Int_Start	Int_End	Risk_Estimate	rs616488	rs11552449	rs11249433	rs12405132	
1	50	80	0.10246926	NA	NA	NA	NA	
2	50	80	0.09008679	2	0	NA	NA	
3	50	80	0.16886889	2	0	1	1	1
	rs12048493	rs6678914	rs4245739	rs72755295	rs12710696	rs4849887	rs2016394	
1	NA	0	0	0	0	0	0	
2	NA	NA	NA	NA	1	1	0	
3	1	1	1	0	2	0	0	
	rs1550623	rs16857609	rs6762644	rs4973768	rs12493607	rs6796502	rs9790517	
1	0	0	0	1	1	0	1	
2	0	2	1	1	1	1	2	
3	0	0	0	2	1	0	1	
	rs6828523	rs10069690	rs13162653	rs2012709	rs10941679	rs10472076	rs1353747	
1	0	1	2	0	0	2	0	
2	0	0	1	0	0	1	1	
3	0	0	1	0	0	0	1	
	rs7707921	rs1432679	rs11242675	rs204247	rs9257408	rs4593472	rs720475	
1	0	1	2	0	0	1	1	
2	0	0	1	2	1	1	0	
3	1	2	1	2	1	1	0	
	rs9693444	rs13365225	rs6472903	rs2943559	rs13267382	rs11780156	rs1011970	
1	1	1	1	0	0	0	0	
2	0	0	1	0	2	1	1	
3	1	1	0	0	1	0	0	
	rs10759243	rs2380205	rs7072776	rs11814448	rs7904519	rs11199914	rs554219	
1	0	2	2	0	0	1	1	
2	1	0	0	0	0	0	0	
3	1	1	1	0	2	0	1	
	rs75915166	rs11820646	rs12422552	rs17356907	rs1292011	rs11571833	rs2236007	
1	0	1	1	0	1	0	1	
2	0	0	0	0	0	0	0	
3	0	1	1	0	2	0	0	
	rs2588809	rs999737	rs941764	rs11627032	rs17817449	rs11075995	rs13329835	
1	0	0	1	0	1	1	1	
2	1	0	0	1	1	1	0	
3	0	0	1	0	0	1	1	
	rs146699004	rs745570	rs527616	rs1436904	rs6507583	rs4808801	rs3760982	
1	0	0	0	0	0	1	0	
2	1	2	0	0	0	1	1	
3	1	2	1	1	0	1	1	
	rs2284378	rs2823093	rs17879961	rs132390	rs6001930	famhist	menarche_dec	parity

1	1	1	0	0	0	0	8	0
2	1	0	0	0	0	0	10	0
3	0	0	0	0	0	0	1	0
	birth_dec	agemeno_dec	height_dec	bmi_dec	rd_menohrt	rd2_everhrt_e		
1	2	2	6	10	1	0		
2	2	1	6	4	1	0		
3	1	7	1	10	1	0		
	rd2_everhrt_c	rd2_currhrt	alcoholweek_dec	ever_smoke				
1	0	0		1	1			
2	0	0		6	0			
3	0	0		1	1			

Session Information

```
> sessionInfo()
```

R version 4.4.0 RC (2024-04-16 r86468)

Platform: x86_64-pc-linux-gnu

Running under: Ubuntu 22.04.4 LTS

Matrix products: default

BLAS: /home/biocbuild/bbs-3.20-bioc/R/lib/libRblas.so

LAPACK: /usr/lib/x86_64-linux-gnu/lapack/liblapack.so.3.10.0

locale:

```
[1] LC_CTYPE=en_US.UTF-8      LC_NUMERIC=C
[3] LC_TIME=en_GB             LC_COLLATE=C
[5] LC_MONETARY=en_US.UTF-8   LC_MESSAGES=en_US.UTF-8
[7] LC_PAPER=en_US.UTF-8     LC_NAME=C
[9] LC_ADDRESS=C              LC_TELEPHONE=C
[11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
```

time zone: America/New_York

tzcode source: system (glibc)

attached base packages:

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

other attached packages:

```
[1] iCARE_1.33.0 Hmisc_5.1-2 gtools_3.9.5 plotrix_3.8-4
```

loaded via a namespace (and not attached):

```
[1] gtable_0.3.5      dplyr_1.1.4      compiler_4.4.0   rpart_4.1.23
[5] tidymodels_1.2.1 htmlTable_2.4.2  stringr_1.5.1   gridExtra_2.3
[9] cluster_2.1.6     scales_1.3.0     fastmap_1.1.1   ggplot2_3.5.1
[13] R6_2.5.1          generics_0.1.3   Formula_1.2-5   knitr_1.46
[17] htmlwidgets_1.6.4 backports_1.4.1  checkmate_2.3.1 tibble_3.2.1
[21] munsell_0.5.1     nnet_7.3-19     pillar_1.9.0    rlang_1.1.3
[25] utf8_1.2.4        stringi_1.8.3    xfun_0.43       cli_3.6.2
```

```
[29] magrittr_2.0.3    digest_0.6.35    grid_4.4.0      rstudioapi_0.16.0
[33] base64enc_0.1-3   lifecycle_1.0.4  vctrs_0.6.5     data.table_1.15.4
[37] evaluate_0.23     glue_1.7.0       fansi_1.0.6     colorspace_2.1-0
[41] rmarkdown_2.26    foreign_0.8-86   tools_4.4.0     pkgconfig_2.0.3
[45] htmltools_0.5.8.1
```