

iCARE(Individualized Coherent Absolute Risk Estimators) Package

April 26, 2022

```
> 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  
10.53    0.30    10.83
```

Compute a summary of the risks and visualize the results

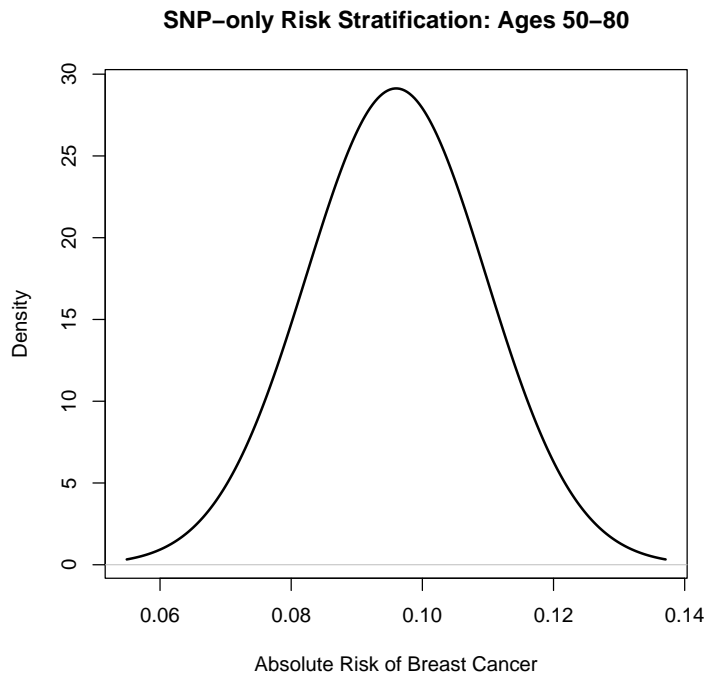
```
> summary(res_snps_miss$risk)
```

```
Risk_Estimate  
Min.      :0.09601  
1st Qu.:0.09601  
Median :0.09601  
Mean     :0.09601  
3rd Qu.:0.09601  
Max.     :0.09601
```

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

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.05457	0.08659	0.09498	0.09601	0.10438	0.15672

```
> 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)
```

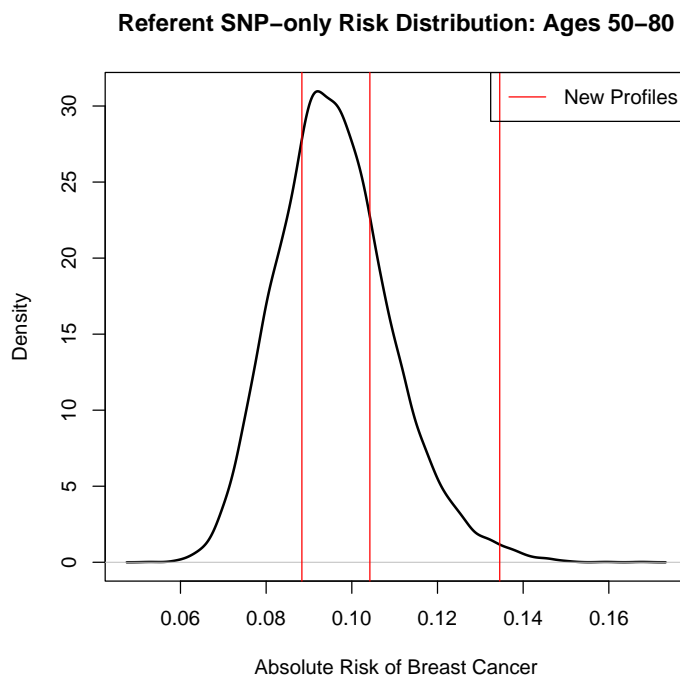
```
[1] "Note: As specified, the model does not adjust SNP imputations for family history."
      user system elapsed
0.64    0.16    0.80
```

```
> 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,
```


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	alcoholdecweek_dec	ever_smoke
1	0	0	1	1
2	0	0	6	0
3	0	0	1	1

Session Information

```
> sessionInfo()
```

```
R version 4.2.0 RC (2022-04-19 r82224 ucrt)
Platform: x86_64-w64-mingw32/x64 (64-bit)
Running under: Windows Server x64 (build 20348)
```

```
Matrix products: default
```

```
locale:
```

```
[1] LC_COLLATE=C
[2] LC_CTYPE=English_United States.utf8
[3] LC_MONETARY=English_United States.utf8
[4] LC_NUMERIC=C
[5] LC_TIME=English_United States.utf8
```

```
attached base packages:
```

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

```
other attached packages:
```

```
[1] iCARE_1.24.0   Hmisc_4.7-0    ggplot2_3.3.5   Formula_1.2-4
[5] survival_3.3-1 lattice_0.20-45 gtools_3.9.2    plotrix_3.8-2
```

```
loaded via a namespace (and not attached):
```

```
[1] tidyselect_1.1.2    xfun_0.30        purrr_0.3.4
[4] splines_4.2.0       colorspace_2.0-3 vctrs_0.4.1
[7] generics_0.1.2      htmltools_0.5.2  base64enc_0.1-3
[10] utf8_1.2.2          rlang_1.0.2      pillar_1.7.0
[13] foreign_0.8-82      glue_1.6.2       withr_2.5.0
[16] DBI_1.1.2           RColorBrewer_1.1-3 jpeg_0.1-9
[19] lifecycle_1.0.1     stringr_1.4.0    munsell_0.5.0
[22] gtable_0.3.0        htmlwidgets_1.5.4 latticeExtra_0.6-29
[25] knitr_1.38          fastmap_1.1.0    fansi_1.0.3
[28] htmlTable_2.4.0     scales_1.2.0     backports_1.4.1
[31] checkmate_2.1.0     gridExtra_2.3    png_0.1-7
[34] digest_0.6.29       stringi_1.7.6    dplyr_1.0.8
```

[37]	grid_4.2.0	cli_3.3.0	tools_4.2.0
[40]	magrittr_2.0.3	tibble_3.1.6	cluster_2.1.3
[43]	crayon_1.5.1	pkgconfig_2.0.3	ellipsis_0.3.2
[46]	Matrix_1.4-1	data.table_1.14.2	assertthat_0.2.1
[49]	rstudioapi_0.13	R6_2.5.1	rpart_4.1.16
[52]	nnet_7.3-17	compiler_4.2.0	