

compareGroups package, updated and improved

Héctor Sanz^{2,1*}, Isaac Subirana^{3,1,4}, Joan Vila^{1,3}

1. Cardiovascular Epidemiology & Genetics group, Inflammatory and Cardiovascular Disease Programme, IMIM, Hospital del Mar Research Institute, Spain

2. UCICEC CAIBER. IMIM-Hospital del Mar

3. CIBER Epidemiology and Public Health (CIBERESP), Spain

4. Statistics Department, University of Barcelona, Spain

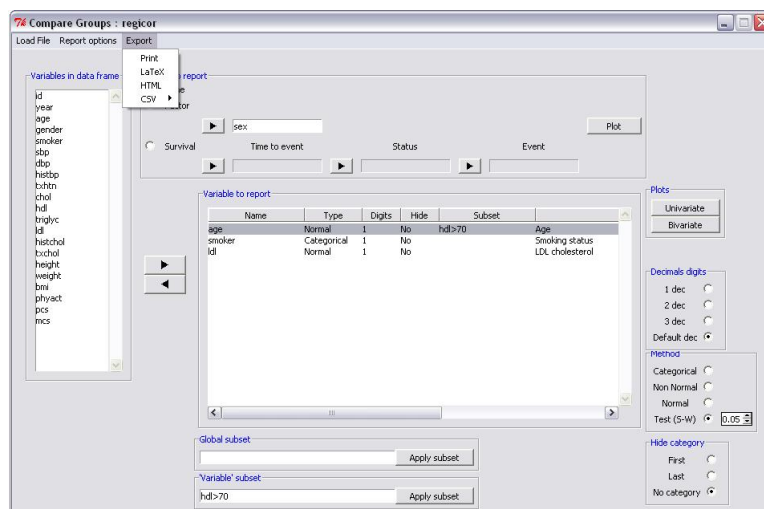
*Contact author: hsanz@imim.es

Keywords: Software Design, Bivariate Table, \LaTeX , Descriptive Analysis

In many studies, such as epidemiological ones, it is needed to compare characteristics between groups of individuals or disease status. Usually these comparisons are presented in the form of tables (also called Bivariate Tables) of descriptive statistics where rows are characteristics, and each column is a group / status. Usually the number of characteristics is large, and thus construction of these tables is laborious.

To build them in an easy, quick and efficient way, we created the **compareGroups** package [1]. Here we present package improvements and extensions in the following issues:

- descriptives by status in a cohort study allowing to include right-censored time-to-response
- descriptives for groups or the entire sample.
- exporting tables to HTML format
- new GUI aspect with a single frame and a main menu
- importing data from workspace using the GUI
- subset specifically for each variable using the GUI
- incidence for right-censored time-to-event row-variables.
- extended and improved vignette
- exporting tables to \LaTeX under the `longtable` environment
- proper plots for survival analysis



Var	Male N=1101	Female N=1193	p.overall
Age	54.8 (11.1)	54.7 (11.0)	0.840
Smoking status:			<0.001
Never smoker	301 (28.1%)	900 (77.5%)	
Current or former < 1y	410 (38.3%)	183 (15.7%)	
Never or former >= 1y	360 (33.6%)	79 (6.80%)	
Systolic blood pressure	134 (18.9)	129 (21.2)	<0.001
Diastolic blood pressure	81.7 (10.2)	77.8 (10.5)	<0.001
History of hypertension	341 (31.1%)	382 (32.1%)	0.644
HTN treatment	189 (17.5%)	239 (20.4%)	0.096
Total cholesterol	217 (42.7)	220 (47.4)	0.140
HDL cholesterol	47.5 (12.6)	57.5 (15.0)	<0.001
Triglycerides	131 (87.4)	101 (55.2)	<0.001
LDL cholesterol	145 (38.5)	142 (40.7)	0.092
Hystory of hypercol	353 (32.3%)	356 (30.2%)	0.308

References

1. Hector Sanz, Isaac Subirana, Joan Vila (2010) "Bivariate Analyses" UserR! 2010, The R User Conference 2010, (National Institute of Standards and Technology, Gaithersburg, Maryland, US), July 2010.