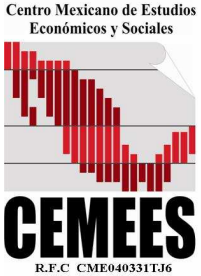
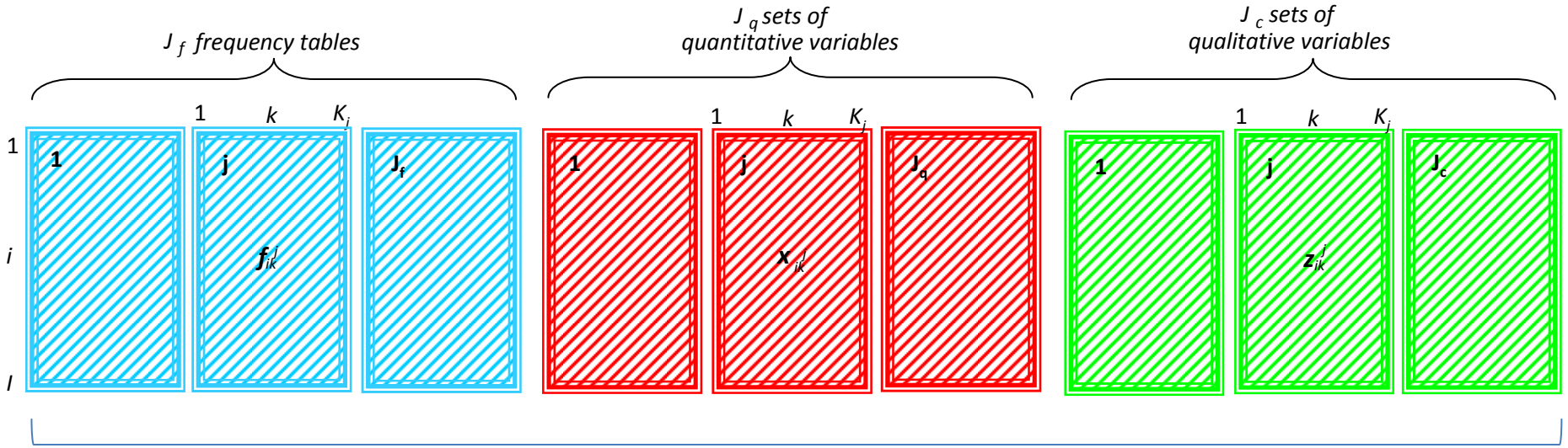


Multiple Factor Analysis for Contingency Tables in FactoMineR Package

Belchin Adriyanov Kostov [a,b], **Mónica Bécue Bertaut** [b], **François Husson** [c], **Daría Micaela Hernández** [d]

[a] Primary Healthcare Center Les Corts - Barcelona
[b] Universitat Politècnica de Catalunya - Barcelona
[c] Agrocampus Rennes - Rennes
[d] Centro Mexicano de Estudios Económicos y Sociales - México D.F.

Multiple tables with a mixture of quantitative, categorical and frequency sets



Analysis

Multiple Factor Analysis for Contingency tables (MFACT)

Multiple Factor Analysis for Contingency Tables (MFACT)

- Proposed by Bécue Bertaut & Pagès (2008)
- Based on MFA (Escofier & Pagès, 1990) principles
  - Balance the influence of set of variables
  - Principal component approach (**X, D, M**)
  - A global non-standardized PCA
  - Global representation of rows and columns
  - Superimposed representation of the rows

MFA function (FactoMineR)

Arguments

- *base:* data frame
- *group:* number of variables in each group
- *type:* type of variables in each group
- row weights, column weights, supplementary individuals and supplementary groups

Values

- *separate.analyses*: results for separate analyses
- *groups:* results for the groups
- results for individuals (*ind*) and variables (*quanti.var, quali.var, freq*)

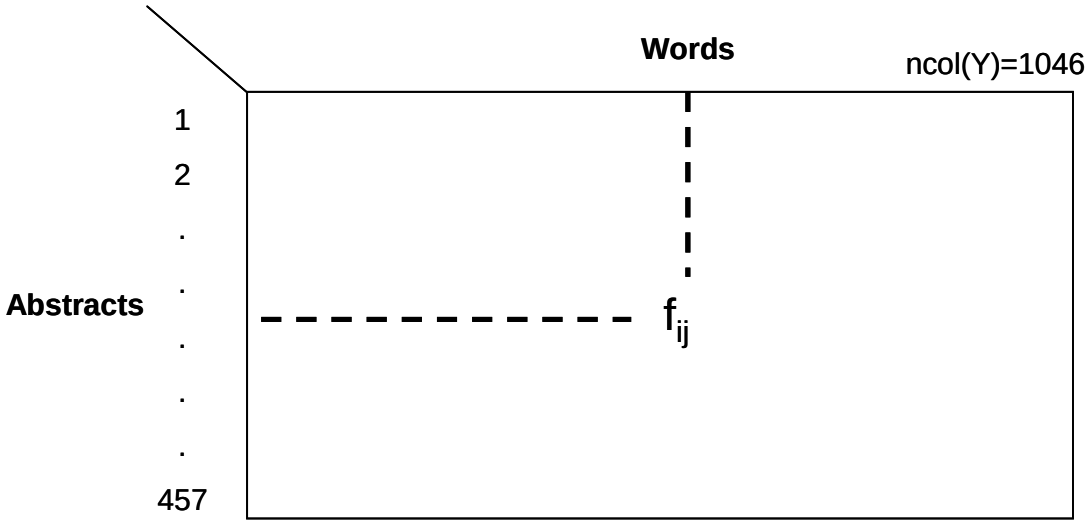
Application of MFACT to a scientometric study in medicine - I

- **Subject** *Systemic Lupus Erythematosus (SLE)*
- **Aims**
  - Study the evolution in the research concerning this rare disease, through the vocabulary changes
  - Look for those abstracts evidencing an important difference between points of view that are the vocabulary of the abstract and its publication year to detect either pioneering works or works returning to topics treated in the past.

Application of MFACT to a scientometric study in medicine - II

- **Methods**
  - Abstracts relative to randomized clinical trials SLE from 1994 to 2011 (457 abstracts)
  - Downloaded from PubMED
  - Abstract, year, journal

Data structure



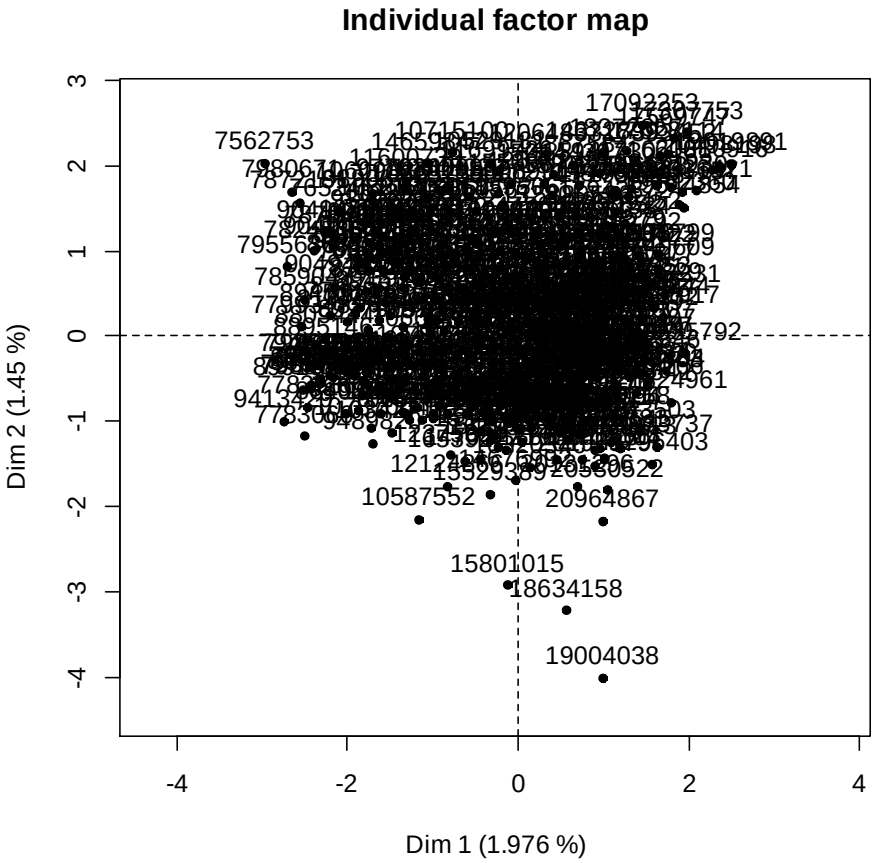
Year	Journal
1994	LUPUS
1994	ARD
.	.
.	.
.	.
.	..
.	
2011	JRHEUM

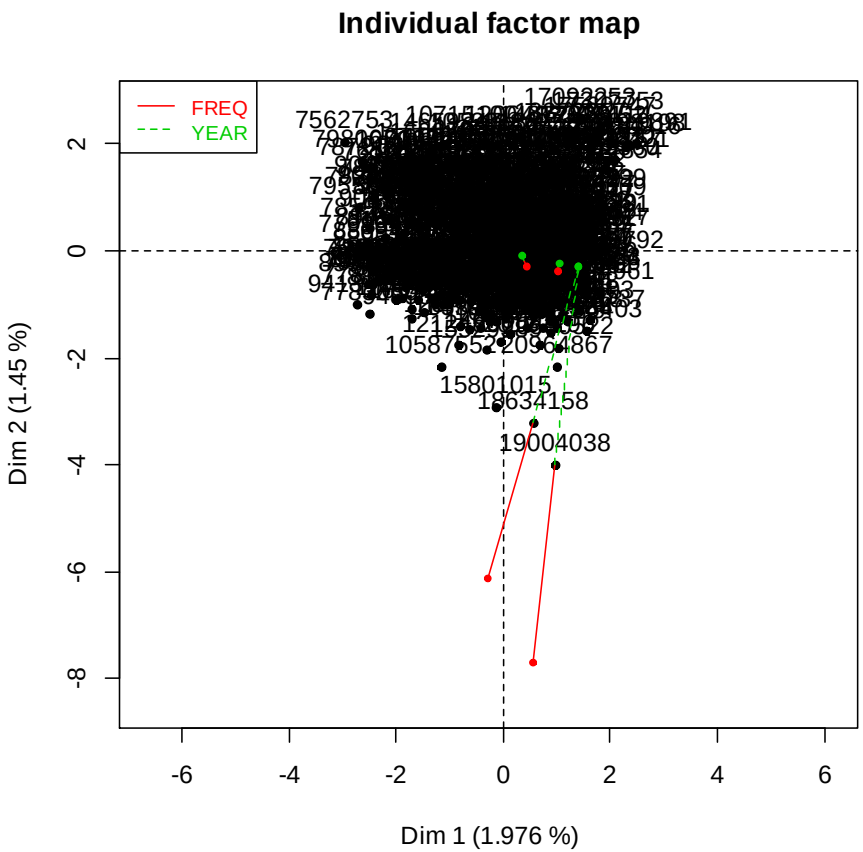
f_{ij}: Frequency of word *j* for abstract *i*

- Abstracts words matrix (457x1046) – *Frequency table*
- Publication year (457x1) – *Quantitative set*
- Journal (457x1) – *Qualitative set*

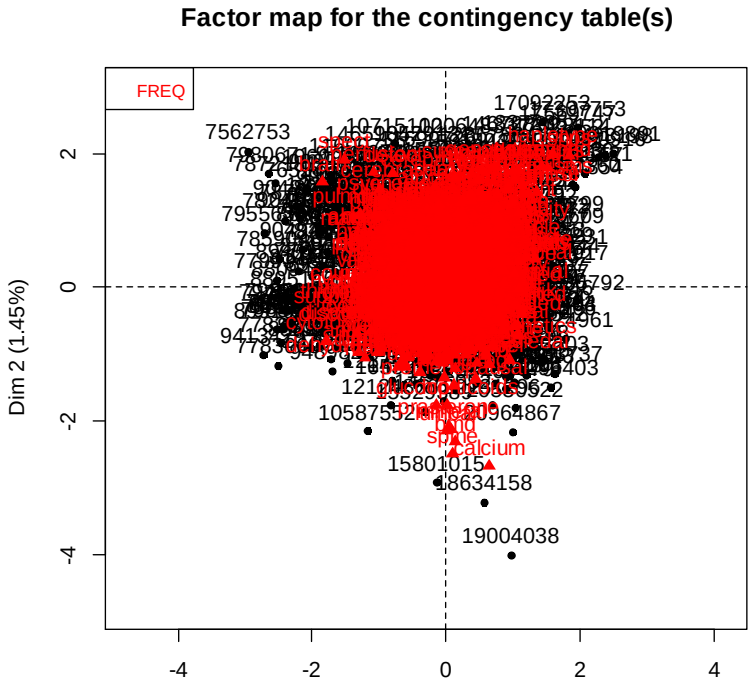
```
mfatc.res<-MFA(base=Data,group=c(1046,1,1),type=c("f","s","n"),
num.group.sup=3, name.group=c("FREQ",“YEAR",“JOURNAL"))
```

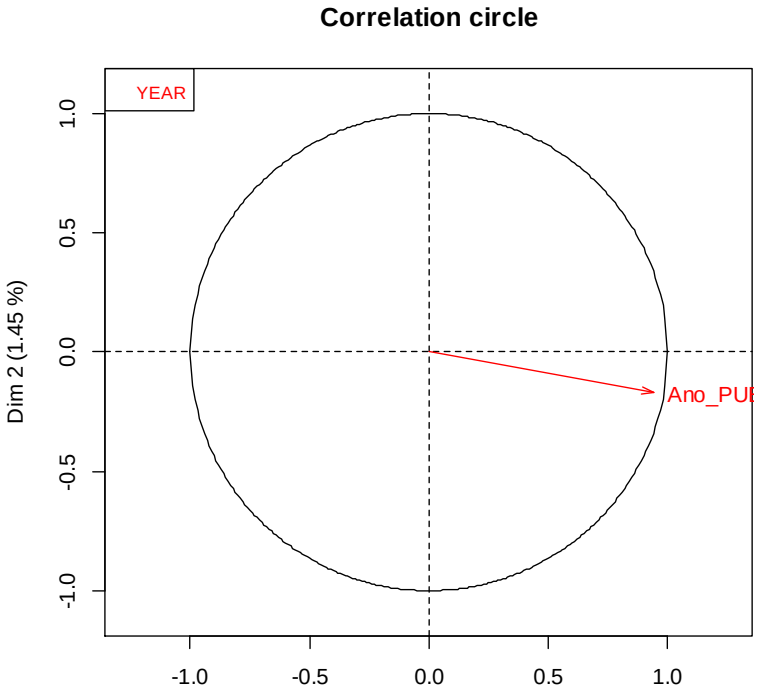
Plots MFACT - Individuals

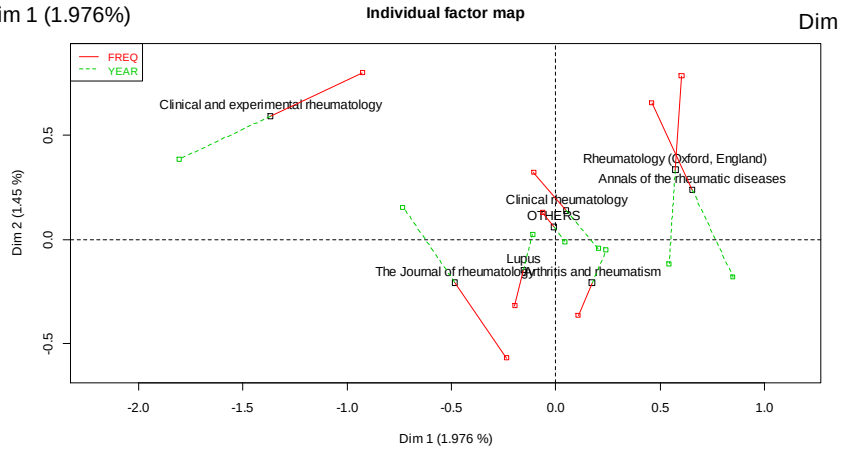




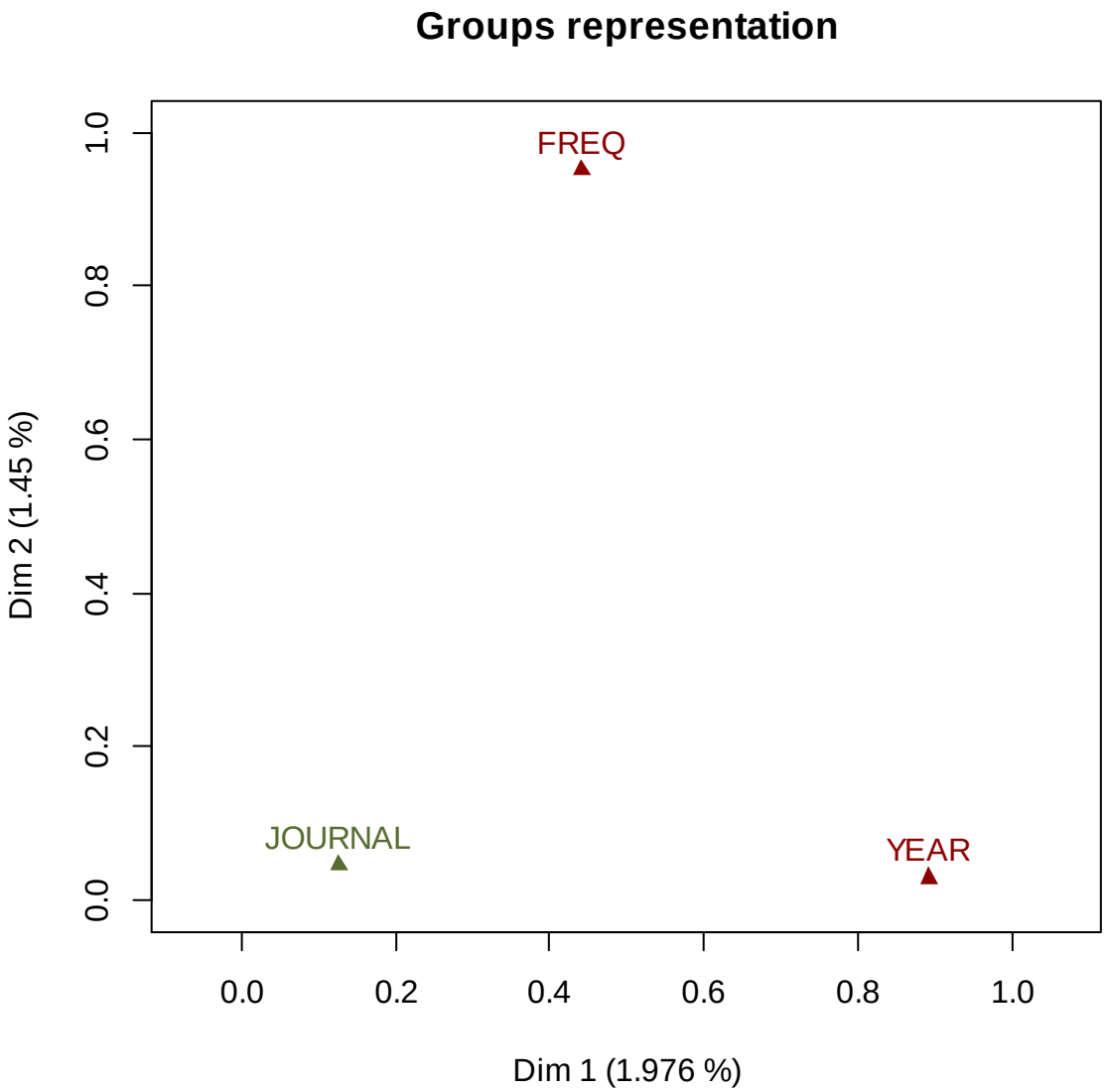
Plots MFACT - Variables



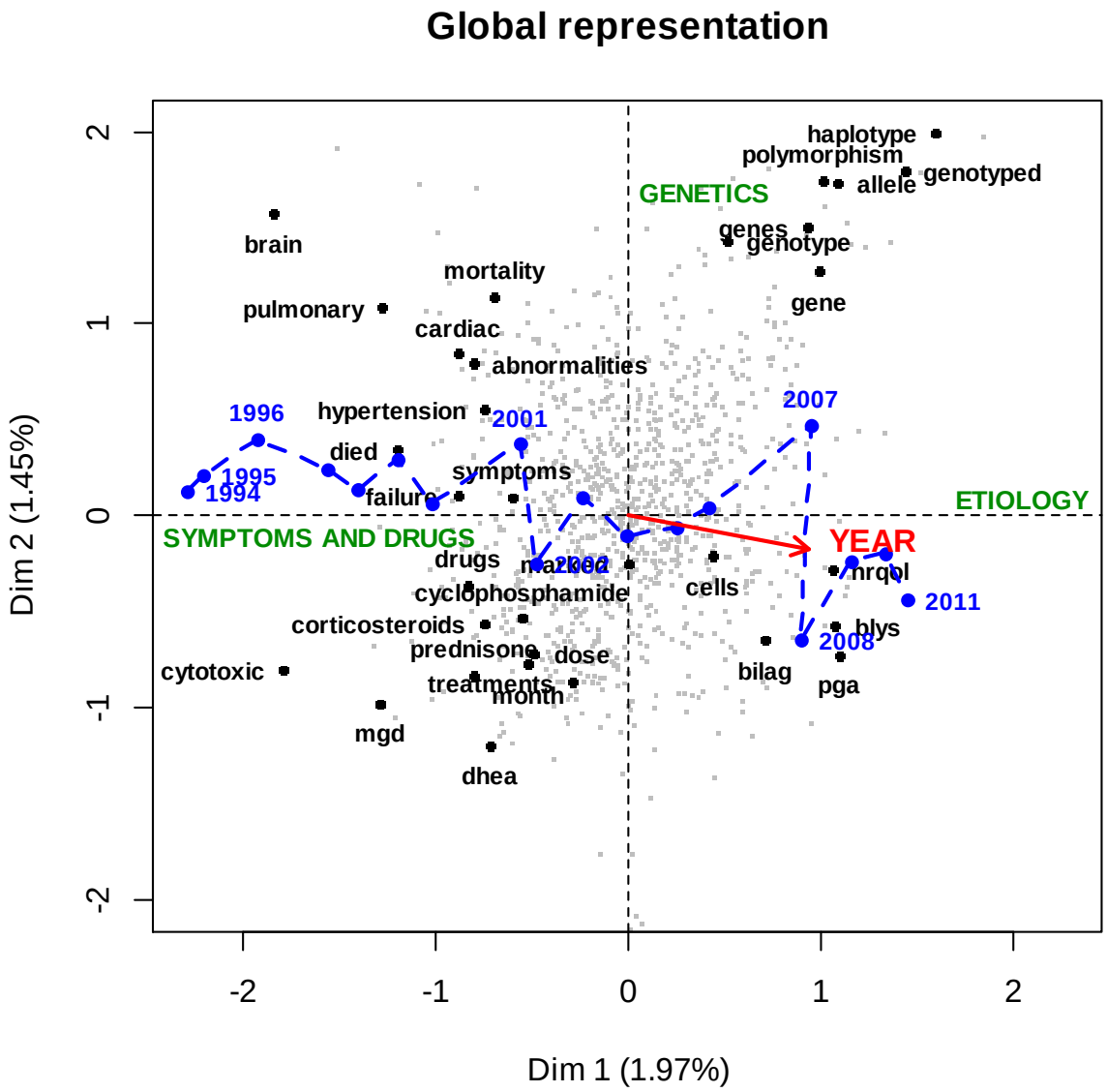




Plots MFACT - Groups

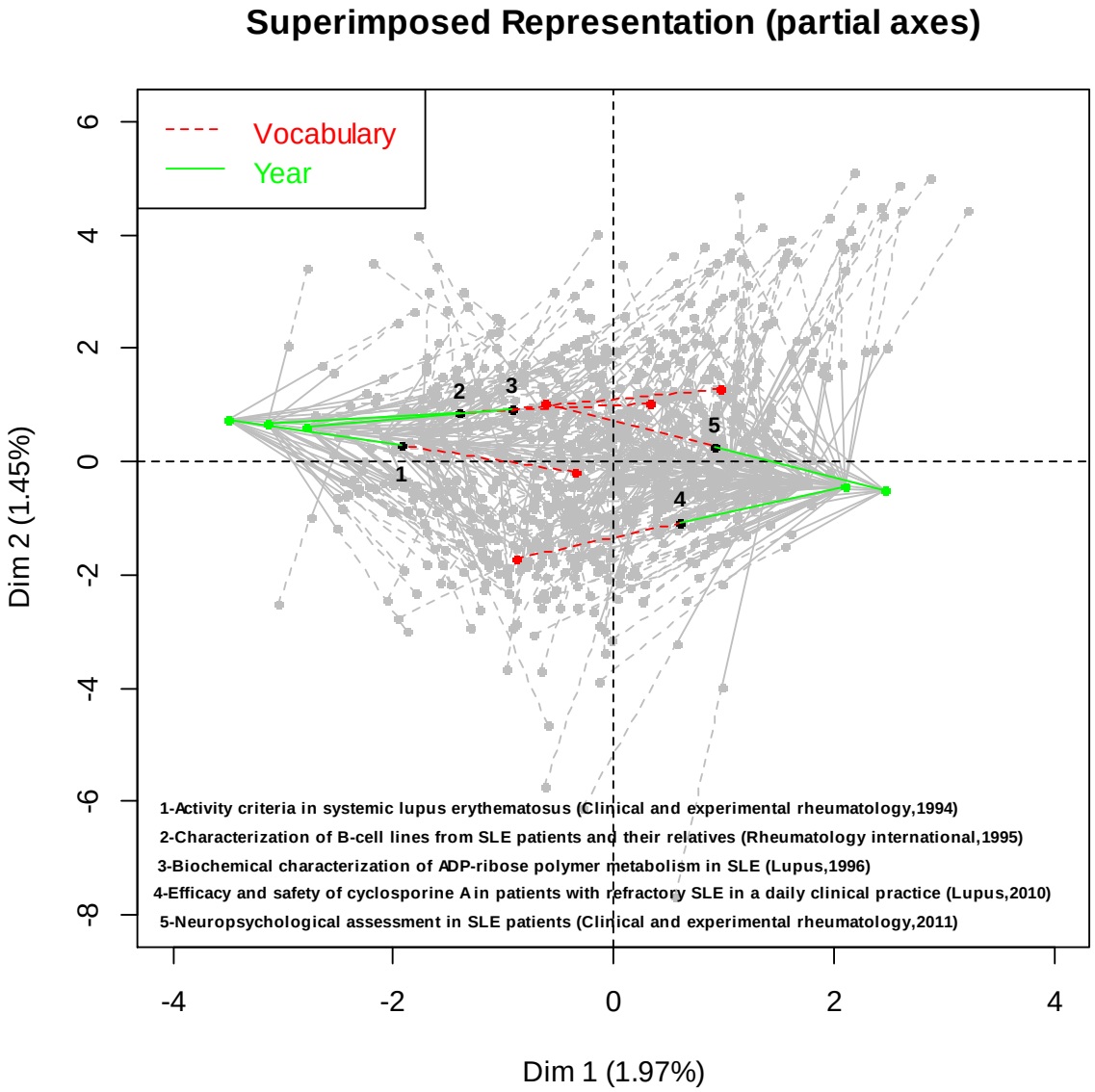


Evolution in the research concerning SLE



- Evolution of the vocabulary linked to chronology (first dimension)
- Early research (symptoms and drugs)
- Recent research (etiology and genetics)

Pioneering works in SLE



- Three pioneer works (1,2 and 3) from 1994 to 1996 dedicated to a genetic approach
- Two recent works from 2010 and 2011 (4 and 5) related to former topic “drugs”

Usefull information

- See also
  - plot.MFA (Draw the Multiple Factor Analysis (MFA) graphs)
  - missMDA (Impute the missing values of a dataset with the Multiple Factor Analysis model)
  - http://factominer.free.fr/