COPULA MODELS AND ITS APPLICATIONS IN DEPENDENCE MODELLING OF BIOMARKERS

Pál Rakonczai[1], Tamás Priksz[1], Zoltán Istenes[1], Bence Nagy[1]
Bernadette Rojkovich[2], Tamás Gáti[2]
[1]Healthware Consulting Ltd, Budapest, Hungary;
[2]Buda Hospital of Hospitaller Brothers of St. John, Budapest, Hungary

Background

- **RA DISEASE**: Rheumatoid arthritis (RA) is a chronic autoimmune disease causing inflammation in the joints and resulting in painful deformity and immobility. During the last decade around 50.000 patients became registered in Hungary .
- **AC**: Data of RA patients were collected from the Clinical Register of the Arthritis Centrum (AC) of Buda Hospital of Hospitaller Brothers of St. John
- **STUDY PERIOD**: from 1998.05.06. to 2015.27.02.
- **STUDY POPULATION**: 489 patients (278 treated with biological therapy) with 9034 visits
- **DATABASE:** Demographic parameters and biomarkers were registered
- **PERMISSION:** 26513-3/2013/EKU (329/2013.) by Scientific and Research Ethics Committee of the Medical Research Council

Biomarkers

- **HAQ index**: HAQ index measures disability/severity of patients with rheumatoid arthritis from 0 (no difficulty) to 3 (unable to do).
- **VAS index:** Visual analogue scale measures pain intensity with visual methods from 0 to 100.
- **DAS28 index:** Disease Activity Score of 28 joints is widely used as an indicator of RA disease activity and response to treatment. It is calculated using previously shown metrics and information about the state of the joints:

$$DAS28 = 0.56 * \sqrt{TCJ28} + 0.28 * \sqrt{SJC28} + 0.70 * ln(CRP) + 0.014 * VAS$$

where TCJ28/SCJ28 are the number of joints with tenderness upon touching/ swollen joints. CRP is produced by the liver which indicates inflammatory diseases. Its scale goes from 0 to 10.

Statistical methods

- **COPULA**: Nowadays there are more and more recent copula models aimed at describing the behavior of multivariate data sets.
- **WHY**? Instead of fitting only trend lines the entire multivariate distribution *H* can be modeled:

$$H(x,y) = Pr\{(X,Y) \leq (x,y)\} = C(F(x),G(y))$$

where *F* and *G* are the marginal distribution functions and *C* is the copula distribution function [Nelsen, 2006] .
- **ADVANTAGES**:
  - Non-linear structures are applicable
  - No restrictions on marginal distributions
  - Covariates can be handled too
- **MODELS**: There are several parametric families available as e.g. Gaussian, Frank or Gumbel.
- **MARGINS:** Any parametric family can be used.
- **COVARIATES**: Biological therapy (0/1) and age groups (under 40 / over 40)
- **METHOD:** 2D or 3D data were transformed into the unit cube by using empirical distribution and several copula families were fitted. For simulation parametric margins were used (rescaled beta distribution).
- **ESTIMATES:** Pseudo maximum likelihood method
- **SOFTWARE**: Estimates of copula parameters are available in the 'copula' package [Kojadinovic and Yan, 2014] of the R statistical programming language. Additional functions for fitting models with covariates were self-made.

Descriptive statistics

Figure 1. Scatterplots of biomarker pairs

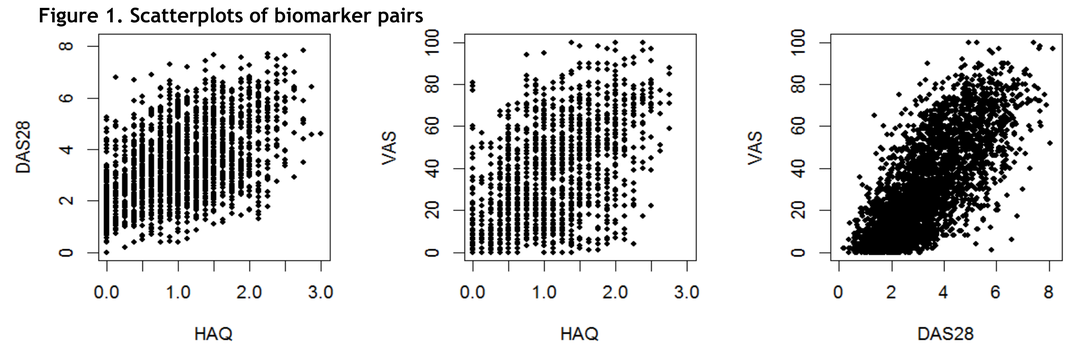


Table 1: Linear and rank correlations of biomarker pairs

Correlation		Linear correlation		
		VAS	DAS28	HAQ
Rank correlation	VAS	1	0.72 (n=3059)	0.54 (n=1568)
	DAS28	0.53 (n=3059)	1	0.52 (n=1994)
	HAQ	0.40 (n=1568)	0.37 (n=1994)	1

Results and conclusion

Copula models without covariates

Figure 2:Next to the sides of the triangles the 2D copula parameters, inside of the triangles the 3D copula parameters are presented.

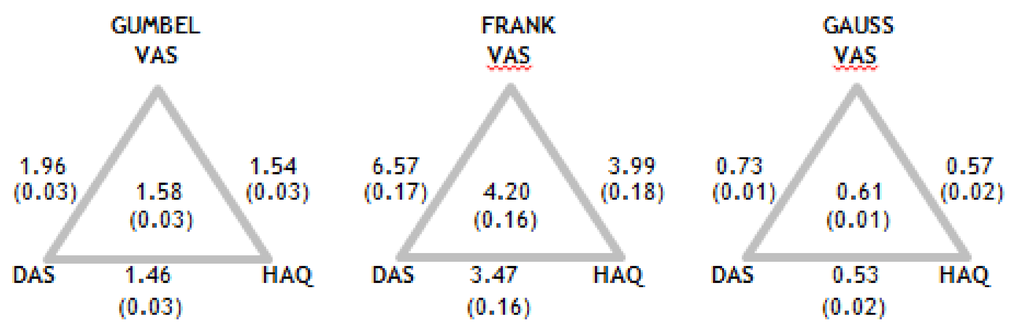
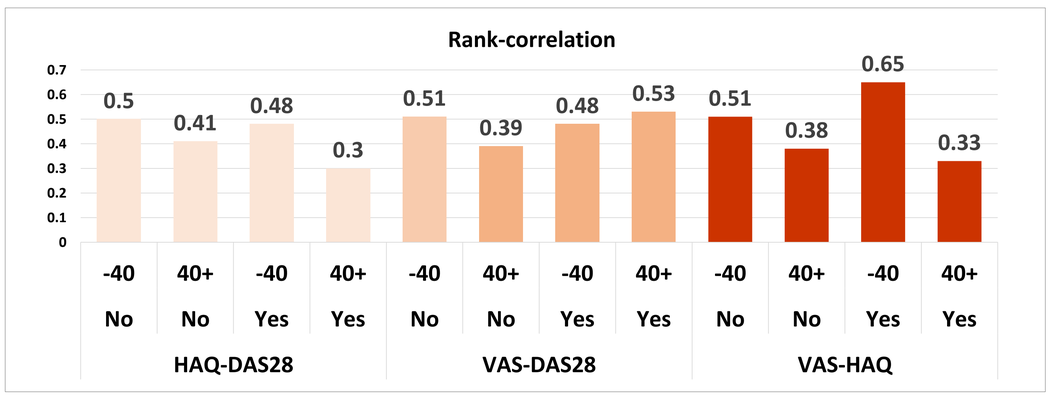


Figure 3: Rank correlations for different subgroups of data



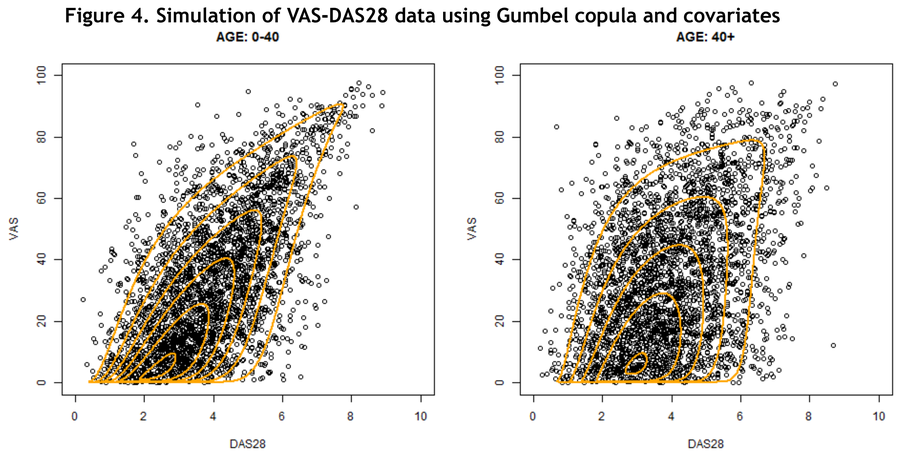
Copula models with covariates

Table 2: Regression coefficients of Gumbel copula models

Biological therapy	Age groups	HAQ-DAS28	HAQ-VAS	VAS-DAS28
No	-40	**1,75**	**1,72**	**1,72**
	40+	**-0,19**	**-0,26**	**-0,21**
Yes	-40	**0,11**	**0,67**	**0,08**
	40+	**-0,39**	**-0,28**	**-0,21**

- **Age and the biological therapy had significant effect on the dependence structure. The strength of the association was systematically milder in 40+ age cohort than in 0-40 age cohort. Biological therapy cohort had significant effect in the 0-40 age cohort, specially in the case of HAQ-VAS.**

Figure 4. Simulation of VAS-DAS28 data using Gumbel copula and covariates



References

[1] Kojadinovic I and Yan J (2014) Package 'copula'. For the manual of the R software package see http://cran.r-project.org/web/packages/copula/copula.pdf
[2] Nelsen, R. B. (2006) An introduction to Copulas. Springer, New York.
[3] Rakonczai P, Balázs T, Nagy B, Rojkovich B, Gáti T (2014) [Modeling Dependence Between Disability Status and Health Service Costs of Patients with Rheumatoid Arthritis in Hungary]

This poster is compiled by Healthware Consulting Ltd. ISPOR 18th Annual European Congress, Milan, 7-17 November 2015.
Healthware Consulting Ltd. / H-1093 Budapest, Közraktár str. 30. River Park Offices, VI. floor
Phone.: +36 1 279 1083 / Fax: +36 1 279 1084 / www.healthware.hu

Download our ISPOR posters directly from our webpage:
Poster code: PRM132