Seminar, 15th December 2016 h16.30
DISMA – Politecnico di Torino

Title: Using penalized regression models to select a microRNAs signature for prostate cancer

Authors: Lidia Sacchetto[1,2], Paola Ostano[3], Maurizia Mello-Grand[3], Ilaria Gregnanin[3], and Giovanna Chiorino[3]

[1]Università degli Studi di Torino
[2]Politecnico di Torino

[3]Lab of Cancer Genomics, Fondazione Edo ed Elvo Tempia, Biella (Italy)

Introduction
Nowadays the correct diagnosis of prostate cancer (PCa) often requires invasive tests, since the prostate-specific antigen (PSA) measurement gives high percentage of false positive. For so, it seems important to find non-invasive bio-markers to better detect PCa: microRNAs (miRNAs) circulating in plasma appear to be the ideal candidates. In this seminar I present the mathematical statistical methodology which allowed us to identify a diagnostic signature.

Methods
The software R (packages *limma, glmnet*) was used to fit:
- linear models for miRNA-array data to highlight differential expressed miRNAs among PCa, benign prostatic hyperplasia (BPH) and normal samples from healthy donors;
- a logistic regression model with lasso penalty on a subset of statistically significant differentially expressed miRNAs to build a classifier able to discriminate between PCa and BPH.
The accuracy of the classifier was assessed through the ROC curves (R package *pROC*).

Results
44 miRNAs were identified as the most differentially expressed in different comparisons. Among them, the lasso penalized logistic regression selected 10 miRNAs with non-zero coefficients, which were then combined in a score (the classifier). This score correctly classified 38 PCa out of 50 and 37 BPH out of 49. Adding PSA, we obtained: accuracy=85.9%, sensitivity=96%, specificity=75.5% and PPV=75%.

Conclusion
The penalized logistic regression allowed to identify a diagnostic signature of circulating microRNAs which classifies PCa better than PSA.