Journal of Statistics on Health Decision

EXTENDED ABSTRACT

P20

Statistical methods for the study of extracellular vesicles content and their potential as biomarkers of multiple myeloma aggressiveness

Carolina Pestana[1,2], Lisete Sousa[1,3], Emilie Carneiro[2], Filipa Barahona[2], Joana Caetano[2,4], Raquel Lopes[2], Bruna Ferreira[2,6],Cristina João[2,4,5]

[1] Departamento de Estatística e Investigação Operacional, Faculdade de Ciências, Universidade de Lisboa
[2] Myeloma and Lymphoma Research Programme, Champalimaud Research, Champalimaud Centre for the Unknown
[3] Centro de Estatística e Aplicações, Faculdade de Ciências, Universidade de Lisboa
[4] Hemato-Oncology Unit, Clinical Centre, Champalimaud Centre for the Unknown
[5] Immunology Department, Nova Medical School
[6] 7th Edition of Doctoral Programme in Medicine, NOVA Medical School

Introduction

Multiple Myeloma (MM) is an aggressive and incurable haematologic malignancy preceded by the pre-malignant stage of monoclonal gammopathy of undetermined significance (MGUS) [1]. At present, almost 46% of MM patients will die within the 5 years following diagnosis [2] and the other half will eventually relapse. Tools such as the disease aggressiveness score (R-ISS) [3] are currently used to estimate prognosis and guide the choice of treatment. However, they fail to accurately predict which patients will relapse early after treatment and die prematurely [4].

Nowadays, one requirement for the evaluation of myeloma prognosis is a bone marrow biopsy/aspirate to quantify and characterize the number of clonal plasma cells [5]. This is an invasive and painful procedure, and so non-invasive alternative methods for biomarkers evaluation are under investigation. Extracellular vesicles such as exosomes, secreted by cancer cells and found in the peripheral blood [6] are good candidates to be used as liquid biopsies to assess disease progression, severity and therapeutic efficacy [7,8].

Additionally, as the overall survival of MM patient´s relies on several variables that change over time, it may be inaccurate to assume that the prognosis of a patient assigned at the time of diagnosis remains unchanged during the course of the disease.

In the present pilot study, we analysed peripheral blood exosomes' content to assess the differences in protein expression between patients' groups at different stages of disease (MGUS and MM) and healthy donors (HD). By increasing the number of samples and analysing multiple time-points, the purpose of our project is to estimate the probabilities of disease progression, relapse and response to treatment through time associated with protein expression and other clinically relevant variables. Once this goal is accomplished, then we will also have achieved the clinical goal of finding biomarkers of disease in peripheral blood.

Methods

In this pilot study we analysed 11 patients from an observational, single-centre, cohort study. Blood samples were obtained from each patient at the time of diagnosis. Exosomal protein content of each sample was analysed by mass spectrometry (MS) in order to obtain the abundance expression of every identified protein in each sample. Then, Rank Product (RankProd R package) statistical method was applied in order to evaluate which proteins were differentially expressed in each of the two interest groups (1) patients' diagnosis and (2) R-ISS score. Although this method was originally designed to identify differentially expressed genes in a single experiment [9], it received widespread acceptance and is now used in several domains of -omics, such as transcriptomics and proteomics. The Rank Product (RP) is a non-parametric statistical method that makes it possible to detect variables consistently high ranked (upregulated or down-regulated) in a number of replicate experiments [10]. The expression levels for all proteins from all replicate samples were ranked, making the proteins with the smallest rank the ones with the most biological interest [11].

Keywords:
Multiple Myeloma, Exosomes, Differential protein expression

Corresponding author:
Carolina Pestana
carolina.pestana@research.fchampalimaud.org

First published: 23 OCT 2020

Some of the main advantages of this method are the fewer assumptions under the model and the strong performance even with noisy data or a low number of biological replicates [12]: since the RP does not depend on estimations of the measurement variance for each single variable, it is very useful when this estimate becomes unreliable in result to a low number of replicates [9].

The false discovery rate (FDR) quantifies the risk of false positives among all significant tests. Since we were in the presence of few samples and the later intention was the validation of the selected proteins, a FDR=0.2 was used.

Results

The diagnostic group included a total of 11 patients: 2 HD, 3 MGUS and 6 MM patients. Since R-ISS is a score of disease aggressiveness only estimated for the MM group, the comparison between R-ISS scores was made including the 6 MM patients, divided by 3 categories: 2 R-ISS I, 1 R-ISS II and 3 R-ISS III.

Using the Rank Product method, a total of 7 proteins with differential expression between diagnostic groups were identified by making the following comparisons: HD vs MGUS, HD vs MM and MGUS vs MM (adjusted p-value<0.05). By comparing the R-ISS III group with the joint group of R-ISS I and R-ISS II, 9 proteins showed to be differentially expressed (p-value<0.05).

Discussion and Conclusion

Our data for this pilot study revealed that EV protein expression differs according to diagnosis and risk score. All selected proteins are already described in the literature as having a direct or indirect relationship with cancer; therefore, we consider that, despite preliminary, our results are promising.

In addition to RP, there is another method for the analysis of protein expression data: the limma package, implemented in the R/Bioconductor software. This method fits a linear model to each protein, using a moderated t-statistics from the empirical Bayes procedure. However, and although it is a robust method in the context of few samples and allows the comparison of more than two groups simultaneously, the fact that it is a parametric method implies some assumptions that were not verified by our current data set.

Therefore, and since more than 100 patients' samples are already undergoing analysis by MS, the next step is the validation of the set of previously selected proteins and potentially identify new ones through the application of the limma package in this larger cohort. Moreover, by applying survival and multi-state models, the goal of our ongoing work is to analyse, not only the subset of selected proteins, but also other variables with clinical interest. These methods will allow us to estimate important measures for several interest groups (such as R-ISS score and diagnostic group) and models (such as "treatment-response-progression" model) in this wider set of samples.

References

1. Landgren, O., Kyle, R., Pfeiffer, R., Katzmann, J., Caporaso, N., Hayes, R. et al. Monoclonal gammopathy of undetermined significance (MGUS) consistently precedes multiple myeloma: a prospective study. Blood, 2009;113(22):5412-5417. https://doi.org/10.1182/blood-2008-12-194241
2. Myeloma - Cancer Stat Facts. (2020). Retrieved 28 July 2020, from https://seer.cancer.gov/statfacts/html/mulmy.html
3. Palumbo, A., Avet-Loiseau, H., Oliva, S., Lokhorst, H., Goldschmidt, H., Rosinol, L. et al. Revised International Staging System for Multiple Myeloma: A Report From International Myeloma Working Group. Journal Of Clinical Oncology, 2015;33(26):2863-2869. https://doi.org/10.1200/JCO.2015.61.2267
4. Manier, S., Liu, C., Avet-Loiseau, H., Park, J., Shi, J., Campigotto, F. et al. Prognostic role of circulating exosomal miRNAs in multiple myeloma. Blood, 2017;129(17):2429-2436. https://doi.org/10.1182/blood-2016-09-742296
5. Moreau, P., San Miguel, J., Sonneveld, P., Mateos, M., Zamagni, E., Avet-Loiseau, H. et al. Multiple myeloma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Annals Of Oncology, 2017;28:iv52-iv61. https://doi.org/10.1093/annonc/mdx096
6. Colombo, M., Giannandrea, D., Lesma, E., Basile, A., Chiaramonte, R. Extracellular Vesicles Enhance Multiple Myeloma Metastatic Dissemination. International Journal Of Molecular Sciences, 2019;20(13):3236. https://doi.org/10.3390/ijms20133236
7. Kalluri, R., LeBleu, V. The biology, function, and biomedical applications of exosomes. Science, 2020;367(6478):eaau6977. https://doi.org/10.1126/science.aau6977
8. Ferreira, B., Caetano, J., Barahona, F., Lopes, R., Carneiro, E., Costa-Silva, B., João, C. Liquid biopsies for multiple myeloma in a time of precision medicine. Journal Of Molecular Medicine, 2020;98(4):513-525. https://doi.org/10.1007/s00109-020-01897-9
9. Breitling, R., Armengaud, P., Amtmann, A., Herzyk, P. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Letters, 2004;573(1-3), 83-92. https://doi.org/10.1016/j.febslet.2004.07.055
10. Del Carratore, F., Jankevics, A., Eisinga, R., Heskes, T., Hong, F., & Breitling, R. RankProd 2.0: a refactored bioconductor package for detecting differentially expressed features in molecular profiling datasets. Bioinformatics, 2017;33(17):2774-2775. https://doi.org/10.1093/bioinformatics/btx292
11. Koziol, J. A. Comments on the rank product method for analyzing replicated experiments. FEBS letters, 2010;584(5):941-944. https://doi.org/10.1016/j.febslet.2010.01.031
12. Hong, F., Breitling, R., McEntee, C. W., Wittner, B. S., Nemhauser, J. L., Chory, J. RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. Bioinformatics, 2006;22(22):2825-2827. https://doi.org/10.1093/bioinformatics/btl476