



A pilot design to propose an apoptosis definition based on gene expression data

Nasim Beigi Boroujeni^{1,2}, Mandana Beigi Boroujeni^{1,2*}

Abstract

Introduction: Apoptosis is a programmed cell death commonly investigated in researches.

Objectives: According to the lack of a gold standard for definition of apoptosis, we conducted a pilot analysis to propose a new definition for apoptosis based on a previous gene expression data.

Materials and Methods: As a secondary analysis, a gene expression data of a vitrification thawing induced model of apoptosis conducted on ten mice ovaries was used. Half of the samples had been treated with selenium. *P53*, *Fas*, *Bax* and *Bcl2* were considered as apoptosis related genes. Their Δ CTs were reported. An apoptosis scoring system was designed based on regression analysis.

Results: In multiple regression of the genes, the only significant association was for *Bcl2* expression for prediction of apoptosis. Then a model was designed consisting of *Bcl2* and some interactions that the calculated amount of its formula was considered as the scoring system ($R^2 = 0.989$, $P (>F) < 0.001$, Root mean square deviation = 0.082). *Bax/Bcl2* ratio showed an acceptable goodness of fit for prediction of this score ($R^2 = 0.845$, $P (>F) < 0.001$, root mean square deviation = 0.219). No conclusive result was found for factor analysis.

Conclusion: The present study used a simple approach to propose statistical models for apoptosis. A comprehensive criterion should be designed apoptosis and other biological systems to be considered as a gold standard.

Keywords: Apoptosis, Systems biology, Statistical modeling, Factor analysis

Please cite this paper as: Beigi Boroujeni N, Beigi Boroujeni M. A pilot design to propose an apoptosis definition based on gene expression data. *J Parathyroid Dis.* 2022;10:e03.

Copyright © 2022 The Author(s); Published by Nickan Research Institute. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Apoptosis is a programmed cell death that can be physiologic or pathologic. Any problem in regulation of apoptosis (either inhibition or activation) may result in different diseases or complications such as cancers, autoimmunity, inflammation, neurodegenerative disorders and developmental defects (1). Apoptosis is a complex process consisting of intrinsic and extrinsic pathways. In extrinsic pathway, cell surface death receptors such as Fas are activated by their ligands. In intrinsic pathway, Bcl2 family or pro-apoptotic proteins lead to mitochondrial changes (2). Both pathways terminate to caspase proteins, and the pathways become common from the level of caspase3 (3). Among the apoptosis proteins, Bcl2 (as a member of Bcl2 family) is an inhibitory protein while Fas, Bax and P53 are activating proteins (4).

Due to the key roles of apoptosis in pathogenesis of many diseases, animal models of this process are used to investigate its mechanisms and effects of treatments (5, 6). Knowing the mechanisms of apoptosis helps the researchers to find better ways for diagnosis, management and treatment of diseases like cancers (1, 7). It is obvious that we need diagnostic methods for detection of

apoptosis. Currently, there are some techniques including electron microscopy (ultra-structural study), TUNEL assay and flow cytometry. Despite many of these methods, there are still some challenges (8). The limitations may also result in some pitfalls and mistakes (9). In addition, some serum and tissue biomarkers are used (10). Some of these biomarkers are apoptosis related molecules including Fas, Bax, Bcl2, P53, etc. Bax/Bcl2 (Bcl2/Bax) ratio is also used as a biomarker for its prognostic roles in some cancers (11).

Objectives

According to the lack of a gold standard for definition of apoptosis as well as little information about the diagnostic role of biomarkers at gene expression level, the present study was designed to perform a pilot analysis to propose a new definition for apoptosis based on a previous gene expression data.

Materials and Methods

Model design

As a secondary analysis study, the gene expression data of our previous study (5) on a vitrification thawing

Received: 10 August 2021, Accepted: 2 June 2022, ePublished: 24 June 2022

¹Razi Herbal Medicines Research Center, Lorestan University of Medical Sciences, Khorramabad, Iran.

²Department of Anatomical Sciences, Lorestan University of Medical Sciences, Khorramabad, Iran.

*Corresponding author: Mandana Beigi Boroujeni, Ph.D., Email: mandana.beigi@lums.ac.ir

■ Implication for health policy/practice/research/medical education

There is no gold standard for definition of apoptosis. We conducted this pilot analysis based on gene expression data to propose a new definition for apoptosis. Statistical modeling should be used further in future for biological systems.

induced model of apoptosis conducted on 10 NMRI mice ovaries was used. Among the samples, half of them had been treated with selenium in their cryomedia to inhibit apoptosis.

Definitions and variables

Apoptosis modeling samples: All the samples of the study as they underwent vitrification thawing process were considered as apoptosis modeling samples. Lack of a gold standard to approve the success of the modeling was a limitation.

Apoptosis positive samples: the half of the apoptosis modeling samples that were not treated with selenium in their cryomedia were considered as apoptosis positive samples.

Apoptosis negative samples: The half of the apoptosis modeling samples that were treated with selenium in their cryomedia were considered as apoptosis negative samples. Lack of a gold standard to rule apoptosis out was limitation.

Apoptosis related genes: *P53*, *Fas*, *Bax* and *Bcl2* were considered as apoptosis related genes. Their efficiency adjusted +ΔCTs (in comparison to GAPDH internal control) were reported as their expression unit.

Bax/Bcl2 ratio: the expression of Bax per Bcl2 was also calculated as a common biomarker.

Score: a scoring system was designed based on the best model of regression formula.

Latent variable: apoptosis was considered as an abstract concept analyzed as a latent variable. Apoptosis related genes were the observed variables.

Statistical analysis

Multiple linear regression was used for prediction of apoptosis negative samples (using logistic regression was not possible as the outcome was fully fitted). Then

a regression model was designed consisting of significant covariates and interactions with a stepwise approach. Thereafter, a scoring system was reported based on the regression formula. Confirmatory factor analysis (CFA) was used to show the apoptosis related genes as predictors of a common abstract concept (i.e. apoptosis). In addition, exploratory factor analysis was performed to find the components predicted by the apoptosis related genes. All the analysis were done in Stata 14 (StataCorp. LLC, USA) with significance level of 0.05.

Results

In the source study, *P53* and *Bcl2* expression were associated with the groups (up-regulation and down regulation respectively in favor of the apoptosis control group). In its multiple regression modeling, only *Bcl2* expression was significantly associated with the groups.

Multiple linear regression was performed to predict apoptosis negative samples. As it had been shown in the source study, the only significant association was for *Bcl2* expression. Then all the possible interactions were added to model and after that the non-significant interactions were removed (Table 1). The prediction formula is shown (equation 1).

$$Y = (-0.972) \times Bcl2 + 0.235 \times Bcl2 \times Fas + (-0.130) \times Bcl2 \times Bax + (-0.082) \times P53 \times Fas + 0.063 \times P53 \times Bax + 1.966 + \varepsilon$$

$$Y = Score$$

$$Score \rightarrow 1 \implies \text{apoptosis negative}$$

$$Score \rightarrow 0 \implies \text{apoptosis positive (Eq. 1)}$$

Since the scoring formula was complex, a multiple linear regression was performed to predict the score based on single genes expression. However, the only significant effect was for *Bcl2* expression (Table 2). Therefore, two simple linear regressions were used to predict the score based on *Bcl2* and *Bax/Bcl2* ratio. Among these two models, *Bax/Bcl2* ratio showed a better goodness of fit (Table 3, equation 2). According to this model, +ΔCT (*Bax/Bcl2*) <3 was in favor of apoptosis (Figure 1). It meant that the fold change (*Bax/Bcl2*) >8 was in favor of apoptosis (fold change = $2^{-\Delta\Delta CT}$).

$$Score = 0.472 \times \frac{Bax}{Bcl2} + (-1.062) + \varepsilon \text{ (Eq. 2)}$$

Table 1. Multiple linear regression for prediction of apoptosis negative samples

Covariates (unit: +ΔCT)	Beta coefficient	Standard error	T value	P (> T)	95% Confidence interval	
					Lower	Upper
Bcl2	-0.972	0.246	-3.95	0.017	-1.655	-0.288
P53#Fas	-0.082	0.014	-6.05	0.004	-0.120	-0.044
P53#Bax	0.063	0.012	5.29	0.006	0.030	0.095
Fas#Bcl2	0.235	0.033	7.17	0.002	0.144	0.326
Bax#Bcl2	-0.130	0.030	-4.30	0.013	-0.213	-0.046
Constant	1.966	0.700	2.81	0.048	0.023	3.908

Interaction sign (all the interactions are continuous).

R² = 0.989, P (>F) <0.001, Root mean square deviation = 0.082

Table 2. Multiple linear regression for prediction of the score

Covariates (unit: +ΔCT)	Beta coefficient	Standard error	T value	P (> T)	95% Confidence interval	
					Lower	Upper
P53	0.010	0.175	0.06	0.956	-0.439	0.459
Fas	0.109	0.155	0.71	0.512	-0.288	0.507
Bax	0.077	0.205	0.38	0.721	-0.449	0.604
Bcl2	-0.413	0.159	-2.60	0.048	-0.822	-0.004
Constant	0.175	3.194	0.05	0.959	-8.035	8.384

$R^2 = 0.856$, $P (>F) < 0.025$, Root mean square deviation = 0.267

Table 3. Simple regression for prediction of the score

Model	Covariates (unit: +ΔCT)	Beta coefficient	Standard error	T value	P (> T)	95% Confidence Interval	
						Lower	Upper
1	Bcl2	-0.373	0.062	-6.010	<0.001	-0.516	-0.230
	Constant	1.971	0.256	7.700	<0.001	1.381	2.561
	Characteristics	$R^2 = 0.819$, $P (>F) < 0.001$, Root mean square deviation = 0.237					
2	Bax/Bcl2	0.472	0.071	6.610	<0.001	0.307	0.637
	Constant	-1.062	0.246	-4.310	0.003	-1.630	-0.494
	Characteristics	$R^2 = 0.845$, $P (>F) < 0.001$, Root mean square deviation = 0.219					

Table 4. Confirmatory factor analysis to find associations of the genes with the latent variable

Pathway	Covariates	Standard Error	T value	P (> T)	95% Confidence Interval	
					Lower	Upper
P53						
Latent variable	1	(constrained)				
Constant	10.538	0.345	30.51	<0.001	9.861	11.215
Fas						
Latent variable	-0.556	1.254	-0.44	0.657	-3.013	1.901
Constant	8.450	0.283	29.83	<0.001	7.895	9.005
Bax						
Latent variable	-0.188	0.465	-0.4	0.686	-1.100	0.724
Constant	11.916	0.145	82.41	<0.001	11.633	12.200
Bcl2						
Latent variable	-1.551	2.661	-0.58	0.560	-6.767	3.665
Constant	3.940	0.408	9.66	<0.001	3.141	4.740
Variance (e.P53)	0.502	0.585				
Variance (e.Fas)	0.589	0.655				
Variance (e.Bax)	0.185	0.145				
Variance (e.Bcl2)	0.000	3.624				
Variance (Latent)	0.691	1.149				

CFA was conducted to show the predictive role of apoptosis related genes for apoptosis as an abstract concept. However, no significant path was found with the latent variable (Table 4). In addition, EFA showed only one factor with Eigenvalue >1 which had positive correlation with *Bcl2* and negative correlation with *P53* expression (considering +ΔCT) (Table 5).

Discussion

Nowadays apoptosis has a very important role in many diseases. This importance needs a gold standard definition and also verified assay methods. Nevertheless, a practical definition was challenging and its assay was difficult (8,

9). The present study was not aimed to design a gold standard, but wanted to conduct statistical modeling at gene expression level. Accordingly, a multiple linear regression model was designed and the score resulted from the model was considered as a variable to be used as the concept of apoptosis. Then Bax/Bcl2 ratio – as a common practical biomarker – was compared with the resulted score. This biomarker had an acceptable goodness of fit with the score.

In general, mathematical approach in systems biology had a good background. Apoptosis was not an exception among the biological systems. So far, few researchers tried to design statistical models for apoptosis. According to the study of Schleich and Lavrik, the first mathematical

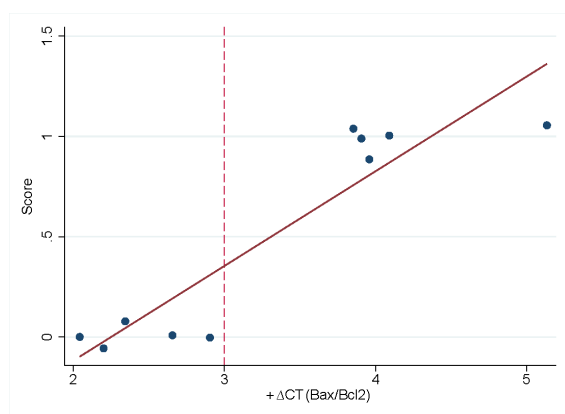


Figure 1. Linear prediction of the score with Bax/Bcl2 ratio table 3 (model 2)

Table 5. Exploratory factor analysis to find the factors and correlation coefficients

Factor characteristics	Factor 1	Factor 2	Factor 3	Factor 4
Eigenvalue	1.797	0.682	0.050	-0.175
Difference	1.115	0.632	0.225	
Proportion	0.764	0.290	0.021	-0.075
Cumulative	0.764	1.053	1.075	1.000
Variables and coefficients				
P53	-0.784	0.462	-0.015	0.172
Fas	0.411	0.679	0.007	0.370
Bax	0.412	0.046	0.197	0.790
Bcl2	0.919	0.069	-0.104	0.141

model of apoptosis was developed by Fussenegger et al which was about caspase activation (12, 13). This model was more mechanistic than statistical in contrast to our study. However, there were some studies with statistical modeling. Yang et al designed a Bayesian neural network for caspase cleavage (14). Afantitis et al designed a multiple linear regression model based on chemical compounds (15). Passante et al conducted principal component factor analysis. They proposed tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) and dacarbazine to induce apoptosis in researches (16).

Conclusion

The present study used a simple approach to propose statistical models for apoptosis. Among the investigated apoptosis related genes, *Bcl2* expression and some gene-gene interactions could predict the samples that underwent vitrification thawing induced model of apoptosis without anti-apoptotic treatments.

Limitations of the study

The limitations of this study were lack of a gold standard test in the source study and hence, the positivity and negativity of apoptosis was just contractual. It seems that in fact the concept of apoptosis is also abstract and

contractual. Although we did not find conclusive results for factor analysis, larger studies were necessary to find its pathophysiologic criteria. Over-fitting of the models was another limitation. A comprehensive criterion should be designed to be considered as a gold standard.

Acknowledgements

We acknowledge Seyyed Amir Yasin Ahmadi for his statistical analysis.

Authors' contribution

Both authors have equal contribution to all the authorship criteria. Both of the authors signed the final draft of the manuscript.

Conflicts of interest

The authors declare that they have no competing interests.

Ethical issues

This study was approved by the ethics committee of Lorestan University of Medical Sciences (IR.LUMS.REC.1397.195). Additionally, the study conforms to the ethical standards in the Declaration of Helsinki. Ethical issues (including plagiarism, data fabrication, double publication) have been completely observed by the authors.

Funding/Support

None.

References

1. Reed JC. Mechanisms of apoptosis. *Am J Pathol.* 2000;157:1415-30. doi: 10.1016/S0002-9440(10)64779-7.
2. Lossi L. The concept of intrinsic versus extrinsic apoptosis. *Biochem J.* 2022;479:357-384. doi: 10.1042/BCJ20210854.
3. Loreto C, La Rocca G, Anzalone R, Caltabiano R, Vespasiani G, Castorina S, et al. The role of intrinsic pathway in apoptosis activation and progression in Peyronie's disease. *Biomed Res Int.* 2014;2014:616149. doi: 10.1155/2014/616149.
4. Boroujeni MB, Peidayesh F, Pirnia A, Boroujeni NB, Ahmadi SAY, Gholami M. Effect of selenium on freezing-thawing damage of mice spermatogonial stem cell: a model to preserve fertility in childhood cancers. *Stem Cell Investig.* 2019;6:36. doi: 10.21037/sci.2019.10.01.
5. Nori-Garavand R, Hormozi M, Narimani L, Beigi Boroujeni N, Rajabzadeh A, Zarei L, et al. Effect of Selenium on Expression of Apoptosis-Related Genes in Cryomedia of Mice Ovary after Vitrification. *Biomed Res Int.* 2020;2020:5389731. doi: 10.1155/2020/5389731.
6. Beigi Boroujeni M, Salehnia M, Khalatbary AR, Pourbeiranvand S, Beigi Boroujeni N, Ebrahimi S. Effect of ovarian stimulation on the endometrial apoptosis at implantation period. *Iran Biomed J.* 2010;14:171-7.
7. Ghobrial IM, Witzig TE, Adjei AA. Targeting apoptosis pathways in cancer therapy. *CA Cancer J Clin.* 2005;55:178-94. doi: 10.3322/canjclin.55.3.178.
8. Martinez MM, Reif RD, Pappas D. Detection of apoptosis: A review of conventional and novel techniques. *Analytical Methods.* 2010;2:996-1004.
9. Darzynkiewicz Z, Bedner E, Traganos F. Difficulties and pitfalls in analysis of apoptosis. *Methods Cell Biol.* 2001;63:527-46. doi: 10.1016/s0091-679x(01)63028-0.
10. Chakraborty JB, Oakley F, Walsh MJ. Mechanisms and biomarkers of apoptosis in liver disease and fibrosis. *Int J Hepatol.* 2012;2012:648915. doi: 10.1155/2012/648915.
11. Ulukaya E, Acilan C, Yilmaz Y. Apoptosis: why and how does it occur in biology? *Cell Biochem Funct.* 2011;29:468-80. doi:

-
- 10.1002/cbf.1774.
12. Schleich K, Lavrik IN. Mathematical modeling of apoptosis. *Cell Commun Signal*. 2013;11:44. doi: 10.1186/1478-811X-11-44.
 13. Fussenegger M, Bailey JE, Varner J. A mathematical model of caspase function in apoptosis. *Nat Biotechnol*. 2000;18:768-74. doi: 10.1038/77589.
 14. Yang ZR. Prediction of caspase cleavage sites using Bayesian bio-basis function neural networks. *Bioinformatics*. 2005;21:1831-7. doi: 10.1093/bioinformatics/bti281.
 15. Afantitis A, Melagraki G, Sarimveis H, Koutentis PA, Markopoulos J, Igglessi-Markopoulou O. A novel QSAR model for predicting induction of apoptosis by 4-aryl-4H-chromenes. *Bioorg Med Chem*. 2006;14:6686-94. doi: 10.1016/j.bmc.2006.05.061.
 16. Passante E, Würstle ML, Hellwig CT, Leverkus M, Rehm M. Systems analysis of apoptosis protein expression allows the case-specific prediction of cell death responsiveness of melanoma cells. *Cell Death Differ*. 2013;20:1521-31. doi: 10.1038/cdd.2013.106.