ISSN print: 1224-5984
ISSN online: 2248-3942

Rom Biotechnol Lett. 2019; 24(1): 100-107
doi: 10.25083/rbl/24.1/100.107

Received for publication: September 20, 2018
Accepted: December 18, 2018

Original paper

Genetic Risk Score for Prostate Cancer in the Romanian Population

RAZVAN DANAU[1,2], DUMITRU CRISTINEL BADIU[1,3], PAUL IORDACHE[1,4], RADU URSU[1,5], VIORICA RADOI[1,4], STEFAN RASCU[1,2], GEORGE DANIEL RADAVOI[1,2], CRISTIAN SORIN SIMA[1,2], CRISTIAN TOMA[1,2], BOGDAN BRATICEVICI[1,2], MIHAELA MANDU[6], TITUS VALENTIN GRIGOREAN[1,3], VIOREL JINGA[1,2]

[1]Carol Davila University of Medicine and Pharmacy, Faculty of General Medicine, Bucharest, Romania
[2]Prof. Dr. Th. Burghele Clinical Hospital, Bucharest, Romania.
[3]Bagdasar-Arseni Clinical Emergency Hospital, Department of General Surgery, Bucharest, Romania
[4]Carol Davila University of Medicine and Pharmacy, Department of Epidemiology, Bucharest, Romania
[5]Carol Davila University of Medicine and Pharmacy, Department of Medical Genetics, Bucharest, Romania
[6]Bagdasar-Arseni Clinical Emergency Hospital, Bucharest, Romania.

Abstract

In our study, we investigated the utility of genome-wide association study results in defining a risk score based on prostate cancer risk variants in the Romanian population. The study population consisted of 1301 unrelated histopathologically confirmed prostate cancer (PCa) cases and 1,073 male controls consisting of patients admitted for urological and surgical conditions, excluding cancer.

None of the tested variants in the Romanian GWAS reached a genome-wide significance (p-value lower than 5x10-8), but 36 markers reached p-values of 1x10-7. 17 of the previously-reported SNPs replicated in the Romanian cohort. Evaluating the total risk observed in the general PCa GWAS compared with the GWAS performed in the Gleason Score <7 subcohorts, we found a 5.1 times difference in-between the two predicted risk scores in individuals carrying all 17 variants.

Keywords : prostate cancer, genetic epidemiology, GWAS, risk score, Romania

To cite this article: DANAU R, BADIU DC, IORDACHE P, URSU R, RADOI V, RASCU S, RADAVOI GD, SIMA CS, TOMA C, BRATICEVICI B, MANDU M, GRIGOREAN TV, JINGA V, Genetic Risk Score for Prostate Cancer in the Romanian Population. *Rom Biotechnol Lett*. 2019; 24(1): 100-107. DOI: 10.25083/rbl/24.1/100.107

*Corresponding author: DUMITRU CRISTINEL BADIU, Carol Davila University of Medicine and Pharmacy, Faculty of General Medicine, Bucharest Romania
E-mail: doctorcristianbadiu@yahoo.com

Introduction

Prostate cancer (PCa) is the most common cancer for the male population in Europe. The heritability of prostate cancer is contributed mostly to a large number of moderate to commonly occurring variants conferring lower risks and rarely to rare, occurring genetic variants with higher penetrance (BENAFIF & al. [1]) . The incidence rates of prostate cancer vary substantially worldwide, with a much higher incidence observed in the Western world than in Asian countries (MATSUDA & al. [2]).

The identification of genetic variation that increases susceptibility for PCa may help to inform screening strategies and clinical management of patients in the future. Despite the genetic variants individually only modestly influencing the risk, their cumulative impact is substantial (DADAEV & al. [3]). There are now more than 100 PCa risk-associated single nucleotide polymorphisms (SNPs) that have been described, and it has been estimated that these genetic loci increase the estimated proportion of the familial risk to 33% (Al OLAMA & al. [4]). These variants can be incorporated into a survival analysis to estimate their effects on the aggressiveness of prostate cancer and their impact can be estimated in the form of a genetic risk score.

The estimated age-standardized incidence of PCA in Romania is 37.9 per 100,000 men in 2012, and the estimated age-standardized mortality rate for PCA is 16.9 per 100,000 men (FERLAY & a [5]). From an epidemiological perspective, several studies were performed recently investigating the genetic profile of PCa in Romania. The results of these studies are indicating that many of the variants showing the most reliable statistical values in the Romanian population reside at loci that have been associated with several cancer types, so-called cancer hubs (IORDACHE PD & al. [6]). This clustering of the previously reported variants can provide a framework for possible risk score estimation in the Romanian population. In the present study, we investigated the utility of genome-wide association study results in defining a risk score based on prostate cancer risk variants in the Romanian population.

Materials and Methods

Study population

The subjects included in this study were male patients admitted between 2008 and 2016 in four clinics from Bucharest for various medical conditions. The study comprises 2374 hospital patients; 1301 unrelated histopathologically confirmed PCa cases, most of which had elevated PSA levels, and 1074 controls, consisting of patients admitted for urological and surgical conditions other than cancer. Blood samples and buccal swabs were collected for genotyping. PSA levels in plasma were measured for all subjects at hospital admission but were not used as exclusion criteria. All subjects gave written informed consent prior to enrolment and accepted the use of personal and clinical data and biological samples for genetic research (MAZILU & al. [7]).

Table 1. Clinical characteristics of the cohort

Decade	**% cases**
under 50	0.30%
50-60	1%
60-70	35.50%
70-80	44%
80-90	15%
Over 90	0.20%
T Staging	**% cases**
1A	2,32%
1B	1,11%
1C	15,65%
2A	1,71%
2B	1,91%
2C	4,04%
3A	43,80%
3B	6,06%
4	23,33%
Gleason score	**% cases**
2	0.20%
3	0.30%
4	1%
5	3.30%
6	13.20%
7	45.10%
8	20.30%
N Staging	**% cases**
N0	21.50%
N1	3.20%
Nx	75.30%
M Staging	**% cases**
M0	22%
M1	10%
Mx	68%

The Bioethical Committee of the Romanian College of Physicians approved the study, and the study protocols were approved by the National Ethical Board of the Romanian Medical Doctors Association in Romania. Trained interviewers performed face-to-face interviews, using standardized questionnaires, to collect personal data (ethnicity, marital status, education, height, and weight), lifestyle data (occupation, smoking, coffee, and tea consumption) and medical history (personal and familial). All subjects were of self-reported European descent. No significant differences were observed in other epidemiological features: BMI, smoking or alcohol consumptions (Table 1).

Genotyping and analysis of SNP data

DNA was extracted from whole blood and saliva at deCODE Genetics (Reykjavik, Iceland) and genotyped using Infinium OmniExpress-24 bead chips (Illumina). A total of 716,503 SNPs were genotyped for each included in the study. The genotype data were filtered using Plink! v1.07 (O'CONNEL & al. [8]) .Approximately 10% of the SNPs genotyped were removed using a Hardy–Weinberg equilibrium significance threshold of 5 × 10−6 and by excluding markers with a minor allele frequency lower than 1%. Prior to the imputation, each chromosome was phased in a single run using SHAPEIT (DELANEAU & al. [9]). Markers from Phase 3 October 2014 of the 1000 Genomes (GIBBS & al. [10]) were imputed into the 2024 chip-typed individuals using the IMPUTE2 software (HOWIE & al. [11]) with a posterior probability of 0.9 as a threshold to call genotypes. The set of genotypes were tested for population heterogeneity using principal component analysis in the ADMIXTURE software (ALEXANDER & al. [12]) and the results were consistent with a homogeneous population.

A total of 24,295,558 markers were generated by imputation for each in the study. Quality control for the imputation results was performed by removing markers with minor allele frequency less than 1%, call rate of 0.95 and info of 0.8. In total, 8,506,022 markers met the filtering criteria. An association test was performed between the 8.5 million imputed markers and a phenotype represented by a positive biopsy for prostate cancer. The association test was calculated using SNPTEST (MARCHINI & al. [13]), using a single binary variable as a response; all reported P-values are two-sided.

Selection of SNPs for replication of previous findings

A systematic literature review of variants associated with prostate cancer from previous GWAS' was completed on 01 September 2018 using the NHGRI catalogue of published genome-wide association studies (WELTER & al. [14)] as a starting point. A search query with 'prostate cancer' as a keyword was performed, and the inclusion criteria for selection were as follows: P-values <5 × 10−8 and a minor allele frequency above 5%. For each study, the following variables were collected: country and ethnicity of the participants, genotyping method, the source of controls and source of replication cohort, and several cases and controls in both discovery and replication study.

A total of 48 articles were initially obtained from the GWAS catalogue based on the keyword search. Twelve of the studies reported results only tangentially related to prostate cancer, while the remaining 25 studies reported associations with prostate cancer risk. We identified 123 unique markers previously associated with PCa.

Results

In an attempt to identify susceptibility loci for PCa, 1301 diagnosed PCa cases, and 1074 controls were tested. A total of 8.5 million imputed markers with a successful imputation score of >0.90% were included in the association test. Results across the genome are graphically illustrated in Figure 1, while the top findings are presented in Table 2.

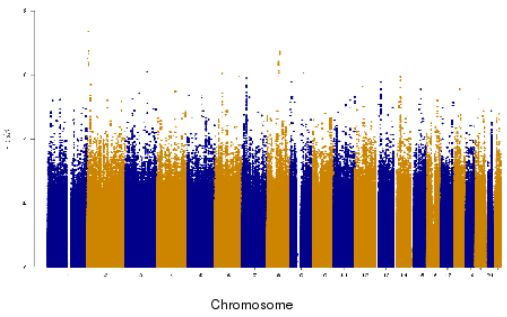

Figure 1. For the plot, the –log10 P-values (y-axis) of sequence variants are shown according to their chromosomal position(x-axis)

Table 2. Variants in the Romanian GWAS with lowest P-values for each locus MAF = Minor Allele Frequency

SNP	Chromosome	Position	Reference allele	Tested allele	MAF* in cases	MAF* in controls	OR	P-Value	Info
rs10490219	chr2	8400895	T	C	0.08	0.05	1.87	1.55E-07	1
rs3923300	chr8	126142476	C	G	0.11	0.07	1.65	2.33E-07	1
rs13253942	chr8	126154649	A	G	0.11	0.07	1.65	2.57E-07	1
rs34903473	chr8	126156276	C	G	0.11	0.07	1.65	2.57E-07	1
rs35215140	chr8	126200988	T	C	0.11	0.07	1.66	3.55E-07	1
rs35401436	chr8	126185276	A	T	0.11	0.07	1.66	3.70E-07	1
rs34246020	chr8	126246328	A	G	0.11	0.07	1.65	3.70E+47	1
rs12675943	chr8	126260092	A	G	0.11	0.07	1.69	4.47E-07	0.99
rs13273034	chr8	126134996	G	A	0.11	0.07	1.66	4.54E-07	0.99
rs35336507	chr8	126256028	G	A	0.11	0.07	1.7	4.82E-07	0.99
rs35955519	chr8	126256027	T	G	0.11	0.07	1.7	5.17E-07	0.99
rs72720569	chr8	126258407	T	G	0.11	0.07	1.68	6.11E-07	99
rs4601329	chr8	126254561	C	T	0.11	0.07	1.69	6.87E-07	0.99
rs72779189	chr2	8395447	G	A	0.08	0.05	1.88	7.99E-07	0.99
rs140930777	chr2	8395953	C	CTTGT	0.08	0.05	1.88	8.03E-07	0.99
rs72779193	chr2	8403566	A	G	0.08	0.05	1.85	8.13E-07	0.99
rs16866634	chr2	8375001	A	G	0.09	0.05	1.78	8.35E-07	0.98
rs6706154	chr2	8398487	C	T	0.08	0.05	1.84	1.07E-06	0.98
rs199704995	chr2	8402794	CA	C	0.08	0.05	1.84	1.17E-06	0.98
rs55937624	chr2	8388038	A	G	0.08	0.05	1.92	1.18E+46	0.97
rs72779188	chr2	8393977	T	C	0.08	0.04	1.95	1.20E-06	0.95
rs34313605	chr8	126271948	G	A	0.12	0.08	1.66	1.20E+46	0.94
rs72779172	chr2	8361400	C	G	0.08	0.05	1.85	1.30E-06	0.94
rs72779171	chr2	8359538	A	G	0.08	0.05	1.86	1.38E-06	0.94
rs16900462	chr8	126250796	A	G	0.14	0.1	1.61	1.43E+46	0.87
is112824993	chr8	126227755	T	G	0.11	0.07	1.65	1.44E-06	1
rs16900452	chr8	126243100	A	G	0.11	0.07	1.65	1.44E-06	1

None of the tested variants in the Romanian GWAS reached genome-wide significance (p-value < 5·10-8) but 36 markers had p-values < 1·10-7. In Table 1 there are presented the strongest association signals seen at 2p25.1 (LINCOO299 gene, intron 7/8) and 8q24.13 (NSMCE2 gene, intron 2/6). The data was obtained by filtering with the following conditions: imputation info >0.8, tested alleles frequency in controls >0.005.

Next, we tested the effect of 123 previously reported PCa variants in the Romanian population. 17 SNPs from 17 different loci were replicated in the Romanian cohort (P-value <0.05).

Also, for the 17 markers, we performed a second GWAS using a subcohort of PCa patients with the values of Gleason score above 7. The 17 markers replicated in the Romanian population are presented in Table 3.

Table 3. Previously reported PCa risk markers associated toPCa risk in the Romanian population with a P value < 0.05

RS-ID	Tested Allele	Ref Allele	OR	P-value	OR for the second test	The p-value for the second test
RS35148638	C	A	1,1905	0,000731	1,4213	0,000236
RS2242652	A	G	0,8721	0,002897	0,9412	0,002546
RS4713266	C	T	1,2541	0,0005421	1,8714	0,0002277
RS8102476	C	T	1,3323	0,003562	1,2224	0,0029581
RS10993994	A	G	0,9830	0,004372	0,9901	0,005376
RS2659124	A	T	1,1245	0,000939	1,3324	0,000672
RS4245739	A	G	0,4212	0,00321	0,7832	0,000913
RS11691517	G	A	0,9643	0,000574	0,7614	0,000729
RS6465657	T	C	1,0676	0,00473	1,1310	0,0061
RS9364554	T	C	0,8235	0,000671	0,9436	0,00064
RS35148638	C	A	0,9245	0,000731	0,9924	0,009722
RS12597458	G	C	0,8244	0,00095	0,9007	0,0012353
RS2005705	C	G	0,8543	0,003994	0,8130	0,003926
RS7127900	C	A	0,9676	0,000802	0,9887	0,000967
RS58262369	G	T	1,0557	0,0044	1,4412	0,00764
RS12791447	G	A	1,3215	0,00014	1,0324	0,00004
RS2238776	G	A	1,0400	0,000315	1,1900	0,000348

Evaluating the total risk observed in the general PCa GWAS compared with the GWAS performed in the Gleason Score <7 subcohorts, we found a 5.1 times difference in-between the two predicted risk scores in individuals carrying all 17 variants. The overall risk score for the 17 variants in the general PCa cohort is 0.63 compared to the risk score of 3.27 observed for the Gleason Score <7 subcohorts.

Discussions

A large amount of information on prostate cancer risk associated SNPs is available for multiple white case-control samples. However, less is known about whether these associations can be consistently replicated in Eastern European populations (JINGA & al.[15]; ARNOLD & al. [16]). Before trying to establish risk-based score on our results, we evaluated the expected p-values compared to the observed p-values. We observe no excess signal in the Q-Q plot when testing all marker (Figure 2-a); the observed p-values (blue line) show a comparable trend to the expected p-values (the red line).

However, we observe an excess of signals in the Q-Q plot when restricting to the set of previously reported variants (Figure 2-b); the observed p-values (blue line) show a comparable trend to the expected p-values (the red line). The Q-Q plots concurred with the plots from the study "Profile of common prostate cancer risk variants in an unscreened Romanian population," our study expanding the cohort of the above-mentioned paper (PANTEA S. & al [17]; NITIPIR C. & al [18]).

When we compare the odds ratios (OR) of the 17 variants in the general PCa GWAS to the GWAS performed in the Gleason Score <7 subcohorts, we observed a constant increase in risk for each replicated variant. The distribution of both OR sets is graphically illustrated in Figure 3.

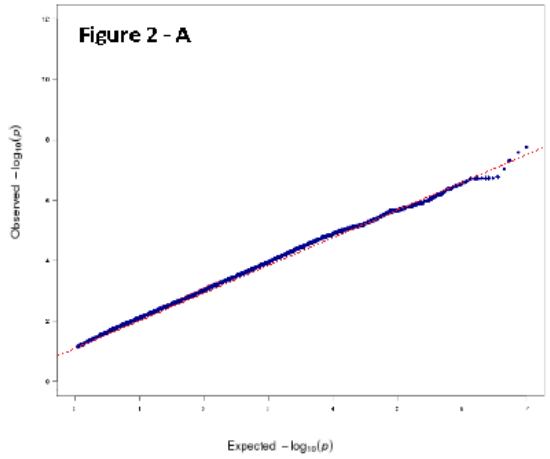

Figure 2(A) shows results from the genome-wide analysis;

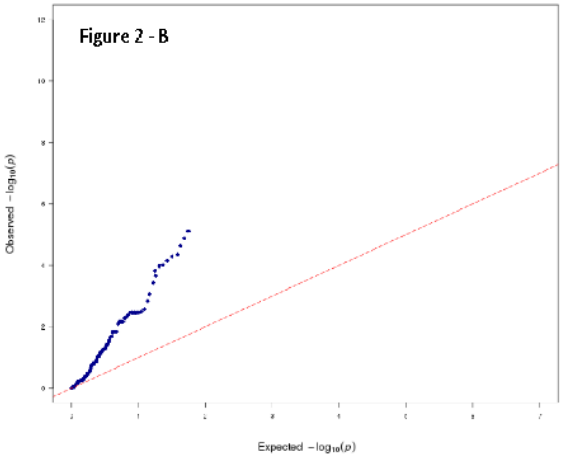

Figure 2 (B) shows results when restricted to GWAS catalog

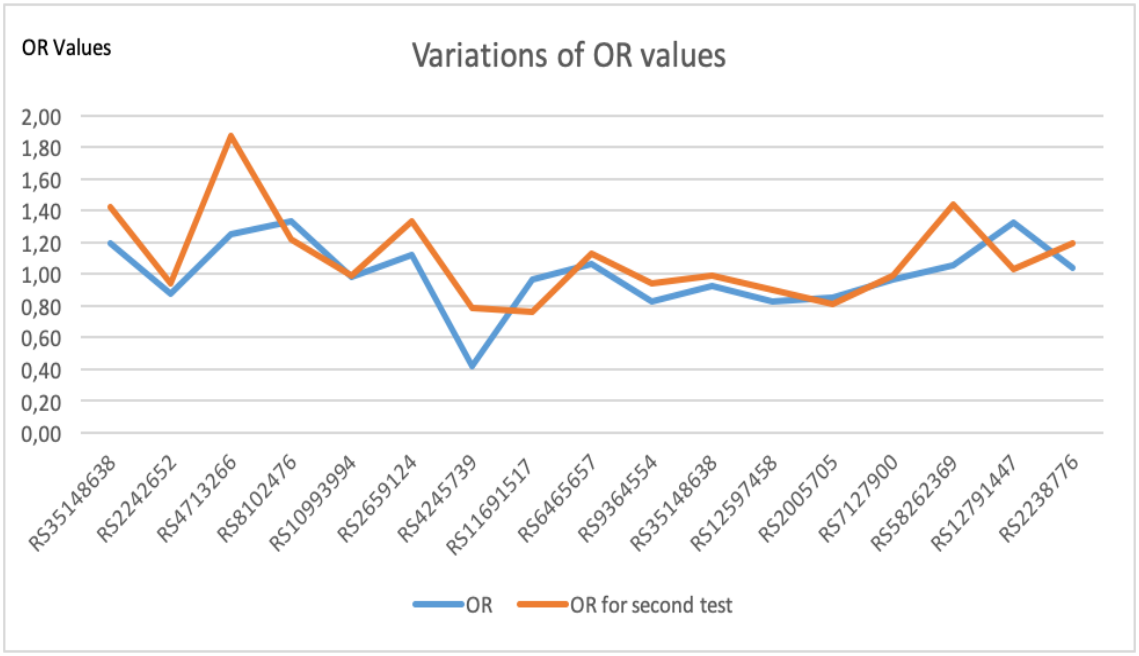

Figure 3: Distribution of OR's values for the 17 replicated markers

Conclusions

Including the 17 replicated Romanian SPNS can provide better risk assessment and can help guide relatives regarding the time of initiation and frequency of PCa screening. Although the genetic risk score can significantly improve screening practices in Romania, the complete clinical significance of this finding requires further evaluation. Furthermore, the inclusion of additional genetic variants from established prostate cancer susceptibility regions will improve disease prediction.

Conflict of interest disclosure

There are no known conflicts of interest in the publication of this article. The manuscript was read and approved by all authors.

Compliance with ethical standards

Any aspect of the work covered in this manuscript has been conducted with the ethical approval of all relevant bodies and that such approvals are acknowledged within the manuscript.

Acknowledgments

This study was funded in part by the European Union FP7 Program (ProMark project 202059) and by the EEA grant (ROMCAN project RO14.0017; EEAJRP.RO.NO.20131.10191).

References

1. S. BENAFIF, Z. KOTE-JARAI, R.A.EELES. A review of prostate cancer genome wide association studies (GWAS). Cancer Epidemiol Biomarkers Prev. January 2018: cebp.1046.2017. doi:10.1158/1055-9965.EPI-16-1046
2. T. MATSUDA, K. SAIKA. Comparison of Time Trends in Prostate Cancer Incidence (1973-2002) in Asia, from Cancer Incidence in Five Continents, Vols IV-IX. Jpn J Clin Oncol. 2009;39(7):468-469. doi:10.1093/jjco/hyp077
3. T. DADAEV, E.J. SAUNDERS, P.J. NEWCOMBE, et al. Fine-mapping of prostate cancer susceptibility loci in a large meta-analysis identifies candidate causal variants. Nat Commun. 2018;9(1):2256. doi:10.1038/s41467-018-04109-8
4. A.A. Al OLAMA, Z. KOTE-JARAI, S.I. BERNDT, et al. A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. Nat Genet. 2014;46(10):1103-1109. doi:10.1038/ng.3094
5. J. FERLAY, E. STELIAROVA-FOUCHER, J. LORTET-TIEULENT, et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer. 2013;49(6):1374-1403. doi:10.1016/j.ejca.2012.12.027
6. P.D. IORDACHE, D. MATES, B. GUNNARSSON, et al. Profile of common prostate cancer risk variants in an unscreened Romanian population. J Cell Mol Med. 2018;22(3):1574-1582. doi:10.1111/jcmm.13433
7. L. MAZILU, Z. NICULESCU, A.I. SUCEVEANU , et al. Ethical aspects of communication with cancer patients, REVISTA ROMANA DE BIOETICA.2010; 8(3):181-188
8. J. O'CONNEL, D. GURDASANI, O. DELANEAU, et al. A general approach for haplotype phasing across the full spectrum of relatedness. PLoS Genet. 2014;10(4):e1004234. doi:10.1371/journal.pgen.1004234
9. O. DELANEAU, B. HOWIE, A.J. COX, J.F. ZAGURY, J. MARCHINI. Haplotype Estimation Using Sequencing Reads. Am J Hum Genet. 2013;93(4):687-696. doi:10.1016/j.ajhg.2013.09.002
10. R.A. GIBBS, E. BOERWINKLE, H. DODDAPANENI, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74. doi:10.1038/nature15393
11. B.N. HOWIE, P. DONELLY, J. MARCHINI, et al. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. Schork NJ, ed. PLoS Genet. 2009; 5(6): e1000529. doi:10.1371/journal.pgen.1000529
12. D.H. ALEXANDER, J. NOVEMBRE, K. LANGE. Fast model-based estimation of ancestry in unrelated individuals. Genome Res. 2009;19(9):1655-1664. doi:10.1101/gr.094052.109
13. J. MARCHINI, B. HOWIE. Genotype imputation for genome-wide association studies. Nat Rev Genet. 2010;11(7):499-511. doi:10.1038/nrg2796
14. D. WELTER, J. MACARTHUR, J. MORALES, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res. 2014;42(Database issue): D1001-6. doi:10.1093/nar/gkt1229
15. V. JINGA, I.E. CSIKI, A. MANOLESCU, et al. Replication study of 34 common SNPs associated with prostate cancer in the Romanian population. J Cell Mol Med. 2016;20(4):594-600. doi:10.1111/jcmm.12729
16. M. ARNOLD, H.E. KARIM-KOS, J.W. COEBERGH, et al. Recent trends in incidence of five common cancers in 26 European countries since 1988: Analysis of the European Cancer Observatory. Eur J Cancer. 2015;51(9):1164-1187. doi:10.1016/j.ejca.2013.09.002

17. S. PANTEA, A. PITURU, S.M.R HAINAROSIE, et al. Testosterone therapy, new opportunities in diabetes mellitus, FARMACIA,2018;66(1):1-7.
18. C. NITIPIR, D. NICULAE, C.,ORLOV, et al. Update on radionuclide therapy in oncology, Oncology Letters,2017,14(6):7011-7