ABO System Blood Groups Distribution in Bulgaria, Based on a Dataset of the Patients of the University Hospital "Saint Anna", Sofia, Bulgaria, from 2015 to 2021

Vassia Atanassova[1[0000-0002-3626-9461]] , Nikolay Andreev[2] and Angel Dimitriev[3]

[1] Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences, 105 Acad. Georgi Bonchev Str., Sofia 1113, Bulgaria
[2] Transfusion Hematology Department, University Hospital "Saint Anna", Sofia 1 D. Mollov Str., Mladost 1, 1750 Sofia
[3] Institute of Biophysics and Biomedical Engineering, Bulgarian Academy of Sciences, 105 Acad. Georgi Bonchev Str., Sofia 1113, Bulgaria
vassia.atanassova@gmail.com

Abstract. The paper presents the analysis of a dataset containing the records of 47562 Bulgarian individuals, patients of the University Hospital "Saint Anna", Sofia, Bulgaria in the period from 1 January 2015 to 31 December 2021, and aims at the establishing the distribution of blood groups of the ABO system and the frequencies of the A_1 and A_2 subgroups and the Rh(D) antigen. In the frames of the conducted research, a series of data cleansing and data extraction procedures have been applied. The statistical analysis established the following frequencies of the ABO system blood groups in this sample of the Bulgarian population: A – 43.66%, B – 16.36%, O – 31.87% and AB – 8.11%. The prevalence of Rh(D) antigen determined was 86.38%. The distributions of A_1 and A_2 subgroups in the A group were determined as 91.92% and 8.09%, respectively, and the distributions of A_1B and A_2B subgroups in the AB group were established as 88.52% and 11.48%. These results are compared to the findings of other researchers regarding the distribution of ABO blood system blood antigens among the Bulgarian population, and slightly updates the results from the dataset of the patients of that Hospital from the period 2015–2020.

Keywords: ABO system, Blood group distribution, Rh(D) antigen distribution.

1 Introduction

Ever since its discovery, the ABO system of blood groups and blood group antigens and their frequencies in population have concentrated researchers' interest. The concept of ABO blood types is crucially important for transfusion medicine, as transfusion of incompatible ABO type of blood is the most common cause of death from blood transfusion, [8]. Antigens furthermore appear to play a significant role in human evolution

because the frequencies of different ABO blood types exhibit significant variability among various human populations, [7, 8, 11, 12], suggesting that a particular blood type conferred a selection advantage like resistance against an infectious disease, etc. [8].

Human migration further outlines the importance of the study of the ABO blood group antigens distribution, since the changing antigen proportions may complicate the provision of the most appropriate blood for blood transfusion, especially at places with lacking or weak blood donation culture.

It is important that every country maintains a regular, up-to-date record regarding the blood group distribution profile of its population. In this sense, the information about Bulgaria according to various publicly available online sources, needs concretization, update and additional references in order to be considered reliable. For instance, in a top-ranking web source dated to 2013, [7], the blood group distribution in Bulgarian population is given as 44% of A, 15% of B, 32% of O and 8% of AB, and one of the most popular websites Wikipedia has provided in [14] as of June 2017 (without citing a particular source) the following frequencies: A Rh(+) 37.4%, A Rh(-) 6.6%, B Rh(+) 12.8%, B Rh(-) 2.2%, O Rh(+) 28.0%, O Rh(-) 5.0%, AB Rh(+) 6.8%, AB Rh(-) 1.2%, summing respectively to the percentages in [7]. The same figures, as rounded to the nearest integer, are reported in another public database dated to 2019, [12]. Another available research from 2015 [13] reports these distributions for the Bulgarian population as differently as 39.96% of A, 16.84% of B, 35.80% of O, and 7.60% of AB. These observations show the definite need for conducting an up-to-date and detailed study aimed at generating recent, accurate and detailed data for Bulgaria, which has been the motivation behind the present research.

The present paper is structured as follows: in Section 2, we give the background of our study: source of the patients' data and methods of data cleansing and retrieval. Section 3 contains the results of the study, namely the established frequencies the ABO system blood groups and subgroups, Rh(D) antigen frequencies, and respective maintained ratios, all of these given across the patients' sex. The subsequent Section 4 makes a comparison with previous studies. Finally, Section 5 gives conclusion and directions for future research.

2 Materials and Methods

The research was conducted in the Centre of Transfusion Hematology in University Hospital "Saint Anna", Sofia, over a dataset of hospital patients registered in the period between January 1, 2015, and December, 31, 2021.

At the stage of primary data collection, we obtained the tabular data of: 6469 samples in 2015, 13716 samples in 2016, 8748 samples in 2017, 10214 samples in 2018, 10534 samples in 2019, 8769 samples in 2020, and 8968 samples in 2021.

We conducted the following procedure of cleansing the primary data, including:

1. Removal of records with incomplete or inaccurate entered personal data about the patients or data about their blood groups (e.g. recorded valued 'A' instead of 'A_1' or 'AB' instead of 'A_1B', missing value of Rh(D) antigen, etc.),

2. Removal of records for patients of foreign nationality (where the personal identity number is not in the "EGN" (unique citizenship number) format for Bulgarian nationals but in the "LNC" format for resident people of foreign nationality),
3. Removal of duplicate records with respect to the "EGN" (including numerous records of the blood samples of one and the same patient within a calendar year, as well as across the full six-year period),
4. Extraction from the "EGN" of the data about the patient's sex,
5. Anonymization of the data.

As a result of the data cleansing procedures, the total number of 67418 records in the primary input was reduced and the dataset over which our investigation was conducted and results are reported, includes data of 47562 unique individuals, including 22194 men and 25368 women. In 2019 and 2021, two similar studies, with a part of this data locked in the period 2015–2019, and 2015–2020, respectively were reported in [2] and [3].

In determining the blood group characteristics, all requirements have been met regarding the blood samples, test reagents and test erythrocytes, as postulated in Ordinance No. 18 / 2004 of the Bulgarian Ministry of Health. In the immune-hematological diagnostics, there has been included: determination of the ABO blood group, the A1 and A2 antigen subgroups, and the Rh(D) antigen. The determination of the blood groups from the ABO system has been performed by the crossmatching method. The blood test results have been recorded in the "Saint Anna" Hospital Information System. The present research does not contain data of the investigated erythrocyte antigens beyond the ABO and Rhesus factor systems.

The atomic approach to data adopted as early as the research design phase allowed the detailed breakdown of data simultaneously in blood group / subgroup, rhesus factor, and sex, and thus gave us the possibility to group and extract the data in various ways and formulate different conclusions.

3 Results

In this section, the results from the conducted research will be presented, from the most general to the most particular results. Thus, on the basis of the 47562 unique patients' records, we have established the following frequencies of the ABO system blood groups: A – 43.66%, B – 16.36%, O – 31.87%, and AB – 8.11% of all patients in the study (Table 1).

Table 1. Established frequencies of the ABO system blood groups

Blood group	Number of samples	Frequency,%
A	20767	43.66
B	7780	16.36
O	15156	31.87
AB	3859	8.11
Total	**47562**	**100.00**

The distribution of the blood groups after determining the Rh(D) antigen is the following: A Rh(+) – 37.70%; A Rh(-) – 5.96%; B Rh(+) – 14.19%; B Rh(-) – 2.17%; O Rh(+) – 27.45%; O Rh(-) – 4.41%; AB Rh(+) – 7.04%; AB Rh(-) – 1.07% of all investigated patients (Table 2).

Table 2. Established frequencies of the ABO system blood groups with the Rh(D) antigen determined.

Blood group	Number of samples	Frequency,%	Rh(D) Frequency within the group,%
A Rh(+)	17931	37.70	86.34
A Rh(-)	2836	5.96	13.66
Total A	**20767**	**43.66**	**-**
B Rh(+)	6748	14.19	86.74
B Rh(-)	1032	2.17	13.26
Total B	**7780**	**16.36**	**-**
O Rh(+)	13057	27.46	86.15
O Rh(-)	2099	4.41	13.85
Total O	**15156**	**31.87**	**-**
AB Rh(+)	3350	7.04	86.81
AB Rh(-)	509	1.07	13.19
Total AB	**3859**	**8.11**	**-**

Additionally, in the last column of Table 2, we have calculated the frequency with which both Rh(D) antigens occur in each blood group, respectively: A Rh(+) are 86.34% and A Rh(-) are 13.66% of all A patients; B Rh(+) are 86.74% and B Rh(-) are 13.26% of all B patients; O Rh(+) are 86.15% and O Rh(-) are 13.85% of all O patients; AB Rh(+) are 86.81% and AB Rh(-) are 13.19% of all AB patients.

Thus the following ratios have been determined: A Rh(+) / A Rh(-) = 6.32; B Rh(+) / B Rh(-) = 6.54; O Rh(+) / O Rh(-) = 6.22; AB Rh(+) / AB Rh(-) = 6.58. Hence, totally, the distribution of the Rh(D) antigen, according to our investigation, is 86.38% (41086 / 47562), while 13.62% (6476 / 47562) of the tested individuals have been determined as Rh(D) negative, and the ratio between all patients with the Rh(D) antigen and those without it is: 41086 / 6476 = 6.34. Compared to our previous research [3] based on the respective patients' dataset from the period 2015–2020, this marks in the year 2021 an increase in the ratio from 6.31 to 6.34.

The characterization of the subgroups of the A and AB blood groups is presented in the next Table 3: 40.13% of all 47562 patients are with the A_1 blood subgroup, 3.53% are with the A_2 blood subgroup, 7.18% are with the A_1B blood subgroup and 0.93% of all patients are with the A_2B blood subgroup. Going into deeper detail regarding the Rh(D) antigen, we have the following picture: A_1 Rh(+) – 34.72%; A_2 Rh(+) – 2.98%; A_1 Rh(-) – 5.41%; A_2 Rh(-) – 0.55%; A_1B Rh(+) – 6.21%; A_2B Rh(+) – 0.83%; A_1B Rh(-) – 0.97; A_2B Rh(-) – 0.10% of all the 47562 studied patients.

Table 3. Established frequencies of the A and AB blood group subgroups with the Rh(D) antigen determined

Blood subgroup	Number of samples	Fre-quency, %	A_1/A_2 Frequency within the blood group,%	Rh(D) Frequency within the blood subgroup,%
A_1 Rh(+)	16513	34.72	79.51	86.51
A_1 Rh(-)	2575	5.41	12.40	13.49
Total A_1	**19088**	**40.13**	**91.91**	-
A_2 Rh(+)	1418	2.98	6.83	84.45
A_2 Rh(-)	261	0.55	1.26	15.55
Total A_2	**1679**	**3.53**	**8.09**	-
A_1B Rh(+)	2956	6.21	76.60	86.53
A_1B Rh(-)	460	0.97	11.92	13.47
Total A_1B	**3416**	**7.18**	**88.52**	-
A_2B Rh(+)	394	0.83	10.21	88.94
A_2B Rh(-)	49	0.10	1.27	11.06
Total A_2B	**443**	**0.93**	**11.48**	-

The frequencies of the A_1 and A_2 subgroups within blood group A in this study has been determined, respectively, as 91.91% (19088 / 20767) and 8.09% (1679 / 20767) of all A patients. The frequencies of the A_1B and A_2B subgroups against blood group AB has been determined, respectively, as 88.52% (3416 / 3859) and 11.48% (443 / 3859) of all AB patients.

These findings are visualized on Figure 1.

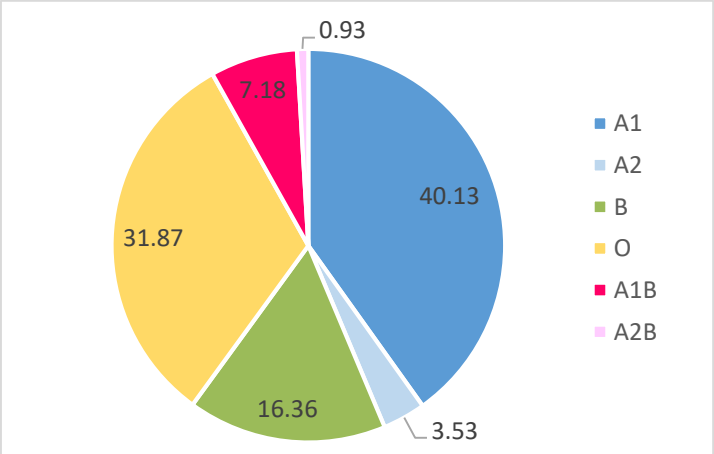

Fig. 1. Distribution of the ABO system blood groups and subgroups

The next Table 4 presents the determined frequencies of the ABO system blood groups in both sexes. The ratio of the blood samples of male to female individuals is 46.66% to 53.34%.

Table 4. Frequencies of the ABO system blood groups – males and females

Blood group	Number of samples		Frequency, %	
	M	F	M	F
A	9572	11195	43.13	44.13
B	3690	4090	16.62	16.12
O	7099	8057	31.99	31.76
AB	1833	2026	8.26	7.99
Total	**22194**	**25368**	**100.00**	**100.00**

The data show the following distribution within the group of male individuals: A – 43.13%, B – 16.62%, O – 31.99%, and AB – 8.26% of all studied male patients; and within the group of female individuals: A – 44.13%, B – 16.12%, O – 31.76%, and AB – 7.99% of all female patients in the study.

The distribution of the blood groups after determining the Rh(D) antigen is the following for males and females (Table 5). The distribution of the Rh(D) antigen in males, according to our investigation, is 86.99% (19306 / 22194), and in females is 85.86% (21780 / 25368). Respectively, determined as Rh(D) negative have been 13.01% of the tested males (2888 / 22194), and 14.14% of the tested females (3588 / 25368). In more details, the last two columns of Table 5 present the frequencies of the Rh(D) antigen within each of the blood groups for both sexes.

Table 5. Established frequencies of the ABO system blood groups with the Rh(D) antigen determined – males and females

Blood group	Number of samples		Rh(D) frequency within the group,%	
	M	F	M	F
A Rh(+)	8307	9624	86.78	85.97
A Rh(-)	1265	1571	13.22	14.03
Total A	**9572**	**11195**	**-**	**-**
B Rh(+)	3201	3547	86.75	86.72
B Rh(-)	489	543	13.25	13.28
Total B	**3690**	**4090**	**-**	**-**
O Rh(+)	6183	6874	87.10	85.32
O Rh(-)	916	1183	12.90	14.68
Total O	**7099**	**8057**	**-**	**-**
AB Rh(+)	1615	1735	88.11	85.64
AB Rh(-)	218	291	11.89	14.36
Total AB	**1833**	**2026**	**-**	**-**

As presented in Table 3 (for all tested patients), the distributions of the A_1 and A_2 antigen in patients with the A blood group is as follows: A_1/A = 91.87% (19088 / 20767) and A_2/A = 8.13% (1679 / 20767). Respectively, these distributions in the group of males and the group of females are the following: Males: A_1/A = 91.50% (8758 / 9572); A_2/A = 8.50% (814 / 9572) ; Females:A_1/A = 92.27% (10330 / 11195); A_2/A = 7.73% (865 / 11195).

Again, as presented in Table 3 (for all tested patients), the distributions of the A_1 and A_2 antigen in patients with the AB blood group is as follows: A_1B/AB = 88.52% (3416 / 3859); A_2B/AB = 11.48% (443 / 3859). Respectively, these distributions in the group of males and the group of females is the following: Males: A_1B/AB = 88.33% (1619 / 1833); A_2B/AB = 11.67% (214 / 1833); Females: A_1B/AB = 88.70% (1797 / 2026); A_2B/AB = 11.30% (229 / 2026).

In details the information about the distributions of the A_1 and A_2 antigen in the patients with A and AB blood groups is presented in the Table 6. The following Figure 2 graphically represents the hitherto presented data.

Table 6. Established frequencies of the A and AB blood group subgroups with the Rh(D) antigen determined – males and females

Blood subgroup	Number of samples		A_1/A_2 Frequency within the blood group,%		Rh(D) Frequency within the blood subgroup,%	
	M	F	M	F	M	F
A_1 Rh(+)	7603	8910	79.43	79.59	86.81	86.25
A_1 Rh(-)	1155	1420	12.07	12.68	13.19	13.75
Total A_1	**8758**	**10330**	**91.50**	**92.27**	-	-
A_2 Rh(+)	704	714	7.35	6.38	86.49	82.54
A_2 Rh(-)	110	151	1.15	1.35	13.51	17.46
Total A_2	**814**	**865**	**8.50**	**7.73**	-	-
A_1B Rh(+)	1423	1533	77.63	75.67	87.89	85.31
A_1B Rh(-)	196	264	10.69	13.03	12.11	14.69
Total A_1B	**1619**	**1797**	**88.32**	**88.70**	-	-
A_2B Rh(+)	192	202	10.47	9.97	89.72	88.21
A_2B Rh(-)	22	27	1.20	1.33	10.28	11.79
Total A_2B	**214**	**229**	**11.67**	**11.30**	-	-

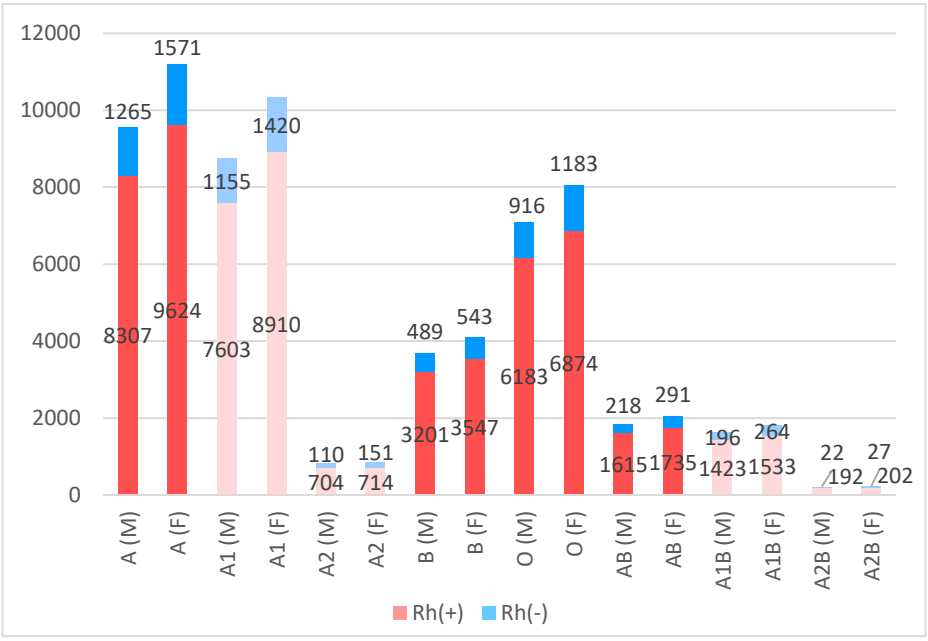

Fig. 2. ABO system blood and subgroups distribution among males and females

4 Comparison with previous studies

In the next Table 7, we present a comparison of the results obtained in the course of our present investigation with the results from similar investigations of the blood group distributions in Bulgarian population, conducted by other authors in the past, which are found in the available Bulgarian literature [1, 9, 11]. As our aim has been to present all the available data in the most comparable way, wherever necessary, missing table entries have been populated on the basis of appropriate computations, [2].

	1957, [8]	**1977, [1]**	**2012, [9]**	**2021 [3]**	**2022**
A	**44.80**	**43.30**	**43.20**	**43.62**	**43.66**
A_1 / A	*87.73*	*86.77*	*90.28*	*91.42*	*92.27*
A_2 / A	*12.27*	*13.23*	*9.72*	*8.58*	*7.73*
B	**16.80**	**16.60**	**16.10**	**16.42**	**16.36**
O	**32.10**	**32.50**	**32.60**	**31.79**	**31.87**
AB	**6.30**	**7.60**	**8.10**	**8.18**	**8.11**
A_1B / AB	*76.40*	*n/a*	*83.95*	*87.56*	*88.70*
A_2B / AB	*23.60*	*n/a*	*16.05*	*12.44*	*11.30*
TOTAL	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**

Table 7. Comparison of the presented results (2022) with results from previous studies (1957, 1977, 2012, 2021)

As presented graphically on Figure 2, we can outline the most notable trend of increase of the AB blood group frequency at the expense of lesser changes of the frequencies of the A, B and O blood groups. At the same time, according to the available data, visualized on Figure 3, there is a noticeable monotonous increase of the A_1 and A_1B frequencies at the expense of the monotonous decrease frequencies of the A_2 and A_2B frequencies, over a period of about 65 years.

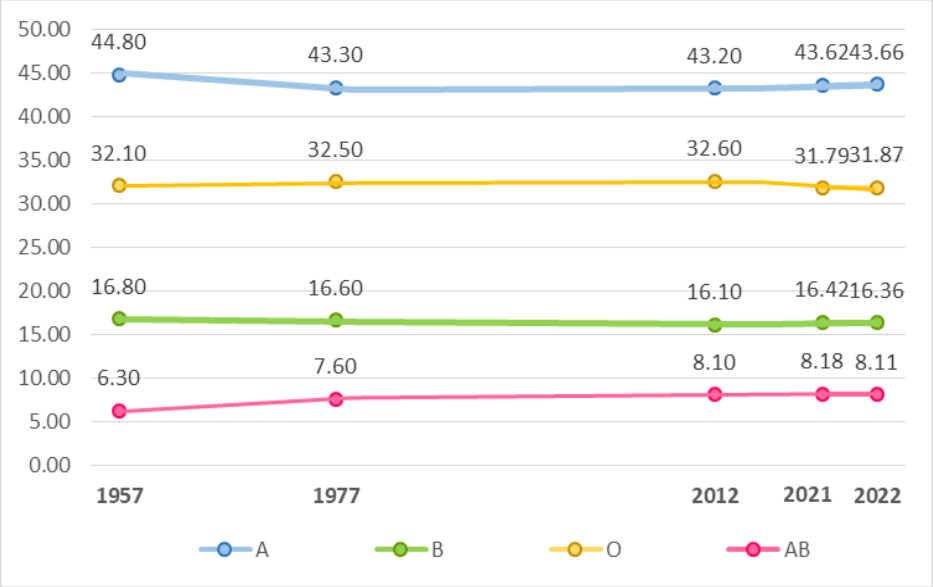

Fig. 3. Temporal change of the distributions
of the A, B, O and AB blood groups (in years 1957, 1977, 2012, 2021)

5 Conclusions and further research

The results from the present research are informative for the specialists of transfusuion hematology in Bulgaria regarding the actual state-of-the-art of the ABO system blood groups distributions, as well as the transfusion capacity of Bulgarian population.

On the basis of the available cleansed dataset of patients of the University Hospital "Saint Anna" collected in the period from 2015 to 2021 year, there have been envisaged at least two new steps of further research in this direction. From the patients' records, we have been able to extract automatically information about their year of birth (which varies in the range from 1913 to 2021 year), as well as information about their district place of birth. Thus we will be able to trace the temporal trends in blood group distribution in a sample of Bulgarian population that has lived for a period of a bit more than a century. Second, the patients' personal identity numbers contain information about the birthplace locations of the patients. Extracting this information would be potentially able to outline some additional patterns in the geographical distribution of the blood groups over the Bulgarian territory. Any significant finding in this latest direction of research would be informative for the easier detection and optimized routing of blood banks across different regional transfusion hematology centres in Bulgaria, especially in cases when rarer types of blood are being urgently sought.

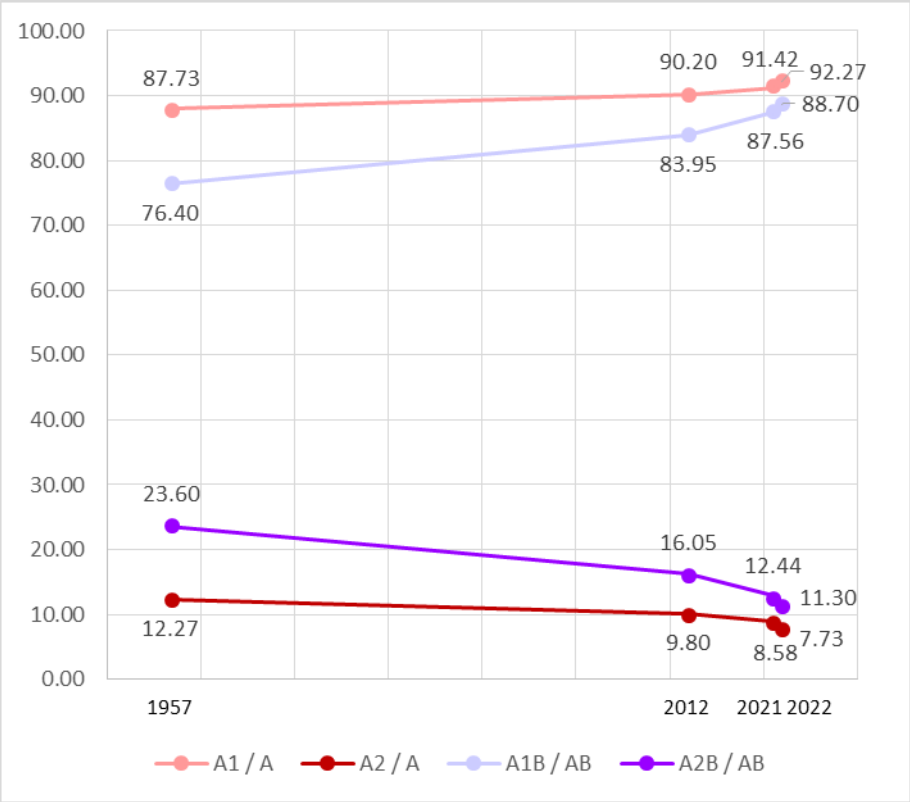

Fig. 4. Distributions of the A_1 and A_2 antigens
in the A and AB blood groups (in years 1957, 1977, 2012, 2021)

Additionally, when collected and treated in this level of detail, the results provide ample opportunities for application of the recently developed multicriteria decision making method called InterCriteria Analysis (ICA), [6. 10]. Such a research has been previously applied to a part of the dataset concerning patient in the University Hospital "Saint Anna" in the period 2015–2019, as reported in [2, 3], and developed in a related research in [4]. The application of ICA can further allow the discovery of trends and patterns of knowledge from the reported data, and further benefit the blood transfusion specialists regarding the blood transfusion capacity of the Bulgarian population.

Acknowledgement

The authors are grateful for the support provided under Grant No. KP-06-N-22/1 "Theoretical research and applications of InterCriteria Analysis" of the National Science Fund of Bulgaria.

References

1. Anastasov, A., Doychinova, N., Iliev, D., Kurteva, B., Nedyalkova, M., Prodanov, P., Shinkov, D. (1977). Immunohematology, Sofia, p. 55 (in Bulgarian).
2. Andreev, N., & Atanassova, V. (2021). InterCriteria Analysis of the blood group distribution of patients of Saint Anna Hospital in 2015–2019. In: Atanassov K.T. et al. (eds) Advances and New Developments in Fuzzy Logic and Technology. Proceedings of IWIFSGN'2019. Advances in Intelligent Systems and Computing, Vol. 1308, pp. 158-165. Springer, Cham. DOI: 10.1007/978-3-030-77716-6_14.
3. Andreev, N., & Atanassova, V. (2021). Distribution of the ABO System Blood Groups in the Patients of the University Hospital "Saint Anna", Sofia, Bulgaria in the Period 2015-2020. 1st Scientific Conference on Transfusion Hematology, Varna, Bulgaria, 24–26 September 2021 (in Bulgarian).
4. Andreev, N., Sotirova, E., Ribagin, S.: InterCriteria analysis of data from the centres for transfusion haematology in Bulgaria. Comptes rendus de l'Acade'mie bulgare des Sci. 72(7), 982–990 (2019)
5. Andreev, N., Vassilev, P., Ribagin, S., & Sotirov, S. (2019). InterCriteria Analysis of data for blood collection in the Transfusion Hematology Department, University Hospital "St. Anna", Sofia. Notes on Intuitionistic Fuzzy Sets. 25(2), pp. 88-95, DOI: 10.7546/nifs.2019.25.2.88-95
6. Atanassov, K., Mavrov, D., Atanassova, V.: InterCriteria decision making: a new approach for multicriteria decision making, based on index matrices and intuitionistic fuzzy sets. Issues Intuitionistic Fuzzy Sets Gen. Nets 11, 1–8 (2014)
7. BloodBook.com (2013, March 18). Racial & Ethnic Distribution of ABO Blood Types. Retrieved on October 14, 2022, from http://www.bloodbook.com/world-abo.html.
8. Dean, L. (2005). Blood Groups and Red Cell Antigens. National Center for Biotechnology Information (US), Bethesda (MD), pp. 31-33. Retrieved on October 14, 2022, from https://www.ncbi.nlm.nih.gov/books/n/rbcantigen/pdf/.
9. Dobreva, A., Doychinova, N., Vasilev, N., eds. (1988). Transfusion Hematology, pp. 68, 114-115. State Publishing House "Medicina I Fizkultura", Sofia (in Bulgarian).
10. Doukovska, L., Atanassova, V., Sotirova, E., Vardeva, I., Radeva, I.: Defining consonance thresholds in InterCriteria analysis: an overview. In: Hadjiski, M., Atanassov, K.T. (eds.) Intuitionistic Fuzziness and Other Intelligent Theories and Their Applications. SCI, vol. 757, pp. 161–179. Springer, Cham (2019). https://doi.org/10.1007/978-3-319-78931-6_11
11. Popov, R., Petrov, N., & Vaseva, V. (2012). Distribution of blood groups of the ABO system in the Military Medical Academy immunohematological diagnostics. Bulgarian Medical Journal, VI(2), pp. 45-48 (in Bulgarian).
12. RhesusNegative.net (2019). Blood Type Frequencies by Country including the Rh Factor. Retrieved on October 14, 2022, from http://www.rhesusnegative.net/themission/bloodtype-frequencies/.
13. Salduz, Z., Cetin, G., Karatoprak, C., Ozder, A. Bilginc, M., Gültepe, I. & Gul, O. (2015). ABO and Rh Blood Group Distribution in Istanbul Province (Turkey). Istanbul Medical Journal. 16, pp. 98-100. DOI: 10.5152/imj.2015.14890. Retrieved on October 14, 2022, from http://cms.galenos.com.tr/Uploads/Article_20948/IMJ-16-98-En.pdf.
14. Wikipedia contributors. (2017, June 21). Blood type distribution by country. In Wikipedia, The Free Encyclopedia. Retrieved from https://en.wikipedia.org/w/index.php?title=Blood_type_distribution_by_country&oldid=786847072.