This is an Accepted Manuscript for *Calitatea Vieții*. Subject to change during the editing and production process.

ETHNIC AND LINGUISTIC DIFFERENCES IN THE COVID-19 MORTALITY IN RURAL LOCALITIES IN MOLDOVA

MAXIM SLAV
VITALIE ȘTÎRBA

The COVID-19 pandemic revealed a visible discrepancy in mortality levels between countries, regions, and populations depending on their socioeconomic and demographic characteristics. Scientific literature shows the influence of individual, behavioural and institutional factors on COVID-19 outcomes, including risk of death. Additionally, state institutions seemed to have a crucial influence on COVID-19 mortality depending on their capacity to respond timely to population health challenges by reducing the risk of death and unnecessary disease sequelae. This research is conducted based on a hypothesis that people respond to the COVID-19 crisis depending on the information available in their mother tongue and their usually spoken language, which ultimately leads to a discrepancy in COVID-19 mortality between populations by characteristics of ethnolinguistic groups (ethnicity, mother tongue, and usually spoken language). Thus, by employing a linear regression model, we compared the level of COVID-19 mortality among the main ethnicities in Moldova. Our results revealed a significantly higher mortality level in the Russian-speaking population. We speculate that our results can be explained by the influence of external, often Russian-language, misinformation about COVID-19. This likely contributed to greater vaccine hesitancy, particularly regarding western-made vaccines. Additionally, higher levels of institutional mistrust among ethnic minorities in Moldova, combined with limited institutional capacity to communicate effectively with these communities, may have further reinforced this trend. The results could serve the central and local authorities in implementing targeted health policies to diminish health inequalities among populations by socioeconomic and ethnolinguistic characteristics.

***Keywords**: health inequalities; COVID-19 mortality; ethnolinguistic differences in coronavirus disease mortality*.

Address of the corresponding authors: Maxim SLAV, Centre for Demographic Research, NIER, Academy of Economic Studies of Moldova, Chisinau, Moldova, e-mail: slav.maxim@ase.md (https://orcid.org/0000-0002-3061-1692); **Vitalie ȘTÎRBA**, Centre for Demographic Research, NIER, Academy of Economic Studies of Moldova, Chisinau, Moldova, PhD candidate, Department of Demography and Geodemography, Charles University, Prague, Czechia, e-mail: stirba.vitalie@ase.md (https://orcid.org/0000-0001-5948-6509).

CALITATEA VIEȚII, XXXVI, nr. 2, 2025, pp. , https://doi.org/10.46841/RCV.2025.02.02

INTRODUCTION

Inequalities in health are often studied in the context of developing and improving health policies, primarily targeting vulnerable population segments. The COVID-19 pandemic has emphasised visible differences in mortality among the population depending on the regional settlement and its ethnic and socioeconomic criteria.

In this research, we have attempted to assess the prospective differences in overall and coronavirus disease mortality during the COVID-19 pandemic in the rural population in Moldova by its ethnic and linguistic characteristics. For this, we analysed ethnically and linguistically homogeneous rural localities, for which we used anonymised individual data on age, date of birth and death, place and cause of death, and the localities' population size. The research includes 841 communes[1] with an overall population of 1.4961 million persons.

Five ethnic groups represent 75%+ of the population in at least one commune in rural Moldova: Moldovans/Romanians[2], Ukrainians, Gagauz, Bulgarians, and Russians (*Figures 1* and *2*). The first four ethnic groups have compact clustering settlements. The Moldovans/Romanians are mostly settled in the centre, northeast, and south along the banks of the Prut and Nistru Rivers. Ukrainian villages are predominantly in the country's northern region, and the Gagauz and Bulgarian ones are in the south. Ethnically and linguistically, the rural settlements are also predominantly homogeneous. The ethnic and linguistic regionalisation allows us to study robustly the differences in registered overall and coronavirus disease mortality in the population by selected characteristics.

The literature offers several explanations for the differences in COVID-19 mortality between ethno-linguistic groups. These include factors such as media consumption, different behavioural patterns, and interaction with state institutions. The reasons can be differentiated with three variables: the spoken language, mother tongue, and ethnicity. While these are strongly correlated with each other, there are groups of people for which these characteristics do not intersect, e.g., there are ethnically Romanian groups that primarily speak Russian and consume Russian media. The data availability does not allow us to further differentiate between the causes, and we hence test the influence of these factors in a set (i.e., all the reasons to believe ethnicity rather than mother tongue is influential etc.). This study presents the existing differences in mortality during the COVID-19 pandemic

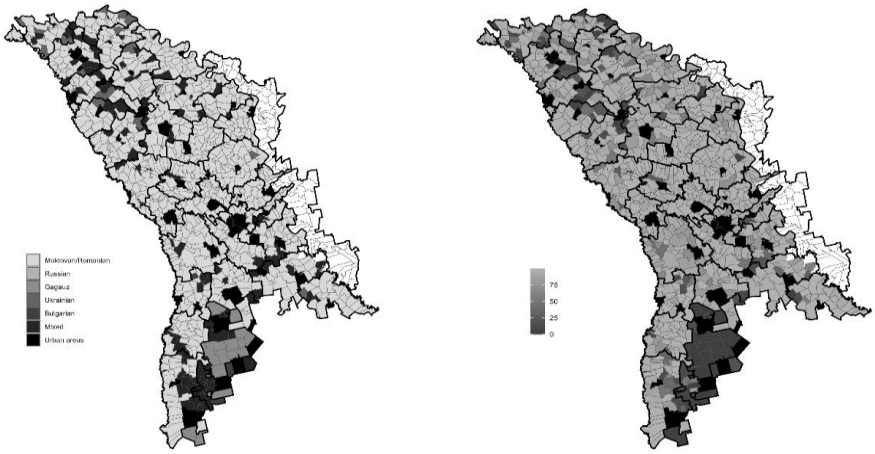
[1] A commune is an administrative-territorial unit comprising one or several small villages grouped for administrative efficiency. In this article, we use interchangeably the terms village and commune.

[2] There are two competing names for the dominating Moldovan ethnicity: “Romanians” and “Moldovans”. In this paper, we use these terms interchangeably to refer to ethnicity. We use the term “Romanian-speakers” to refer to the linguistic group that uses the state language as their mother tongue.

among the population based on ethnic and linguistic strata. Heterogeneities in socioeconomic and household characteristics were also considered.

Figures 1 and 2

The distribution of ethnicities (on the left) and Romanian-speakers (on the right) at the level of communes in Moldova



Notes: Except for the Mixed group, all the others stand for 75% ethnic group majorities. Black areas represent urban localities; white areas represent the Transnistrian region for which data is unavailable.

Source: The 2014 Population and Housing Census.

While several polls registered the differences between the ethno-linguistic groups' behaviour, the data they offer are strongly limited. Our main analysis is based on the data offered by the 2014 census and the exclusive dataset on the causes of death in Moldova in 2020 and 2021. The census data is available at the communal level (between several hundred and several thousand individuals each). While we recognise that individual-level data is more appropriate to our study, it does not prohibit us from drawing several conclusions. We tested the hypotheses with regression analysis. We used three main variables: the shares of the spoken languages, those of the mother tongues, and those of the ethnicities. Overall, our results emphasise a higher COVID-19 mortality rate among the non-Romanian speakers. However, the variables performed differently, and we interpret it via a greater importance of the spoken (Russian) language compared to the mother tongue or ethnicity. Additionally, several control variables had robust performance. Specifically, we found that communes with high out-migration and a greater proportion of one-person households tend to have higher COVID-19 mortality rates. We further built the argument with the sensitivity analysis, reviewing the

outliers and influential observations. We also replicated the regressions with the excess mortality as the dependent variable.

The results are subject to the following important limitations. The study does not include the localities uncontrolled by authorities on the left bank of the Nistru River, which have a significant Russian media influence. Additionally, at the level of communes, important data on media consumption, vaccination against COVID-19, the level of social and institutional trust etc., were not available. The small populations at the level of communes are not making it possible to standardise the death rate; thus, as a dependent variable, the Crude Death Rate was employed, where the potential effect of the population structure on mortality was adjusted by using the proportion of the population 65+ as a control variable. The excess mortality rate also suffers from this drawback.

The results might be valuable for central and regional authorities in implementing targeted policies to improve population health and reduce morbidity and lifespan inequalities.

SUMMARY OF RESEARCH EVIDENCE

During the COVID-19 pandemic, most European countries have encountered unprecedented population health challenges that perturbed optimistic life expectancy trends previously registered (Schöley *et al.* 2022). Even though the situation has been to some extent ameliorated, theCOVID-19 pandemic revealed existing regional disparities in mortality between the countries and regions within the countries (Bonnet *et al.* 2024; Pizzato *et al.* 2024). Generally, the regional disparities in coronavirus disease mortality are obviously resulting from the efficiency of the national health systems and their infrastructure (Lupu and Tiganasu 2022). Additionally, the bureaucratic capacity in responding to the pandemic challenges (e.g., providing the timely possibility to vaccination against COVID-19) is essential in reducing the risk of cause-specific mortality and possible disease sequelae (Bayati *et al.* 2022). However, mortality disparities vary between the populations depending on their socioeconomic status (Dukhovnov and Barbieri 2022) and other behavioural and lifestyle factors (Foster *et al.* 2022).

Among the main determinants that lead to discrepancies in COVID-19 mortality between the regions and localities are population density and age structure (Chang *et al.* 2022; Pascoal and Rocha 2022). Nevertheless, the size of the household appears to have a significant influence on COVID-19 mortality. An overcrowded household is positively associated with a risk of death (Varshney, Glodjo and Adalbert 2022), while persons living alone have a higher risk of developing a severe COVID-19 outcome (Gillies *et al.* 2022).

Education, as one of the key indicators used in measuring health inequalities (Galobardes *et al.* 2006), has significantly contributed to COVID-19 mortality disparities among different population strata. The population with higher levels of

education had, on average, more opportunities in regard to remote jobs and higher income, which allowed them to reduce the risk of exposure to coronavirus disease (Pathak *et al.* 2022; Lee *et al.*2022). Ultimately, education shapes the population's social behaviour to a higher resilience towards epidemiological challenges (Zhuo and Harrigan 2023).

The evidence shows that vaccination against COVID-19 has significantly contributed to a decline in the number of hospitalisations and deaths associated with coronavirus disease (Hoxha *et al.* 2022). Nevertheless, among the factors that influenced disparities in COVID-19 mortality across countries and regions were vaccination availability and hesitancy among different population groups. The evidence shows a higher prevalence of refusal and hesitancy in vaccination against COVID-19 in Eastern European countries (Toshkov 2023), where differentiation in vaccination hesitancy is influenced by multiple factors such as education (Paul, Steptoe and Fancourt 2021; Sylvester 2021), media and social media consumption (Wilson and Wiysonge 2020), political views and institutional trust (Costa-Font, Garcia-Hombrados and Nicińska 2023).

The literature also gives evidence for the existence of centralised COVID-19 misinformation campaigns (Foster 2021; Keegan 2022;Magdin 2020), which aimed at undermining trust in the press and state institutions, increasing the lack of clarity amid public opinion on authorities' actions to control the spreading of coronavirus diseases and vaccination campaigns. As a part of the post-Soviet space, Moldova is consistently listed as one of the targets of such campaigns (Marsili 2021), but its multi-ethnicity makes the exposure to propaganda vary depending on the spoken language. Nevertheless, a significant proportion of the Moldovans/Romanians emphasised disbelief in the efficacy of the vaccination against coronavirus disease, while a visible proportion of the population viewed COVID-19 rather as a myth (Institutul de Opinii Publice 2021).

Thus, the past research suggests three competing hypotheses on the mortality differences between ethnolinguistic groups. If the basis of the COVID-19 mortality differences is genetic, ethnicity would be the primary explanatory factor. If the basis of the differences is in the behavioural patterns more common in certain cultures, the mother tongue (as a proxy of culture) would have a stronger link. If the differences are due to propaganda, the mortality will have a stronger link with spoken languages. The three groups must intersect significantly, but the strength of the effect is likely to vary accordingly.

DATA AND METHODS

The research incorporates data from three official administrative and open sources:

1. The National Agency for Public Health provided individual data on death counts based on the date of birth and death, place where the person died, diagnosed

cause of death and existing comorbidities, out of which we selected COVID-19 deaths that correspond to U071 and U072 of the ICD-10 special codes (NAPH). This data was converted from individual to communal level. To estimate the excess mortality during the COVID-19 pandemic, we filtered out for each commune the overall number of deaths.

2. The National Bureau of Statistics provided data on population size at the commune level for the 2020–2021 period, which is based on 2014 census data and adjusted by vital statistics according to information on births and death counts, but also internal and international migration (NBS). This source is presenting the overall male and female population at the level of communes, which is not disaggregated by age structure. To avoid the effect of the population structure on overall and COVID-19 mortality, we used the proportion of the population 65+ for each commune from 2014 Population and Housing Census (NBS 2014).

3. The data on independent variables on population, such as ethnicity, mother tongue, the spoken language, education, household characteristics etc., were obtained from the (unadjusted by vital events) 2014 census data (NBS 2014).

The datasets were merged with CUATM codes (Classification of administrative-territorial units) for the settlements' region and commune.

Additionally, we use the data provided by the Public Opinion Barometer (Institutul de Opinii Publice 2021) for the preliminary study of the heterogeneities of the rural Moldovan trust in the government and the vaccines.

The deaths from COVID-19 were unevenly distributed across the two calendar years. As a result, we constructed the dependent variable in the following way:

$$covid\_rate = \frac{coviddeaths_{2020} + coviddeaths_{2021}}{(population_{2020} + population_{2021})/2}$$

The main regression formula is as follows:

$$\begin{aligned} Crude\ COVID-19\ death\ rate \\ = \beta_0 + \beta_1 * Population_{65+} + \beta_2 * Nonmigrants + \beta_3 \\ * Computerization + \beta_4 * Solitary\ People + \beta_5 \\ * Unheated\ households + \beta_6 * Secondary\ education \\ + \beta_7 * Ethnolinguistic\ Group + \varepsilon \end{aligned}$$

We follow these regressions with the sensitivity analysis section, where we use Leave One Out Cross-Validation (LOOCV) and Cook's distance to find and study outliers. We also do regressions using a different dependent variable, excess mortality. The number of expected deaths for 2020 and 2021 is calculated as average of the number of deaths in 2015−2019, and the number of observed deaths is calculated as the average number of deaths in 2020−2021. In other words, the formula is as follows:

$$excess\ mortality = 100 * \frac{observed\ deaths - expected\ deaths}{expected\ deaths}$$
$$= 100 * \frac{\overline{deaths_{2020}^{2021}} - \overline{deaths_{2015}^{2019}}}{\overline{deaths_{2015}^{2019}}}$$

The main advantage of excess mortality is its incorporation of the additional deaths resulted from the systematic pressure on the healthcare system. It also accounts for the misclassified COVID-19 deaths. The evidence shows that the crude death rate that we use underestimates up to 35% of the additional number of deaths (Kung *et al.* 2021). At the same time, our research question primarily concerns COVID-19 mortality, and differentiations between the hypotheses are based on the COVID-19 specifics.

DESCRIPTIVE ANALYSIS

Since the dependent variable is continuous and has no conglomerated value, linear regression was employed for this research. See *Table no. 2* for descriptive statistics on the dependent (as well as the independent) variables.

Most of the rural localities in Moldova are ethnically and linguistically homogeneous (see *Table no. 1*). The language and ethnicity operationalisations do not fully overlap (e.g., there are far more Russian speakers than there are Russians, and Romanian is spoken by some minorities). Still, the inclusion of both will result in severe multicollinearity, and we used the operationalisations separately. Four main variables were used: (a) the share by spoken languages, (b) the shares by mother tongues, (c) the shares by ethnicities, and (d) dummy variables for the main (more than 75%) ethnolinguistic group. We used the three former operationalisations twice: once to measure the share of the Moldovans/Romanians in the commune and once to measure the shares of the Russians, the Ukrainians, and the Gagauz separately. With respect to the dummy variables, we constructed one for each of the five categories: Ukrainians, Gagauz, Russians, Bulgarians, and Mixed. The Moldovans/Romanians were the base category. We did not include the Bulgarians in the first three operationalisations since they are rarely encountered outside of the few villages in which they are the majority. We did not include Romani people, as, according to the census, there are no communes where they dominate. The census did not provide data on other ethnolinguistic groups.

The size of communes varies based on the ethnic groups that inhabit them (see *Figure 3*), where Gagauz, Ukrainian, and Moldovan/Romanian communes diverge in population size, while Russian and Bulgarian-dominating communes tend to be medium-sized. Moldovan-dominating communes comprise more than 85% of all rural localities, with about 88% of the overall rural population. *Table*

*no. 1*summarises the distribution of the communes by ethnicity and the number of people living there. *Figure 3* also shows the villages' distribution with regard to the dominating ethnicity and the dependent variable.

Table no. 1

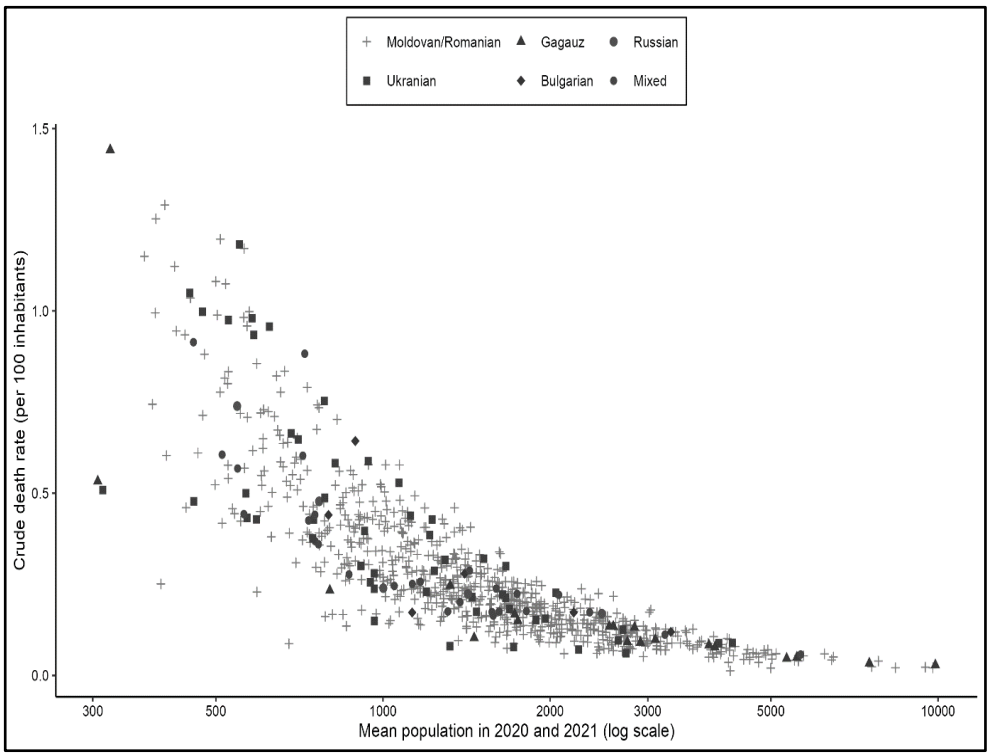
The description of the rural localities by ethnicities

Ethnicity	Moldovan/ Romanian	Russian	Gagauz	Ukrainian	Bulgarian	Mixed	Total
Number of communes with 50% ethnic group majority	733	5	19	51	6	27	841
Number of people in the communes with 50% ethnic group majority	1 316 268	6 215	60 594	65 213	9 729	38 091	1 496 109
Mean crude death rate	0.271	0.40	0.204	0.413	0.305	0.334	0.281
Number of communes with 75% ethnic group majority	698	1	15	20	3	104	841
Number of people in the communes with 75% ethnic group majority	1 268 691	1001	57 183	27 782	4 407	137 044	1 496 109
Mean crude death rate	0.268	0.24	0.104	0.333	0.298	0.382	0.281

Note: The commune is classified as mixed when none of the ethnic groups outweigh 50% and 74% of the population, respectively.
Source: National Bureau of Statistics and 2014 Population and Housing Census.

Figure 3

The distribution of the dependent variable by the mean population and dominant ethnicity



Source: National Agency for Public Health, National Bureau of Statistics, 2014 Population and Housing Census.

The literature review highlighted several factors that affected the chances of contracting COVID-19 and dying from it. We use the variables that correlate with differing behaviour, specifically, the share of migrating people, the share of people with secondary education only, and the share of households with computers. We also used the variable that correlates with the course of the disease: the share of households without heating and the share of solitary people. Since we expected the population structure to affect the deaths from COVID-19, we also included the share of the population aged 65+. We chose this cut-off point to find the margin for 90% of deaths due to COVID-19. The 65+ age group was the closest for both genders (taken separately). The descriptive statistics are provided in *Table no. 2.*

Table no. 2

Descriptive Statistics for the Variables Used

Variable	Number of zeroes	Min	Max	Median	Mean	St. deviation
Crude death rate	0	0.013	1.442	0.22	0.281	0.217
Excess mortality	0	-77.2	160	6.45	7.61	21.41
Mean commune size in 2020−2021	0	306	9891	1454	1779	1228
People who have not migrated two years prior	0	0.916	1	0.987	0.985	0.009
Households with computers	0	0.019	0.208	0.098	0.099	0.027
Households with one person only	0	0.011	0.22	0.072	0.078	0.032
Households without heating	511	0	0.121	0.001	0.004	0.01
Old people	0	0	0.894	0.006	0.072	0.179
Secondary education	0	0	0.935	0.005	0.018	0.062
Spoken Romanian	2	0	0.982	0.001	0.027	0.137
Spoken Ukrainian	346	0	0.854	0	0.017	0.081
Spoken Russian	14	0	0.927	0.001	0.05	0.158
Spoken Gagauz	741	0	0.988	0.01	0.072	0.166
Spoken Bulgarian[3]	732	0	0.929	0	0.017	0.115
Romanian mother tongue	728	0	0.862	0	0.009	0.069
Ukrainian mother tongue	46	0	0.901	0.004	0.058	0.162
Russian mother tongue	16	0	0.956	0.007	0.044	0.108
Gagauz mother tongue	473	0	0.976	0	0.024	0.132
Bulgarian mother tongue	486	0	0.881	0	0.014	0.077
Ethnic Romanians	0	0.005	0.999	0.975	0.857	0.259
Ethnic Ukrainians	27	0	0.998	0.006	0.071	0.178
Ethnic Russians	31	0	0.999	0.005	0.018	0.062
Ethnic Gagauz	383	0	0.982	0	0.016	0.081
Ethnic Bulgarians	405	0	0.854	0	0.026	0.14

Note: Every variable except death rate and mean commune size is the respective share.
Source: National Agency for Public Health, National Bureau of Statistics and 2014 Population and Housing Census.

[3] As mentioned, the Bulgarians (ethnicity) were only used as a dummy variable category. The data on the Bulgarians in this table are provided for general information.

RESULTS

This section is organised as follows. First, we show the heterogeneities of the rural Moldovans' trust in the government and the vaccines with the data of the Public Opinion (*Tables no. 3* and *no. 4*). Second, we make a regression analysis of the crude death rate calculated as shown in the previous section. Third, we conduct a sensitivity analysis and the robustness analysis changing both the dependent variable and the independent variable as shown in the Data and Methods section. Finally, we discuss our results in the context of the country.

Exploratory analysis

We conducted a t-test on the proportions of the two samples. We found significant differences between non-Romanian- and Romanian-speakers for the first and the third questions but not the second one (*Table no. 3*). The results change their statistical power depending on the addition of the bilinguals, but the p-value stays between 0.0089 and 0.048. The first question demonstrates substantial heterogeneities in trust in the government between the Romanian and non-Romanian speakers, which is explained by the partisan politics. The poll was conducted in February 2021, before the parliamentary elections and the defeat of the Socialist pro-minorities' government. While the differences in the views on the vaccine efficiency are also both substantial and significant, they do not yield a significant difference in the vaccination intentions.

Table no. 3

The public opinion on COVID-19 in February 2021 (share of positive answers)

	Sample size	How do you evaluate the activity of the Moldovan leadership in managing the COVID-19 pandemic?	Will you personally get vaccinated?	Do you think vaccination will help address the pandemic?
Total	588	29.4%	32.5%	25.3%
Romanian-speakers	452	25.9%	31.4%	22.6%
Romanian-speakers + bilinguals	517	28.0%	31.9%	23.6%
Non-Romanian-speakers + bilinguals	135	41.2%	36.0%	34.6%
Non-Romanian-speakers	71	39.4%	36.6%	38.0%

Source: Public Opinion Barometer.

In the second round of the Public Opinion Barometer survey, when the vaccination became available for all population strata, the respondents were asked

about their opinion on vaccination and their vaccination status (*Table no. 4*). The results emphasise a higher attendance among Moldovans/Romanians, compared to other ethnicities, while regarding vaccination intentions, the proportion of Moldovans/Romanians is lower. The difference in vaccination rates cannot be explained by the differences in vaccination capacities, considering the significant number of unused and expired vaccination doses (Radio Europa Liberă 2023).

The increase in vaccination attendance could result from factors other than protection against coronavirus disease, e.g. travel and lockdown restrictions. However, the proportion of those who are not intending to vaccinate is high for all ethnicities.

Table no. 4

The public opinion on vaccination against COVID-19 in June 2021

Ethnicity	Vaccinated	Planning to get vaccinated	Do not want to get vaccinated	NA
Moldovan/Romanian	28.2%	23.1%	46.5%	2.2%
Other	14.6%	31.5%	50.8%	3.1%

Source: Public Opinion Barometer.

Regression analysis

To assess the efficiency of the language and ethnicity variables on the overall data, we first did the regression analysis without the control variables and then added those. In rural areas, the language and the ethnicities intersect to a great extent. For example, the correlation between the ethnically Romanians and those with the Romanian native language is 0.991, and the correlation between the Romanian native language and the Romanian spoken language is 0.993. In other words, their simultaneous usage could yield a severe multicollinearity effect. This can be observed in *Table no. 5* (below), although at the end the language shows itself as a more robust independent variable (the native tongue shows analogous results).

Table no. 5

Regressions with the main independent variables only and crude COVID-19 death rate as the dependent variable

	(1)	(2)	(3)
Intercept	0.349 ***	0.348 ***	0.311 ***
	(0.026)	(0.022)	(0.030)
Romanians (ethnicity)	-0.079 **		0.269
	(0.029)		(0.150)
Romanians (spoken language)		-0.080 **	-0.314 *
		(0.025)	(0.132)

N	841	841	841
R2	0.009	0.012	0.016

Note: *** p < 0.001;** p < 0.01;* p < 0.05.
Source: National Agency for Public Health, National Bureau of Statistics and 2014 Population and Housing Census.

Table no. 6

Crude COVID-19 Death Rate Models

Main variable type	**Spoken language**	**Spoken language**	**Mother tongue**	**Mother tongue**	**Ethnicity**	**Ethnicity**	**Ethnic 75% majority**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Moldovans / Romanians	-0.061*		-0.063*		-0.059*		Base category
	(0.024)		(0.026)		(0.028)		
Russians		0.131**		0.135		0.079	0.034
		(0.044)		(0.069)		(0.114)	(0.199)
Ukrainians		0.023		0.025		0.037	-0.027
		(0.048)		(0.048)		(0.042)	(0.046)
Gagauz		-0.056		0.005		0.036	-0.057
		(0.062)		(0.054)		(0.052)	(0.053)
Bulgarians							0.118
							(0.116)
Mixed							0.084***
							(0.022)
People aged 65+	0.089	0.135	0.088	0.134	0.095	0.138	0.102
	(0.231)	(0.230)	(0.232)	(0.231)	(0.232)	(0.231)	(0.232)
People have not migrated in the last two years	-2.38**	-2.011*	-2.393**	-2.174*	-2.406**	-2.364**	-1.948*
	(0.848)	(0.859)	(0.848)	(0.863)	(0.849)	(0.861)	(0.856)
Households with a computer	-1.427***	-1.41***	-1.427***	-1.413***	-1.427***	-1.421***	-1.36***
	(0.297)	(0.297)	(0.297)	(0.298)	(0.297)	(0.299)	(0.297)
Households with one person only	1.879***	1.788***	1.885***	1.814***	1.888***	1.851***	1.826***
	(0.344)	(0.355)	(0.344)	(0.356)	(0.344)	(0.356)	(0.351)
Houses without heating	-0.944	-1.063	-0.924	-0.954	-0.886	-0.806	-1.056
	(0.741)	(0.743)	(0.741)	(0.745)	(0.740)	(0.743)	(0.737)
Secondary education	0.089	0.096	0.088	0.093	0.088	0.084	0.085
	(0.115)	(0.115)	(0.115)	(0.116)	(0.115)	(0.116)	(0.115)
Intercept	2.615**	2.182**	2.630**	2.345**	2.639**	2.534**	2.126*
	(0.831)	(0.841)	(0.831)	(0.845)	(0.833)	(0.842)	(0.838)
N	841	841	841	841	841	841	841
R^2	0.166	0.170	0.166	0.165	0.164	0.162	0.178

Note: *** p < 0.001; ** p < 0.01; * p < 0.05.
Source: National Agency for Public Health, National Bureau of Statistics and 2014 Population and Housing Census.

According to the results, among the large ethno-linguistic groups, the Moldovans/Romanians had the lowest crude death rates in 2020–2021. The effect is sustained if the share of Moldovans/Romanians is measured by their mother tongue (rather than spoken language) or their ethnicity. The Russian speakers, on the other hand, have significantly higher death rates, if measured by the spoken language. When measured by the mother tongue or ethnicity, the share of the Russian speakers (or the Russians) loses its significance. Combined with the robustness of the effect of the share of the Moldovans/Romanians, these findings allow us to reject the hypothesis that ethnic or cultural characteristics increase COVID-19 vulnerability in Moldova. The hypothesis on the language had three parts: one based on the interactions with the state institutions, one based on the identity, and one based on the COVID-19 misinformation propaganda campaigns (which were more common in Russian than in Romanian). We would place a higher weight on the latter part. The Russian speakers are the only group that showed a significant positive relation with the death rates. Such a relation was not detected for the Ukrainians and is actually (non-statistically significant) negative for the Gagauz.

It may also suggest that there is a group of Russian speakers from non-Russian ethnicities that are associated with higher COVID-19 crude mortality rates, which may not be detected because of the sample size. If the binary variable is used for the measurement of ethnicity, the communes that do not have a 75% ethnic majority have significantly more people dying in 2020–2021. The aforementioned group can be in these mixed communes, which comprise several small villages with populations of different ethnolinguistic characteristics.

While the results show that the effect of language is statistically significant, it should be checked if the effect is socially substantial. There are villages with overwhelming domination of the Romanian or Russian language, so the following interpretation can be suggested according to regressions #2 and #3. All things equal, a fully Romanian-speaking village, on average, had 6 fewer deaths per 10000 people compared to a fully non-Romanian village. Similarly, a fully Russian-speaking village had, on average, 13 more deaths per 10000 people compared to a village with no Russian-, Ukrainian-, or Gagauz-speakers[4]. Most villages (585 out of 841) had between 500 and 2000 inhabitants, which further scales down such effects. The Romanian-speaking communes are also larger than the Russian-speaking ones, on average (1806 vs. 1207, accordingly), so the central estimates for the average-size village are, respectively, 1.08 and 1.56.

Finally, the control variables that we considered had a varying effect. The share of older people, the heating and the proportion of people with only secondary education were not significant in any of the regressions. The share of households

[4] The overwhelming majority of such data points are Romanian-speaking villages, but they can technically be composed of Bulgarian-speakers, for example.

with a computer, the share of households consisting of one person, and the share of the people that did not change their place of living in the two years before the census are all statistically significant. While the coefficients per se are higher than those of the main variables, their variability is also smaller. For example, only 16 villages have a share of the people who have not changed their place of living less than 96%. A shift in this variable from 96% to 100% results in 9 fewer deaths per 10000 in these two years.

Sensitivity Analysis

In the spirit of Treiman (2009), we establish the robustness of our results, we conduct two-step sensitivity analysis, specifically, the data sensitivity analysis and regression specification sensitivity analysis.

In the first step, we attempt to find if there is a pattern among the outliers and influential observations. They were detected with two techniques: the LOOCV made with the *caret* package (Kuhn 2008) and Cook's distance. The former is done as follows: an observation is deleted from the sample one at a time, and the regression is run once again. Then, the prediction is done for the observation that was left out from the calculations and the difference between the real value of the observation and its predicted value is calculated. Two metrics are calculated from this: RMSE, which is the root of the mean standard error, and MAE, which is the mean absolute error. If the errors are relatively uniform, the RMSE and the MAE must be the same. In our case, for each of the regressions offered in the previous section, RMSE was 0.2 and MAE was 0.14. In other words, there are indeed strong outliers in the dataset, and on average the model estimates the number of Covid deaths with an average error of 14 individuals per 10000. When reviewing the residuals, we found that the model underestimated the number of deaths. With Cook's distance, we detected 51 (out of 841) influential observations. The analysis of such influential observations is out of the scope of this paper. However, there was a pattern of the mixed villages (defined as the lack of 75% domination of an ethnicity in the village) being overrepresented among the underestimated outliers. The pattern of mixed villages overrepresentation among the outliers also existed in Model 7 that used dummy variables for minority groups.

In *Table no. 7*, the results for the additional regressions are shown. These regressions used the excess mortality as the dependent variable. Its additional independent variable is the commune's size in 2020–2021. We added this variable to reflect the relation observed in *Figure 3* in the Data and Methods. The relation could not have been accounted for in the previous regressions set since the average population size is used to calculate the crude covid death rate.

Table no. 7

Excess 2020−2021 Mortality Models

	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Moldovans /Romanians	-0.073 **		-0.088 **		-0.092 **		Reference group
	(0.026)		(0.028)		(0.029)		
Russians		0.154 **		0.205 **		0.139	9.317
		(0.047)		(0.073)		(0.119)	(4.891)
Ukrainians		0.019		0.047		0.099 *	-18.272
		(0.051)		(0.050)		(0.044)	(21.048)
Gagauz		0.026		0.061		0.076	5.007
		(0.065)		(0.057)		(0.054)	(5.580)
Bulgarians							-2.621
							(12.294)
Mixed							7.357 **
							(2.294)
Intercept	135.945	100.584	140.323	108.142	143.288	129.770	117.574
	(86.905)	(88.020)	(86.859)	(88.081)	(86.911)	(87.772)	(87.921)
People aged 65+	0.300	0.312	0.280	0.303	0.274	0.311	0.293
	(0.242)	(0.241)	(0.242)	(0.241)	(0.243)	(0.241)	(0.244)
Commune size (in 100s)	0.080	0.097	0.079	0.087	0.078	0.073	0.087
	(0.066)	(0.067)	(0.066)	(0.067)	(0.066)	(0.067)	(0.068)
People have not migrated in the last two years	-1.423	-1.143	-1.448	-1.213	-1.472	-1.431	-1.312
	(0.875)	(0.887)	(0.874)	(0.888)	(0.874)	(0.885)	(0.886)
Households with a computer	1.050 ***	1.021 **	1.047 ***	1.037 **	1.045 ***	1.074 ***	1.114 ***
	(0.317)	(0.317)	(0.316)	(0.317)	(0.316)	(0.317)	(0.317)
Households with one person only	0.049	0.063	0.058	0.012	0.065	-0.038	-0.002
	(0.367)	(0.378)	(0.367)	(0.378)	(0.367)	(0.377)	(0.375)
Houses without heating	0.108	-0.057	0.079	-0.041	0.095	0.107	0.077
	(0.778)	(0.780)	(0.777)	(0.779)	(0.776)	(0.777)	(0.776)
Secondary education	-1.055	-1.001	-1.090	-1.029	-1.092	-1.025	-1.042
	(0.654)	(0.654)	(0.653)	(0.654)	(0.653)	(0.655)	(0.655)
N	841	841	841	841	841	841	841
R^2	0.054	0.060	0.057	0.060	0.057	0.056	0.062

Note: *** $p < 0.001$;** $p < 0.01$; * $p < 0.05$.
Source: National Agency for Public Health, National Bureau of Statistics and 2014 Population and Housing Census.

The regressions with the excess mortality as the dependent variable repeat the main results we discussed above, i.e., the lower mortality in the Romanian-speaking communes and higher mortality in the Russian-speaking communes. It also showed higher mortality in the ethnically mixed communes. However, unlike

the previous models, Model 13 suggests that the excess mortality positively correlates with the share of the ethnically Ukrainian rural residents. These results are rather in a disagreement with Model 14, which suggests that the ethnically mixed communes suffer from excess mortality rather than predominantly Ukrainian ones.

Brief results overview and research limitations

Even though we are not able to establish the causation of the emphasised differentiation in COVID-19 mortality between the population ethnolinguistic groups among rural localities in Moldova, we speculatively explain our results as being influenced by joint factors. Firstly, external misleading information on the COVID-19 pandemic (which was more common in the Russian language) might be contributing to an increase in the prevalence of hesitancy in vaccination against coronavirus disease. Secondly, we are aware of the potential higher institutional mistrust among ethnic and linguistic minorities, which might challenge health interventions. And lastly, central and local authorities might not be capable of communicating efficiently with ethnic minorities during the crises. We also do not exclude that villages with a larger population have more sizeable budgets for the local medical infrastructure, implementing local initiatives, and supporting vulnerable population strata. Moreover, the discrepancies in COVID-19 mortality between ethnolinguistic groups might be underlined due to socioeconomic and educational characteristics, which, at this point, haven't been studied.

In regard to the health system, Moldova has a solid infrastructure of hospitals and health facilities, which is relatively accessible for the entire population[5]. Even though most specialised hospitals are established in the capital city, every town has an operating hospital, while villages with a larger population have Centres of health and family doctors. However, the rural population, especially in the remote villages, is disadvantaged regarding the health system accessibility and has a higher mortality rate compared to the urban population (Penina 2022).

Understanding the overall influence of health system accessibility on COVID-19 mortality outcome, at this point, we are unable to estimate the inequalities in health care (prevention and treatment) that different populations faced. This concern might be the subject of future studies.

The presented research has the following essential limitations. Due to a lack of available data, we did not include in our research most of the localities on the left bank of the Nistru River, which are predominantly exposed to Russian media broadcasts. Additionally, we were limited in data on the population regarding media consumption, institutional trust, hesitancy and refusal to be vaccinated

[5] Except for most of the localities on the left bank of the Nistru River, which are controlled by unconstitutional authorities. However, in emergency cases, the patients from the Transnistrian region could be transferred to the health facilities on the right bank of the Nistru River.

against COVID-19 etc. Further, we might underestimate the effect of bilingualism since a significant proportion of the population can receive information from different sources and broadcasting services.

CONCLUSIONS

Naturally, people consume media and interact within the community in the mother or usually spoken language. Considering the country's size and multi-ethnic population structure, media consumption in Moldova is highly influenced by centralised broadcastings from Romania, Russia, Ukraine, and other countries. Thus, in the presented research, based on the hypothesis that people respond to the COVID-19 crisis depending on the information they have available in their usually spoken language, differences in mortality from COVID-19 disease and other chronic diseases during the COVID-19 pandemic might be emphasised among the population by ethnicity, mother tongue, and usually language spoken.

The results emphasise a statistically significant higher mortality rate among Russian speakers compared to the Romanian-speaking population in the rural areas. Among the control variables, such as households with a computer, the share of households consisting of one person, and the share of the people that did not change their place of living in the two years before the census are all statistically significant. The proportion of people with secondary education as a control variable was insignificant in any regressions. We speculatively explain the differences in COVID-19 mortality rate among ethnolinguistic population groups in rural Moldova by a joint influence of external propaganda, which led to a higher hesitancy of vaccination among some population strata, a higher institutional mistrust among the ethnic minorities in Moldova and a lack of institutional capacity to communicate efficiently with the ethnic minorities.

The research has several limitations due to the lack of data availability at the village level, especially concerning the consumption of information from media and social media, attendance, refusal and hesitancy in vaccination against COVID-19, social and institutional trust etc. Additionally, the study does not include villages on the left bank of the Nistru River, which are not under the control of the constitutional authorities. The effect of exposure to COVID-19 misinformation on Romanian speakers could have been underestimated since a significant proportion of the population in Moldova is bilingual. The findings of the presented research could serve as evidence for central and local authorities to implement targeted health policies to diminish the differences in mortality levels of the population by socioeconomic and ethnic-linguistic characteristics.

Acknowledgement

The article was developed within the framework of the research programme 030102 "The demographic transition in the Republic of Moldova: particularities, socioeconomic implications and demographic resilience strengthening (TDRM, 2024−2027)".

REFERENCES

Bayati, Mohsen, RayeheNoroozi, Mohadeseh Ghanbari-Jahromi, and Faride Sadat Jalali. 2022. "Inequality in the distribution of Covid-19 vaccine: a systematic review." *Int J Equity Health* 21 (122). doi:10.1186/s12939-022-01729-x.

Bonnet, Florian, Pavel Grigoriev, Markus Sauerberg, Ina. Alliger, Michael Mühlichen, and Carlo-Giovanni Camarda. 2024. "Spatial disparities in the mortality burden of the COVID-19 pandemic across 569 European regions (2020−2021)." *Nature Communications* 15 (4246). doi:10.1038/s41467-024-48689-0.

Chang, Dianna, Xin Chang, Yu He, and Kelvin Jui Keng Tan. 2022. "The determinants of COVID-19 morbidity and mortality across countries." *Sci Rep* 12 (5888). doi:10.1038/s41598-022-09783-9.

Costa-Font, Joan, Jorge Garcia-Hombrados, and Anna Nicińska. 2023. "The institutional origins of vaccines distrust: Evidence from former-Soviet countries." *PLoS ONE* 18 (3). doi:10.1371/journal.pone.0282420.

Dukhovnov, Denys, and Magali Barbieri. 2022. "County-level socio-economic disparities in COVID-19 mortality in the USA." *International Journal of Epidemiology* 51 (2): 418–428. doi:10.1093/ije/dyab267.

Foster, Hamish M.E., Frederick K. Ho, Frances S. Mair, Bhautesh D. Jani, Naveed Sattar, Srinivasa Vittal Katikireddi, Jill P. Pell *et al.* 2022. "The association between a lifestyle score, socioeconomic status, and COVID-19 outcomes within the UK Biobank cohort." *BMC Infectious Diseases* 22 (273). doi:10.1186/s12879-022-07132-9.

Foster, Noel 2021. "Propaganda Gone Viral: A Theory of Chinese and Russian "COVID Diplomacy" in the Age of Social Media." In *The Russian Federation in Global Knowledge Warfare*, by H. Mölder, V. Sazonov, A. Chochia and T. Kerikmäe, 123–145. doi:10.1007/978-3-030-73955-3.

Galobardes, Bruna, Mary Shaw, Debbie A. Lawlor, John W. Lynch, and George Davey Smith. 2006. "Indicators of socioeconomic position (part 1)." *Journal of Epidemiology & Community Health* 60 (1): 7–12.

Gillies, Clare L., Alex V. Rowlands, Cameron Razieh, VahéNafilyan, Yogini Chudasama, Nazrul Islam, Francesco Zaccardi, *et al.* 2022. "Association between household size and COVID-19: A UK Biobank observational study." *Journal of the Royal Society of Medicine* 138-144. doi:10.1177/01410768211073923.

Hoxha, Ilir, Riaz Agahi, Altina Bimbashi, Mrika Aliu, Lul Raka, Ilirjana Bajraktari, Petrit Beqiri, and Lisa V. Adams. 2022. "Higher COVID-19 Vaccination Rates Are Associated with Lower COVID-19 Mortality: A Global Analysis." *Vaccines* 11 (74). doi:10.3390/vaccines11010074.

Institutul de OpiniiPublice. 2021. "Barometer of Public Opinion (Feb. 2021)." http://bop.ipp.md/.

Institutul de OpiniiPublice. 2021. "Barometer of Public Opinion (June 2021)." doi:https://ipp.md/wp-content/uploads/2021/07/Raport-BOP-2021_.pdf.

Keegan, Katrina M. 2022. "Clarity for friends, confusion for foes: Russian vaccine propaganda in Ukraine and Serbia." *Harvard Kennedy School (HKS) Misinformation Review* 3 (3).

Kuhn, Max. 2008. "Building Predictive Models in R Using the caret Package." *Journal of Statistical Software* 28 (5): 1-26. doi:10.18637/jss.v028.i05.

Kung, Stacey, Marjan Doppen, Melissa Black, Irene Braithwaite, Ciléin Kearns, Mark Weatherall, Richard Beasley, and Nethmi Kearns. 2021. "Underestimation of COVID-19 mortality during the pandemic." *ERJ Open Res* 7 (1). doi:10.1183/23120541.00766-2020.

Lee, Haena, Theresa Andrasfay, Alicia Riley, Qiao Wu, and Eileen Crimmins. 2022. "Do social determinants of health explain racial/ethnic disparities in COVID-19 infection?" *Riley, Qiao Wu, Eileen Crimmins* 306. doi:10.1016/j.socscimed.2022.115098.

Lupu, Dan, and Ramona Tiganasu. 2022. "COVID-19 and the efficiency of health systems in Europe." *Health Econ Rev* 12 (14). doi:10.1186/s13561-022-00358-y.

Magdin, Radu. 2020. "Disinformation Campaigns in the European Union: Lessons Learned from the 2019 European Elections and 2020 COVID-19 Infodemic in Romania." *Romanian Journal of European Affairs* 20 (2): 49–61.

Marsili, Marco. 2021. "The Russian Influence Strategy in Its Contested Neighbourhood." In *The Russian Federation in Global Knowledge Warfare. Contributions to International Relations*, by H. Mölder, V. Sazonov, A. Chochia and T. Kerikmäe. Springer.

NAPH. n.d. Ministry of Health of the Republic of Moldova. National Agency for Public Health.

NBS. 2014. *2014 Population and Housing Census.* https://recensamant.statistica.md.

NBS. n.d. *Number of population with usual residence, by sex, at the level of first-level administrative-territorial unit.* National Bureau of Statistics.

Pascoal, Rui, and Humberto Rocha. 2022. "Population density impact on COVID-19 mortality rate: A multifractal analysis using French data." *Physica A: Statistical Mechanics and its Applications* 593. doi:10.1016/j.physa.2022.126979.

Pathak, Elizabeth B., Janelle M. Menard, Rebecca B. Garcia, and Jason L. Salemi. 2022. "Joint Effects of Socioeconomic Position, Race/Ethnicity, and Gender on COVID-19 Mortality among Working-Age Adults in the United States." *Int. J. Environ. Res. Public Health* 19. doi:10.3390/ijerph19095479.

Paul, Elise, Andrew Steptoe, and Daisy Fancourt. 2021. "Attitudes towards vaccines and intention to vaccinate against COVID-19: Implications for public health communications." *The Lancet Regional Health − Europe.*doi:10.1016/j.lanepe.2020.100012.

Penina, Olga. 2022. "Rural−urban disparities of mortality and causes of death in the Republic of Moldova." *Archives of the Balkan Medical Union* 57 (3): 222−230. doi:10.31688/ABMU.2022.57.3.01.

Pizzato, Margherita, Alberto Giovanni Gerli, Carlo La Vecchia, and Gianfranco Alicandro. 2024. "Impact of COVID-19 on total excess mortality and geographic disparities in Europe, 2020–2023: a spatio-temporal analysis." *The Lancet Regional Health* 44. doi:10.1016/j.lanepe.2024.100996.

Radio Europa Liberă. 2023. *Republica Moldova a pierdut peste jumătate de milion de vaccinuriîmpotriva Covid-19.*13 06. https://moldova.europalibera.org/a/r-moldova-a-pierdut-peste-jumatate-de-milion-de-vaccinuri-impotriva-covid-19-/32455876.html.

Schöley, Jonas, José ManuelAburto, Ilya Kashnitsky, Maxi S. Kniffka, Luyin Zhang, Hannaliis Jaadla, Jennifer B. Dowd, and Ridhi Kashyap. 2022. "Life expectancy changes since COVID-19." *Nature Human Behaviour* 6: 1649–1659. doi:10.1038/s41562-022-01450-3.

Sylvester, Steven M. 2021. "COVID-19 and Motivated Reasoning: The Influence of Knowledge on COVID-Related Policy and Health Behavior." *Social Science Quarterly.* doi:10.1111/ssqu.12989.

Toshkov, Dimiter. 2023. "What accounts for the variation in COVID-19 vaccine hesitancy in Eastern, Southern and Western Europe?" *Vaccine* 41 (20): 3178-3188. doi:10.1016/j.vaccine.2023.03.030.

Treiman, Donald J. 2009. *Quantitative Data Analysis: Doing Social Research to Test Ideas (1st ed.).* Edited by CA: Jossey-Bass.

Varshney, Karan, Talia Glodjo, and Jenna Adalbert. 2022. "Overcrowded housing increases risk for COVID-19 mortality: an ecological study." *BMC Research Notes* 15 (126). doi:10.1186/s13104-022-06015-1.

Wilson, Steven L., and Charles Wiysonge. 2020. "Social media and vaccine hesitancy." *BMJ Global Health.* doi:10.1136/bmjgh-2020-004206.

Zhuo, Jie, and Nicholas Harrigan. 2023. "Low education predicts large increase in COVID-19 mortality: the role of collective culture and individual literacy." *Public Health* 201−207. doi:10.1016/j.puhe.2023.06.016.

Pandemia COVID-19 a evidențiat o discrepanță în nivelul mortalității dintre diferite grupuri ale populației în funcție de caracteristicile sociale, economice și demografice. Literatura științifică arată influența factorilor individuali și comportamentali asupra incidenței bolii coronavirus și riscului de deces. În același timp, nivelul mortalității prin COVID-19 este influențat de capacitatea instituțiilor statului de a răspunde eficient la provocările legate de sănătatea populației prin reducerea riscului de deces și evitarea sechelelor bolii. Această cercetare se bazează pe ipoteza că populația își ajustează comportamentul și răspunde la criza pandemică COVID-19 în funcție de informația disponibilă în limba maternă și cea vorbită de obicei, ceea ce duce la o discrepanță în mortalitatea prin boala coronavirus între populații după caracteristicile etnolingvistice. Astfel, prin utilizarea unui model de regresie liniară, am comparat nivelul mortalității cauzate de COVID-19 în rândul principalelor etnii din Moldova. Rezultatele obținute au arătat un nivel de mortalitate mai ridicat în rândul populației vorbitoare de limbă rusă, care poate fi explicat prin efectul sinergic al multiplilor factori, precum: dezinformarea privind situația pandemică în presa rusolingvă, care a contribuit la ezitarea în administrarea vaccinurilor produse în țările occidentale, gradul mai înalt de neîncredere instituțională în rândul minorităților etnice, dar și lipsa unei capacități instituționale de a comunica eficient cu minoritățile etnice. Rezultatele studiului pot servi autorităților centrale și locale în implementarea unor politici de sănătate care vizează diminuarea inegalităților în sănătate în rândul populațiilor în funcție de caracteristicile socioeconomice și etnolingvistice.

***Cuvinte-cheie**: inegalități în sănătatea populației; mortalitatea prin COVID-19; diferențe etnolingvistice în mortalitatea prin boala coronavirus.*

Received: 17.01.2025 *Accepted: 30.05.2025*