







RESEARCH ARTICLE

# Volga German surnames and Alzheimer's disease in Argentina: an epidemiological perspective

Arturo Leonardo Morales<sup>1,2,3,4†</sup> , Marcelo Isidro Figueroa<sup>5,6†</sup> , Pablo Navarro<sup>1,2,3,4</sup> ,  
Estela Raquel Chaves<sup>7</sup>, Anahí Ruderman<sup>1,4</sup> , José Edgardo Dipierri<sup>4,5</sup>  and Virginia Ramallo<sup>1,4</sup> 

<sup>1</sup>Instituto Patagónico de Ciencias Sociales y Humanas (IPCSH), Centro Nacional Patagónico, Consejo Nacional de Investigaciones Científicas y Técnicas, Puerto Madryn, Argentina, <sup>2</sup>Laboratorio de Ciencias de las Imágenes, Departamento de Ingeniería Eléctrica y Computadoras, Universidad Nacional del Sur, Bahía Blanca, Argentina, <sup>3</sup>Departamento de Informática, Facultad de Ingeniería, Universidad Nacional de la Patagonia San Juan Bosco, Trelew, Argentina, <sup>4</sup>Programa de Referencia y Biobanco Genómico de la Población Argentina (PoblAr), Buenos Aires, Argentina, <sup>5</sup>Unidad de Genética, Hospital Materno Infantil Doctor Hector Quintana, San Salvador de Jujuy, Argentina, <sup>6</sup>Instituto de Ecorregiones Andinas, Universidad Nacional de Jujuy-CONICET, San Salvador de Jujuy, Argentina and <sup>7</sup>Instituto de Biología de la Altura, Universidad Nacional de Jujuy, San Salvador de Jujuy, Argentina

**Corresponding author:** Virginia Ramallo; Email: [ramallo@cenpat-conicet.gob.ar](mailto:ramallo@cenpat-conicet.gob.ar)

(Received 12 January 2024; revised 24 February 2024; accepted 4 April 2024; first published online 29 April 2024)

## Abstract

The N141I variant (PSEN1 gene) is associated with familial forms of early-onset Alzheimer's disease (AD) in descendants of Volga Germans, whose migration to Argentina is well documented. As a proxy for geographic origin, surnames can be a valuable tool in population studies. The 2015 Argentine Electoral Registry provided geographic data for 30,530,194 individuals, including 326,922 with Volga German surnames. Between 2005 and 2017, the Ministry of Health recorded 4,115,216 deaths, of which 17,226 were attributed to AD and related causes. The study used both diachronic and synchronic data to identify patterns of territorial distribution and co-spatiality, using Moran's I and generalised linear model statistics. The frequency of surnames of Volga German origin accounts for 43.53% of the variation in deaths from AD and three clusters of high non-random frequency were found. Almost 150 years later, people descending from the Volga migration remain highly concentrated and may have a different risk of developing AD. The identification of spatial patterns provides reliable guidance for medical research and highlights the importance of specific health policies for particular populations.

**Keywords:** surname analyses; epidemiology; Volga migration-descended; Alzheimer's disease

## Introduction

### Migration and health

There is a well-established link between population mobility, migration, and the epidemiology of certain diseases (Gushulak and MacPherson, 2006). While communicable diseases have traditionally been the focus of attention, the increasing global importance of migration has led to a renewed interest in other aspects of population health, in particular non-communicable diseases that are linked to genetic characteristics.

By combining various analytical methods, available data on Alzheimer's disease (AD) can be used to identify determinants of health disparities in AD and their impact at the individual,

<sup>†</sup>Arturo Leonardo Morales and Marcelo Isidro Figueroa contributed equally to this manuscript.

community, and societal levels (Akushevich *et al.*, 2023). The determinants may include ethnic, gender, and geographic factors.

### **Alzheimer's disease and Volga Germans**

AD can be either familial or sporadic. The familial form is autosomal dominant and has an early onset, occurring in individuals under 65 years of age. It accounts for only 1–5% of cases and is caused by mutations in the genes PSEN1 (presenilin 1) (MIM \*104311) and PSEN2 (presenilin 2) (MIM \*600759). PSEN1 is associated with Alzheimer's disease type 3 (MIM 607822) and PSEN2 with AD type 4 (MIM 606889). Late-onset AD is classified as occurring in patients over the age of 65. The sporadic form, which accounts for 95% of cases, has no known genetic cause (Andrade-Guerrero *et al.*, 2023).

In 1988, Bird *et al.* reported on presenile AD in five families in the United States. The disease was confirmed by autopsy and inherited as an autosomal dominant trait, present in both males and females for several generations (Bird *et al.*, 1988). All five families were descended from immigrants known as Volga Germans who arrived in the United States between 1870 and 1920. In 1992, the authors studied 28 families of Volga German descendants with AD. Eighteen families originated from the villages of Frank and Walter, near Saratov, Russia. The authors suggested that the cause of AD may be a founder effect. However, a common affected ancestor could not be identified (Bird *et al.*, 1992).

### **The diaspora of Volga Germans**

To understand the context of the migration of Germans to Russia, it is necessary to consider the Seven Years' War. This international conflict over the control of colonies in North America and India began in 1756 and ended in 1763. It involved several empires, kingdoms, and other political structures established in territories that are now nation-states, including Germany. Catherine II, the Empress of the Russian Empire, aimed to populate the border regions near Asia with settled European peasants. In 1763, she promoted a series of benefits for migrants to make the colonisation of the Volga attractive to the post-war weakened German population. These benefits included freedom of religion, temporary tax exemption, interest-free loans, internal self-government, and permanent exemption from military conscription. Between 1764 and 1769, settlers primarily from Hesse, a state in Germany with its current capital in Wiesbaden, arrived on the lower Volga near the city of Saratov. During this time, 104 colonies were established, with a total population of 22,246 inhabitants (Pohl, 2009).

In 1864, Tsar Alexander II made amendments to the original agreement. By 1874, German immigrants were required to register for military service. This event marked the beginning of the diaspora, particularly in the United States, Canada, Australia, and Brazil (Pohl, 2009). In 1878, 1,100 Volga Germans arrived in Argentina. They settled in the provinces of Entre Ríos, Buenos Aires, Santa Fe, Chaco, and Córdoba, where numerous rural colonies were founded that still exist today. Immigration to Argentina continued until the First World War (AufderHeide, 2006). Estimates from descendants' associations suggest that there are 2,000,000 Volga Germans in the country. Individual migration stories are well documented and accessible through the websites of the descendant societies and/or their social networks, including Alemanes del Volga en Argentina (<http://www.alemanesdelvolga.com.ar>), Centro Cultural Argentino Wolgadeutsche, Federación Argentina de Descendientes de Alemanes del Volga (<http://fadadav.org.ar>), and Asociación Argentina de Descendientes de Alemanes del Volga Unser Licht.

This database contains family registers, a list of surnames, and information on the colonies in which they settled.

### **Epidemiology of Alzheimer's disease in Argentina**

Research on AD has been concentrated on particular geographic regions and groups, as evidenced by studies conducted by Larraya *et al.* (2004), Melcon *et al.* (2010), Méndez *et al.* (2018), Itzcovich *et al.* (2020), García and Comesaña (2021), and Dalmaso *et al.* (2023). Currently, there is no national institutional registry of individuals diagnosed with AD. However, death certificates can serve as a valuable source of data. The Department of Health Statistics and Information (DEIS) is responsible for preparing periodic statistical reports that include an official registry of all deaths, their immediate causes, and any associated or pre-existing conditions. The International Classification of Diseases (ICD-10) is used to calculate morbidity and mortality statistics. The quality of this information is validated and serves as a critical input, although there may be variations between years or regions.

The aim of this study is to contribute to the epidemiology of AD by analysing the spatial distribution of all deaths related to AD, the spatial distribution of surnames of Volga German origin, and the association between these two phenomena in the country.

### **Materials and methods**

#### **Data sources**

Argentina, with a population of 46,654,581 according to the National Institute of Statistics and Censuses (INDEC, 2023), is the second largest country in South America. It is divided into 23 provinces, one federal district, and 529 minor subdivisions known as departments. The provinces are further grouped into five geographical regions (refer to Fig. 1). Northwest (comprising the provinces of Catamarca, Jujuy, La Rioja, Salta, Santiago del Estero, and Tucumán), Northeast (Corrientes, Chaco, Formosa, and Misiones), Cuyo (Mendoza, San Juan, and San Luis), Central or Pampean (Buenos Aires, Córdoba, Entre Ríos, La Pampa, Santa Fe, and the Autonomous City of Buenos Aires), and Patagonia (Río Negro, Neuquén, Chubut, Santa Cruz, and Tierra del Fuego). The Central/Pampean region has the highest population density and income levels. In the 19th and 20th centuries, it was the most common destination for transcontinental migrations.

Since 2012, all Argentine citizens over the age of 16 have the right to vote, and the electoral roll is compiled annually. For this study, the 2015 Electoral Registry and national death records from 2005 to 2017 were used, as they share the same territorial data organisation. We analysed publicly available anonymous information. According to Argentine legislation, ethical approval was not required in these cases.

#### **The Surnames of the Volga Germans**

Surnames are sociocultural variables that result from historical and cultural processes. They are an important resource in bioanthropology and human population genetics. The linguistic and/or geographic origin of surnames can serve as a proxy for ethnicity in demographic and spatial analyses (Mateos, 2014; Albeck *et al.*, 2017). A list of surnames was compiled using the information available on the websites of the aforementioned associations of Volga German descendants. The study compared the 50 most common Volga surnames (VSs) in Argentina to the 50 most common surnames in the State of Hesse, Germany, as provided by Forebears (<https://forebears.io/germany/hesse#surnames>), and the 50 most common German surnames (excluding the State of Hesse) based on telephone user records of 30,000,000 individuals used by Rodríguez-Larralde *et al.* (1978). To ensure accuracy, a conservative approach was adopted for the comparative analysis, due to the potential for incorrect surname recording during migration in the 19th century. The threshold for homonymy was set at the difference of a diacritical mark or accent, such as a dieresis.



Figure 1. Map of Argentina’s major administrative divisions and regions.

Using the Bulsarapp application (Morales *et al.*, 2021), surnames of Volga origin found in the 2015 electoral register were georeferenced. The frequency of these surnames was calculated by determining the number of individuals with VSs per 1000 voters ( $VS \times 1000$ ) at the departmental, provincial, and regional levels.

### Deaths due to Alzheimer’s disease

Although the death certificate is an official document signed by a medical professional, it may contain inaccuracies if data on the general state of health prior to death is not well known or would require forensic examinations to be conclusive. Therefore, we have used a broad criterion to select the following ICD 10 codes: G30.0 (AD with early onset, usually before the age of 65), G30.1 (AD with late onset, usually after the age of 65), G30.8 (other AD), G30.9 (Alzheimer disease, unspecified), and G31 (other degenerative diseases of nervous system, not elsewhere classified).

The DEIS (2023) provided the time series of the number of deaths reported in Argentina from 2005 to 2017. Specific death rates (SDRs) related to AD were calculated per 1000 deaths ( $AD \times 1000$ ) at departmental, provincial, and regional levels for the entire period.

### Statistical and spatial analysis

To assess the impact of VS on SDRs, we used a generalised linear model with mixed effects. We employed three distribution types: Poisson (Mod. Poiss), Negative Binomial (Mod. BN), and zero-inflated Poisson (Mod. ZIP). We diagnosed the goodness of fit using the Q-Q statistic and assessed the model's relative quality using the Akaike information criterion (AIC). We considered the political division (departments within provinces) as a random term. The study calculated the incidence rate ratio by exponentiating the regression beta. Furthermore, the Nagelkerke coefficient of determination was used to assess the proportion of the model's variance explained.

The study utilised the Global Moran Index and Local Indicator of Spatial Autocorrelation to analyse the spatial distribution of VSs and SDRs, as well as the combination of both variables. Spatial autocorrelation is a statistical method used to determine whether a group of entities, such as regions, provinces, and departments, and their attributes, such as VS and SDR values, exhibit clustered, sparse, or random patterns across a territory. The analysis can aid in identifying spatial relationships and patterns that may not be immediately apparent (Anselin, 1995). Significance was determined at the 0.05 confidence level using the Monte Carlo test (999 permutations) under the null hypothesis of no spatial association. The analyses were conducted using GeoDa 1.14 software and the GeoPandas and PySAL function libraries for the Python programming language.

### Results

Table 1 compares the 50 most common VSs in Argentina, Hesse (HS), and Germany (GS). Argentina and HS share 24% of the same surnames, while only 14% of the family names are common to all three lists.

According to the National Institute of Statistics and Censuses (INDEC, 2023), Argentina's total population in 2015 was 43,131,966. The electoral roll for that year had 30,530,194 registered voters with 373,709 different surnames, representing over 70% of the country's population. Out of these, 326,922 individuals (1.22%) had a VS, which accounts for a total of 1,109 surnames with this origin.

Table 2 displays the geographic distribution of VS carriers. The province of La Pampa, located in the Central region, had the highest frequency ( $VS \times 1000$ ), while San Juan (Cuyo region) had the lowest. Additionally, Table 2 summarises death certificate data, which shows that between 2005 and 2017, there were 4,115,216 recorded deaths in Argentina, with 17,226 (4.19%) attributed to or associated with AD. At the provincial level, La Pampa had the highest SDR, while the province of Santa Cruz had the lowest. Of all deaths related to Alzheimer's, 68% were women and 31% were men. In the remaining cases, the biological sex of the deceased could not be determined. The most common code in the database was G30.9 (AD, unspecified), appearing in 85% of certificates. For death certificates with code G30.0 (early-onset AD), the mean age was 73.82 for women (SD 14.89) and 68.23 for men (SD 12.01).

All spatial analyses indicate positive autocorrelation with  $p$ -values below 0.001. Figure 2 displays choropleth maps that clearly visualise the variability in VS and SDR by department. Several departments in the Central-Pampean region have the highest concentration of voters with surnames of Volga origin (see Fig. 2a). The maps show notably low SDR values in the departments of the NWA, NEA, and southern regions of the country (Fig. 2b). After calculating the bivariate Moran's I for the two data sources (VS and SDR), three clusters with high and non-random values (Moran's  $I = 0.19$ ) were identified. Hot spots are located in the Central/Pampean region and in northern Patagonia, with high specific mortality rates related to AD and a high frequency of VSs.

**Table 1.** Comparison between the 50 most frequent Volga surnames in Argentina, the State of Hesse (HS), and Germany (GS). The surnames that are common between Argentina and HS are highlighted in yellow, those common between Argentina and GS are highlighted in red, and the surnames that are common in all databases are shown in grey

Argentina	State of Hesse	Germany
MULLER	MÜLLER	HANSEN
COLMAN	SCHMIDT	KELLER
SCHMIDT	SCHNEIDER	LANG
SCHNEIDER	SCHÄFER	FRANKE
WAGNER	BECKER	FRANK
FRANK	WEBER	BRANDT
WEBER	WAGNER	JUNG
MAYER	FISCHER	LORENZ
MEYER	SCHMITT	ALBRECHT
FISCHER	HOFMANN	VOGEL
KLOSTER	KOCH	HAHN
BECKER	WOLF	WINKLER
SMITH	KLEIN	FRIEDRICH
WALTER	JUNG	FUCHS
KUHN	HARTMANN	WEISS
ROTH	BAUER	REGNER
KLEIN	HOFFMANN	MAJER
JACOB	RICHTER	MAYER
SCHULZ	WERNER	GÜNTHER
MILLER	MÖLLER	HUBER
SCHWINDT	ROTH	SCHUBERT
BAUER	SCHWARZ	KÖNING
HOFFMANN	BRAUN	KAISER
HANSEN	SCHULZ	SCHMID
KOCH	MEYER	SCHOLZ
KELLER	KÖHLER	HERRMANNN
BRAUN	KAISER	WALTER
GRAFF	NEUMANN	KÖHLER
VOGEL	WALTER	SCHMITT
WEISS	ZIMMERMANN	BRAUN
FRITZ	LANG	LEHMANN
RUPPEL	HAHN	HOFMANN
SCHWAB	JÄGER	SCHULZE
FUHR	KELLER	MEIER

(Continued)

Table 1. (Continued)

Argentina	State of Hesse	Germany
HEIT	FRIEDRICH	SCHMITZ
HUBER	SIMON	MÜLLER
BERGER	FUCHS	PETERS
KESSLER	STEIN	WERNER
REGNER	SCHRÖDER	HARTMANN
BROWN	DIEHL	ZIMMERMANN
MEIER	KRAFT	WOLF
MAIER	SAUER	KRAUSE
ZIMMERMANN	HERRMANN	LANGE
ULRICH	BECK	SCHWARTZ
GETTE	MICHEL	KLEIN
RIEDEL	WINTER	BAUER
SCHEFFER	KÖNIG	KRÜGER
LANG	KRÄMER	SCHÄFER
KLOSTERBECKER	LUDWIG	NEUMANN

Conversely, cold spots, or areas with low values, are located in the Northwest and Northeast regions (Fig. 2c). No significant spatial associations were found in the rest of the country.

The Q-Q and residual plots for Poisson, Negative Binomial, and zero-inflated Poisson models are shown in Fig. 3. The AIC is a measure of the quality of data usage in a model that penalises complexity. A model with a lower AIC is considered better than one with a higher AIC. The Negative Binomial distribution model had the lowest AIC values (Table 3) and suggests that for each unit of variation of VS, the SDR increases by 0.4%. The high frequency of VSs alone explains 43.53% of the observed variability of SDR, according to Nagelkerke R square.

## Discussion

Diagnosing AD requires a combination of clinical evaluation, neuroimaging, and biomarkers. However, economic constraints often limit comprehensive studies, leading to underrepresentation or a lack of knowledge about certain mutations' frequency. Molecular studies of genes associated with this disease are infrequent in Argentina. In their study, *Bird et al.* (1988) described families that shared a single N141I mutation in the PSEN2 gene. Although more than 10 additional mutations in PSEN2 have been reported, the N141I mutation has only been found in families of Volga origin, suggesting its specificity to this population group (Blauwendraat *et al.*, 2016). Llibre-Guerra *et al.* (2020) conducted a meta-analysis in Latin America to identify pathogenic variants of autosomal dominant AD, including an investigation into the presence of N141I. Twenty-four variants were detected in 3,583 individuals at risk, mostly of European ancestry and typically attributable to founder effects. The frequency of these variants was higher in Colombia, followed by Puerto Rico and Mexico. A meta-analysis of 47 countries that reported variants in at least one of the three genes APP, PSEN1, and PSEN2 showed that the N141I variant is only found in Argentina, Germany, and the United States (Dehghani *et al.*, 2021).



**Table 2.** Frequency of Volga surnames (VSs) and specific death rates (SDRs) by regions, provinces, and the entire country

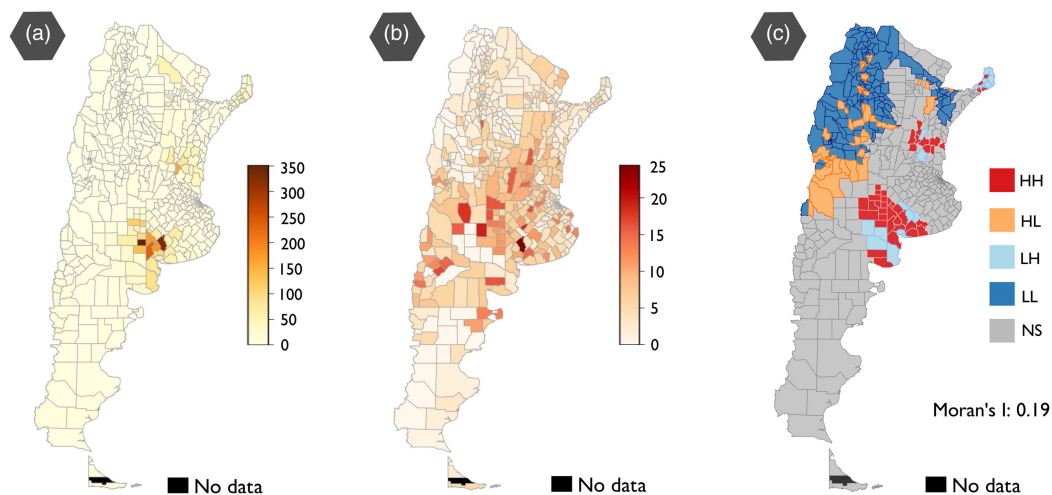
Region	Provinces	Voters 2015	Volga surnames (VSs) carriers	VS*1000	Number of Deaths 2005–2017	Deaths related to Alzheimer's diseases	Specific death rates (SDRs)
<b>Central/ Pampean</b>	Buenos Aires	11,384,389	142,590	12.53	1,677,007	5,258	3.13
	Ciudad Autónoma de Buenos Aires	2,541,076	20,750	8.17	422,702	1,129	2.67
	Córdoba	2,645,525	13,766	5.20	360,432	2,148	5.96
	Entre Ríos	979,546	46,768	47.74	128,128	736	5.74
	La Pampa	262,030	14,769	56.36	32,465	338	10.41
	Santa Fe	2,552,338	30,694	12.03	378,409	2,795	7.39
	<b>Total</b>	<b>20,364,904</b>	<b>269,337</b>	<b>13.23</b>	<b>2,999,143</b>	<b>12,404</b>	<b>4.14</b>
<b>Northwest (NWA)</b>	Catamarca	278,151	361	1.30	28,986	82	2.83
	Jujuy	478,463	477	1.00	51,760	101	1.95
	La Rioja	250,537	368	1.47	25,779	45	1.75
	Salta	885,984	1,351	1.52	92,080	271	2.94
	Santiago del Estero	648,777	1,102	1.70	70,362	132	1.88
	Tucumán	1,079,057	1,519	1.41	127,163	515	4.05
	<b>Total</b>	<b>3,620,969</b>	<b>5,178</b>	<b>1.43</b>	<b>396,130</b>	<b>1,146</b>	<b>2.89</b>
<b>Cuyo</b>	Mendoza	1,307,278	2,634	2.01	167,537	1,271	7.59
	San Juan	504,837	487	0.96	60,880	185	3.04
	San Luis	334,603	1,093	3.27	37,218	183	4.92
	<b>Total</b>	<b>2,146,718</b>	<b>4,214</b>	<b>1.96</b>	<b>265,635</b>	<b>1,639</b>	<b>6.17</b>
<b>Northeast (NEA)</b>	Chaco	815,907	8,359	10.25	92,655	268	2.89
	Corrientes	765,271	2,972	3.88	85,889	257	2.99
	Formosa	392,863	2,238	5.70	43,792	268	6.12
	Misiones	787,588	19,489	24.75	82,463	310	3.76
	<b>Total</b>	<b>2,761,629</b>	<b>33,058</b>	<b>11.97</b>	<b>304,799</b>	<b>1,103</b>	<b>3.62</b>

(Continued)



Table 2. (Continued)

Region	Provinces	Voters 2015	Volga surnames (VSs) carriers	VS*1000	Number of Deaths 2005–2017	Deaths related to Alzheimer's diseases	Specific death rates (SDRs)
<b>Patagonia</b>	Chubut	388,934	3,208	8.25	38,443	194	5.05
	Neuquén	436,081	3,397	7.79	36,851	332	9.01
	Río Negro	474,634	5,979	12.60	50,768	373	7.38
	Santa Cruz	220,278	1,481	6.72	17,353	19	1.09
	Tierra del Fuego	116,042	1,070	9.22	6,094	16	2.63
	<b>Total</b>	<b>1,635,969</b>	<b>15,135</b>	<b>9.25</b>	<b>149,509</b>	<b>934</b>	<b>6.25</b>
<b>Argentina</b>	<b>Total</b>	<b>30,530,189</b>	<b>326,922</b>	<b>10.71</b>	<b>4,115,216</b>	<b>17,226</b>	<b>4.19</b>



**Figure 2.** Choropleth map of frequency by departments of Volga surnames (a), and specific death rates related to AD (b). LISA maps illustrate the geographic clustering of both variables (c). Departments with high values that are surrounded by neighbours with high values are coloured in red (HH). High values surrounded by low values are coloured in orange (HL). Departments with low values surrounded by high values are coloured in light blue (LH). Low values clusters or cold spots are coloured in blue (LL). In grey, non-significant (NS).

In 2010, Yu *et al.* identified a haplotype based on six single-nucleotide polymorphisms that cover the PSEN2 gene. The frequency of this haplotype was 0.64 in Volga Germans carrying the N141I mutation, compared to 0.26 in Volga Germans without the mutation and 0.13 in Europeans typed by the Centre d'Etude du Polymorphisme Humain. The study suggests that the N141I mutation in PSEN2 may have occurred before the emigration from the Hesse region to Russia. The discovery of families with this mutation living in Argentina and Germany suggests the possibility of additional cases sharing this common ancestry (Yu *et al.*, 2010; Muchnik *et al.*, 2015). The mutations responsible for familial forms have known biochemical consequences that are likely to be at the root of sporadic AD. Early interventions can delay or even prevent dementia in asymptomatic individuals and families at risk, as well as slow progression in those with symptoms (Bateman *et al.*, 2011).

A study conducted at the beginning of the 21st century in Germany provided estimated values on the prevalence and incidence of dementing illness through large-scale epidemiological and meta-analyses (Bickel, 2000). The prevalence of dementia is higher in the states of Baden-Württemberg, Bavaria, Lower Saxony, North Rhine-Westphalia, and Hesse. This region is also the homeland of the Volga Germans. However, in recent times, there has been a change in this trend, particularly due to the increased frequency of cases of late-onset AD in female patients (Ziegler and Doblhammer, 2009; Lange *et al.*, 2017). These differences are primarily attributed to women's longer life expectancy, as advanced age remains the greatest risk factor for AD (Chêne *et al.*, 2015). Of all the registered deaths due to Alzheimer's in Argentina, 68% were women and 31% were men. A significant difference in the distribution of SDR and VS was observed between the Central/Pampean, Cuyo, and Patagonia regions compared to the NWA region. The NWA region comprises various environments, including the Andean foothills, where populations live at altitudes above 2,500 metres. Genomic studies indicate that the NWA populations have the highest proportion of Central Andean ancestry component (Muzzio *et al.*, 2018; Luisi *et al.*, 2020). Meanwhile, individuals from the province of Misiones (NEA) have the highest proportion of Central/Northern European ancestry. This aligns with the historical record of settlement of Polish, German, Danish, and Swedish colonies in this province (Luisi *et al.*, 2020).

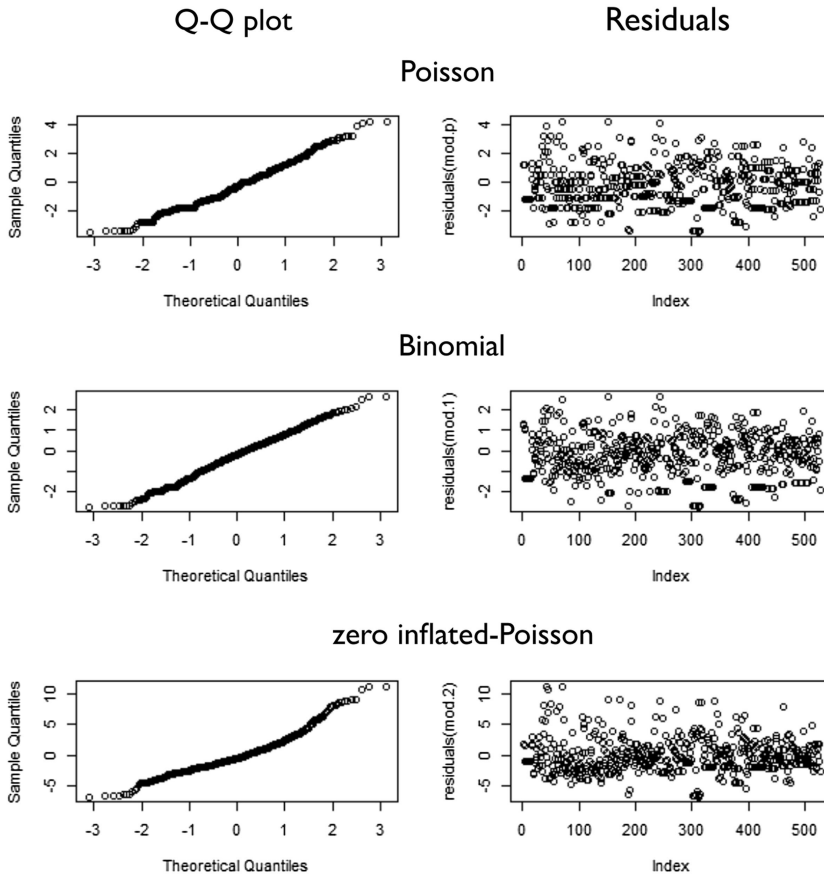


Figure 3. Q-Q and residuals plots for Poisson, Negative Binomial, and zero-inflated Poisson models.

Geographical factors and the biological history of a population can also have an impact on the risk of developing dementia (Alzheimer's Association, 2020). The migration of Volga Germans to Argentina should be distinguished from the migration of other German groups. According to Germanic sources, the massive migration took place between 1883 and 1890. However, the majority of the contingents that left through the Port of Bremen until 1870 went to the United States, Canada, and Brazil. Argentina only became a destination of interest after 1870. German migrants left from ports in Hannover and Bremen, and mainly came from cities in northern Germany, but also including the Duchy of Hesse (de Flachs, 1994). According to the Argentina National Censuses, German migrants represented a minority, comprising no more than 2.35%, 1.70%, and 1.14% of the total population in the 1869, 1895, and 1914 censuses, respectively. The immigrants were mainly concentrated in the same provinces as the Volga Germans, but they also settled in other destinations in Argentina. For example, 2.3% of the German immigrants registered in the aforementioned censuses resided in the NOA region and the majority of them were in the province of Tucumán. Volga migration remained stable in the provinces of Buenos Aires and La Pampa (Central/Pampean region). As shown in Table 3, the highest SDR in the country was recorded in the province of La Pampa (10.41). For this calculation, all deaths were considered according to ICD codes G30.0 (early-onset AD), G30.1 (late-onset AD), G30.8 (other AD), G30.9 (AD, unspecified), and G31 (other degenerative diseases of the nervous system). The rate was also

**Table 3.** Akaike information criterion (AIC) values for each model. Df, degrees of freedom

Model	df	AIC
Poisson (Mod. Poiss)	3	2563,192
Binomial (Mod. BN)	4	2445,948
zero-inflated Poisson (Mod. ZIP)	5	2458,591

calculated specifically for codes G30.0 and G30.9 to better weigh the impact of deaths from early-onset AD and to account for the possibility of under-diagnosis. Once again, the province of La Pampa had the highest SDR (9.61), with 483 deaths in the period analysed.

German surnames display considerable lexical, phonological, and morphological variation, which is reflected in their distribution across different regions (Dräger and Schmuck, 2009). The comparison of the 50 most common surnames shows that the Volga Germans in Argentina are more closely related to the population of Hesse than to any other German state. This makes them excellent markers for analysing large databases. Several dialects are spoken in Hesse. This linguistic diversity is also reflected in the specificity of family names. According to Rodríguez-Larralde *et al.* (1978), who based their research on the Lasker isonymic distance between German cities, this state is part of a cluster in southern Germany. The region is home to speakers of Rhine Franconian (including Hessian), Alemannic, and Bavarian (Konig, 1978). In Russia, the Volga German colonies were linguistic isolates with limited bilingualism and a situation of internal diglossia. Hipperdinger (2017) suggests that this sociolinguistic characteristic led to the replacement of Russian with Spanish in Argentina, a situation that lasted until the second half of the twentieth century. The VSs maintained their distinctiveness in both contexts.

## Conclusion

AD is a significant health problem, especially due to the ageing population. Familial forms of AD represent a small percentage of cases but are critical to study. This requires a comprehensive understanding of the biology of the population. The combination of historical and official documents, such as electoral rolls, census data, and health information, can be used to analyse population dynamics. Our study uses a methodological approach that provides a coherent view of the spatial distribution of deaths from AD in Argentina and its relationship to migration processes by combining diachronic (surnames) and synchronous (death certificates) data. Nearly 150 years after the establishment of the first colonies, the population descended from the Volga migration remains highly concentrated in the southeastern departments of La Pampa Province and the southwestern departments of Buenos Aires Province. Within the Central Region, these two areas are contiguous. This demographic behaviour has health consequences. These departments are included in a statistically significant cluster with a high frequency of surnames of Volga origin and high SDRs from AD. Tracing surnames by origin is a cost-effective method for distinguishing structures within a seemingly homogeneous social group. These analyses provide a reliable basis for guiding patient recruitment in medical research and reducing sampling error by identifying where and when a pre-existing genetic pattern is likely to persist. This approach may be useful for describing complex migration scenarios in other Latin American countries undergoing similar population processes.

**Acknowledgements.** We express our gratitude to Dr. Rolando González-José and the authorities of the Cámara Nacional Electoral Argentina for providing us with access to the 2015 Electoral Register.

**Funding statement.** This work was supported by Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), grant PIP 2021–2023 number 11220200101902.

**Competing interests.** Authors declare no competing interests.

**Ethical standard.** The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

## References

- Akushevich I, Kravchenko J, Yashkin A, Doraiswamy PM, Hill CV, and Alzheimer's Disease and Related Dementia Health Disparities Collaborative Group (2023) Alzheimer's Disease and Related Dementia Health Disparities Collaborative Group. Expanding the scope of health disparities research in Alzheimer's disease and related dementias: Recommendations from the "Leveraging Existing Data and Analytic Methods for Health Disparities Research Related to Aging and Alzheimer's Disease and Related Dementias" Workshop Series. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* 15(1), e12415. <https://doi.org/10.1002/dad2.12415>
- Albeck ME, Alfaro EL, Dipierri JE, and Chaves ER (2017) Los apellidos de Salta en el siglo XXI: origen geo-lingüístico, diversidad y frecuencia. *Andes* 28(2), 2–21. <https://www.redalyc.org/articulo.oa?id=12755958008>.
- Alzheimer's Association (2020) 2020 Alzheimer's disease facts and figures. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* 16(3), 391–460. <https://doi.org/10.1002/alz.12068>
- Andrade-Guerrero J, Santiago-Balmaseda A, Jeronimo-Aguilar P, Vargas-Rodríguez I, Cadena-Suárez AR, Sánchez-Garibay C, Pozo-Molina G, Méndez-Catalá CF, Cardenas-Aguayo MD, Diaz-Cintra S, Pacheco-Herrero M, Luna-Muñoz J, and Soto-Rojas LO (2023) Alzheimer's disease: an updated overview of its genetics. *International Journal of Molecular Sciences* 24(4), 3754. <https://doi.org/10.3390/ijms24043754>
- Anselin L (1995) Local indicators of spatial association—LISA. *Geographical Analysis* 27, 93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- AufderHeide E (2006) Representations of German-Speaking Exiles and Immigrants in Argentina [Bowling Green State University/OhioLINK]. URL: [http://rave.ohiolink.edu/etdc/view?acc\\_num=bgsu1162858784](http://rave.ohiolink.edu/etdc/view?acc_num=bgsu1162858784)
- Bateman RJ, Aisen PS, De Strooper B, Fox NC, Lemere CA, Ringman JM, Salloway S, Sperling RA, Windisch M, and Xiong C (2011) Autosomal-dominant Alzheimer's disease: a review and proposal for the prevention of Alzheimer's disease. *Alzheimer's Research & Therapy* 3(1), 1. <https://doi.org/10.1186/alzrt59>
- Bickel H (2000) Dementia syndrome and Alzheimer disease: an assessment of morbidity and annual incidence in Germany. *Gesundheitswesen* 62(4), 211–218. <https://doi.org/10.1055/s-2000-10858>
- Bird TD, Lampe TH, Nemens EJ, Miner GW, Sumi SM, and Schellenberg GD (1988) Familial Alzheimer's disease in American descendants of the Volga Germans: probable genetic founder effect. *Annals of Neurology* 23(1), 25–31. <https://doi.org/10.1002/ana.410230106>
- Bird TD, Nemens EM, Nochlin D, Sumi SM, Wijsman EM, and Schellenberg GD (1992) Familial Alzheimer's disease in Germans from Russia: a model of genetic heterogeneity in Alzheimer's disease. Heterogeneity of Alzheimer's Disease. In *Research and Perspectives in Alzheimer's Disease*. Berlin, Heidelberg: Springer. [https://doi.org/10.1007/978-3-642-46776-9\\_13](https://doi.org/10.1007/978-3-642-46776-9_13)
- Blauwendraat C, Wilke C, Jansen IE, Schulte C, Simón-Sánchez J, Metzger FG, Bender B, Gasser T, Maetzel W, Rizzu P, Heutink P, and Synofzik M (2016) Pilot whole-exome sequencing of a German early-onset Alzheimer's disease cohort reveals a substantial frequency of PSEN2 variants. *Neurobiology of Aging*, 37, 208.e11–208.e17. <https://doi.org/10.1016/j.neurobiolaging.2015.09.016>
- Chêne G, Beiser A, Au R, Preis SR, Wolf PA, Dufouil C, and Seshadri S (2015) Gender and incidence of dementia in the Framingham Heart Study from mid-adult life. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* 11(3), 310–320. <https://doi.org/10.1016/j.jalz.2013.10.005>
- Dalmasso MC, de Rojas I, Olivari N, Muchnik C, Angel B, Gloger S, Sanchez Abalos MS, Chacón MV, Aránguiz R, Orellana P, Cuesta C, Galeano P, Campanelli L, Novack GV, Martínez LE, Medel N, Lisso J, Sevillano Z, Irureta N, Castaño EM, Montreal L, Thoenes M, Hanses C, Heilmann-Heimbach S, Kairiyama C, Mintz I, Villella I, Rueda F, Romero A, Wukitsevtis N, Quiroga I, Gona C, Lambert JC, Solís P, Politis DG, Mangone CA, Gonzalez-Billault C, Boada M, Tarraga L, Slachevsky A, Albala C, Fuentes P, Kochen S, Brusco LI, Ruiz A, Morelli L, and Ramírez A (2023) The first genome-wide association study in the Argentinian and Chilean populations identifies shared genetics with Europeans in Alzheimer's disease. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*. [10.1002/alz.13522](https://doi.org/10.1002/alz.13522). Advance online publication. <https://doi.org/10.1002/alz.13522>
- de Flachs MCV (1994) Emigraciones transoceánicas. Los alemanes en América. 1850-1914. El caso argentino. *Cuadernos de Historia Contemporánea* 16, 65.
- Dehghani N, Bras J, and Guerreiro R (2021) Erratum to: How understudied populations have contributed to our understanding of Alzheimer's disease genetics. *Brain* 144(8), e69. <https://doi.org/10.1093/brain/awab028>
- Dirección de Estadísticas e Información de la Salud – DEIS (2023) Bases de datos de defunciones. URL: <https://www.argentina.gob.ar/salud/deis/datos/defunciones> (accessed on 02nd April 2023).

- Dräger K, and Schmuck M** (2009) The German Surname Atlas Project—Computer-based surname geography, Mainz. Available from: <https://www.germanistik.uni-mainz.de/files/2015/01/Dr%C3%A4ger-Schmuck2009.pdf>.
- García MJ, and Comesaña A** (2021) Prevalence of neurocognitive disorders in a rural area of Argentina. *Revista de la Facultad de Ciencias Médicas* 78(4), 347–352.
- Gushulak BD, and MacPherson DW** (2006) The basic principles of migration health: population mobility and gaps in disease prevalence. *Emerging Themes in Epidemiology* 3, 3. <https://doi.org/10.1186/1742-7622-3-3>
- Hipperding YH** (2017) Las lenguas inmigratorias en la Argentina: El caso de los alemanes del Volga. *Sociedad y Discurso* 30, 92–114. <https://doi.org/10.5278/ojs.s%20&%20d.v0i30.1851>
- Instituto Nacional de Estadísticas y Censos de la República Argentina – INDEC** (2023) Estadísticas de población por año y región. URL: <https://www.indec.gov.ar/indec/web/Nivel3-Tema-2-41> (accessed on 15th November 2023).
- Itzcovich T, Chrem-Méndez P, Vázquez S, Barbieri-Kennedy M, Niikado M, Martinetto H, Allegri R, Sevlever G, and Surace EI** (2020) A novel mutation in PSEN1 (p.T119I) in an Argentine family with early- and late-onset Alzheimer's disease. *Neurobiology of Aging* 85, 155.e9–155.e12. <https://doi.org/10.1016/j.neurobiolaging.2019.05.001>
- König W** (1978) *dtv-Atlas zur deutschen Sprache* (First Edition). Deutscher Taschenbuch Verlag, Munich.
- Lang L, Schulte T, Dittmann B, and Hildebrandt H** (2017) 2 Regionale Verteilung der Demenz sowie Inanspruchnahme vor und nach Erstdiagnose. *Monitor Versorgungsforschung* 41–46. <https://doi.org/10.24945/MVF.05.18.1866-0533.2098>
- Larraya FP, Grasso L, and Mari G** (2004) Prevalence of dementia of the Alzheimer type, vascular dementia and other DSM-IV and ICD-10 dementias in the Republica Argentina. *Revista Neurologica Argentina* 29, 148–153.
- Llibre-Guerra JJ, Li Y, Allegri RF, Mendez PC, Surace EI, Llibre-Rodríguez JJ, Sosa AL, Aláez-Verson C, Longoria EM, Tellez A, Carrillo-Sánchez K, Flores-Lagunes LL, Sánchez V, Takada LT, Nitri N, Ferreira-Frota NA, Benevides-Lima J, Lopera F, Ramírez L, Jiménez-Velázquez I, Schenk C, Acosta D, Behrens MI, Doering M, Ziegemeier E, Morris JC, McDade E, and Bateman RJ** (2020) Dominantly inherited Alzheimer's disease in Latin America: genetic heterogeneity and clinical phenotypes. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association* 17(4), 653–664. <https://doi.org/10.1002/alz.12227>
- Luisi P, García A, Berros JM, Motti JMB, Demarchi DA, Alfaro E, Aquilano E, Argüelles C, Avena S, Bailliet G, Beltramo J, Bravi CM, Cuello M, Dejan C, Dipierri JE, Jurado Medina LS, Lanata JL, Muzzio M, Parolin ML, Pauro M, Paz Sepúlveda PB, Rodríguez Golpe D, Santos MR, Schwab M, Silvero N, Zubrzycki J, Ramallo V, and Dopazo H** (2020) Fine-scale genomic analyses of admixed individuals reveal unrecognized genetic ancestry components in Argentina. *PLoS One* 15(7), e0233808. <https://doi.org/10.1371/journal.pone.0233808>
- Mateos P** (2014) *Names, Ethnicity and Populations. Series: Advances in Spatial Science*. Berlin and Heidelberg, Germany: Springer. ISBN: 978-3-642-45412-7
- Melcon CM, Bartoloni L, Katz M, Del Mónaco R, Mangone CA, Melcon MO, and Allegri RF** (2010) Propuesta de un Registro centralizado de casos con Deterioro Cognitivo en Argentina (ReDeCAr) basado en el Sistema Nacional de Vigilancia Epidemiológica. *Neurología Argentina* 2(3), 161–166.
- Méndez PC, Calandri I, Nahas F, Russo MJ, Demey I, Martín ME, Clarens MF, Harris P, Tapajoz F, Campos J, Surace EI, Martinetto H, Ventrice F, Cohen G, Vázquez S, Romero C, Guinjoan S, Allegri RF, and Sevlever G** (2018) Argentina-Alzheimer's disease neuroimaging initiative (Arg-ADNI): neuropsychological evolution profile after one-year follow up. *Archivos de neuro-psiquiatria* 76(4), 231–240. <https://doi.org/10.1590/0004-282x20180025>
- Morales AL, Navarro P, Cintas C, González-José R, Ramallo V, and Delrieux C** (2021) Bulsurapp: interactive visual analysis for surname trend exploration. *IEEE Computer Graphics and Applications* 42, 28–33.
- Muchnik C, Olivar N, Dalmasso MC, Azurmendi PJ, Liberczuk C, Morelli L, and Brusco LI** (2015) Identification of PSEN2 mutation p.N141I in Argentine pedigrees with early-onset familial Alzheimer's Disease. *Neurobiology of Aging* 36(10), 2674–7.e1. <https://doi.org/10.1016/j.neurobiolaging.2015.06.011>
- Muzzio M, Motti JMB, Paz Sepúlveda PB, Yee MC, Cooke T, Santos MR, Ramallo V, Alfaro EL, Dipierri JE, Bailliet G, Bravi CM, Bustamante CD, and Kenny EE** (2018) Population structure in Argentina. *PLoS One* 13(5), e0196325. <https://doi.org/10.1371/journal.pone.0196325>
- Pohl JO** (2009) Volk auf dem Weg: transnational migration of the Russian-Germans from 1763 to the present day. *Studies in Ethnicity and Nationalism* 9(2), 267–286.
- Rodríguez-Larralde A, Barrai I, Nesti C, Mamolini E, and Scapoli C** (1998) Isonymy and isolation by distance in Germany. *Human Biology* 70(6), 1041–1056.
- Yu CE, Marchani E, Nikisch G, Müller U, Nolte D, Hertel A, Wijsman EM, and Bird TD** (2010) The N141I mutation in PSEN2: implications for the quintessential case of Alzheimer disease. *Archives of Neurology* 67(5), 631–633. <https://doi.org/10.1001/archneurol.2010.87>
- Ziegler U, and Doblhammer G** (2009) Prävalenz und Inzidenz von Demenz in Deutschland—Eine Studie auf Basis von Daten der gesetzlichen Krankenversicherungen von 2002. *Das Gesundheitswesen* 71(05), 281–290.

**Cite this article:** Morales AL, Figueroa MI, Navarro P, Chaves ER, Ruderman A, Dipierri JE, and Ramallo V (2024). Volga German surnames and Alzheimer's disease in Argentina: an epidemiological perspective. *Journal of Biosocial Science* 56, 625–638. <https://doi.org/10.1017/S002193202400018X>