Alpha-1 antitrypsin Pi*Z gene frequency and Pi*ZZ genotype numbers worldwide: an update

Ignacio Blanco¹
Patricia Bueno²
Isidro Diego³
Sergio Pérez-Holanda⁴
Francisco Casas-Maldonado⁵
Cristina Esquinas⁶
Marc Miravitlles⁶,⁷

¹Alpha-1 Antitrypsin Deficiency Spanish Registry (REDAAT), Fundación Respira, Spanish Society of Pneumology and Thoracic Surgery (SEPAR), Barcelona, ²Internal Medicine Department, County Hospital of Jarrio, ³Materials and Energy Department, School of Mining Engineering, Oviedo University, ⁴Surgical Department, University Central Hospital of Asturias (HUCA), Oviedo, Principality of Asturias, ⁵Pneumology Department, Complejo Hospitalario Universitario de Granada, Granada, ⁶Pneumology Department, Hospital Universitari Vall d’Hebron, ⁷CIBER de Enfermedades Respiratorias (CIBERES), Barcelona, Spain

Abstract: In alpha-1 antitrypsin deficiency (AATD), the Z allele is present in 98% of cases with severe disease, and knowledge of the frequency of this allele is essential from a public health perspective. However, there is a remarkable lack of epidemiological data on AATD worldwide, and many of the data currently used are outdated. Therefore, the objective of this study was to update the knowledge of the frequency of the Z allele to achieve accurate estimates of the prevalence and number of Pi*ZZ genotypes worldwide based on studies performed according to the following criteria: 1) samples representative of the general population, 2) AAT phenotyping characterized by adequate methods, and 3) measurements performed using a coefficient of variation calculated from the sample size and 95% confidence intervals. Studies fulfilling these criteria were used to develop maps with an inverse distance weighted (IDW) interpolation method, providing numerical and graphical information of Pi*Z distribution worldwide. A total of 224 cohorts from 65 countries were included in the study. With the data provided by these cohorts, a total of 253,404 Pi*ZZ were estimated worldwide: 119,594 in Europe, 91,490 in America and Caribbean, 3,824 in Africa, 32,154 in Asia, 4,126 in Australia, and 2,216 in New Zealand. In addition, the IDW-interpolation maps predicted Pi*Z frequencies throughout the world even in some areas that lack real data. In conclusion, the inclusion of new well-designed studies and the exclusion of the low-quality ones have significantly improved the reliability of results, which may be useful to plan strategies for future research and diagnosis and to rationalize the therapeutic resources available.

Keywords: SERPINA1, alpha-1 antitrypsin deficiency, protease inhibitor, genetic epidemiology, inverse distance weighted interpolation, geographic information system

Introduction

Human alpha-1 antitrypsin (AAT), also named α1 proteinase inhibitor (α1-Pi) and SERPINA1 (serine protease inhibitor, group A, member 1), is a water-soluble, tissue-diffusible, medium-sized circulating glycoprotein, with a molecular weight of 52 kDa, and a blood half-life of around 5 days. AAT is mainly synthesized by hepatocytes, reaching plasma concentrations of 1–2 g L⁻¹. From the plasma, 80% of AAT diffuses to interstitial tissues, and up to 10% reaches biological fluids. AAT acts as a major inhibitor of leukocyte elastase, proteinase-3, and bacterial serine proteases. In addition, it has anti-inflammatory and immunomodulatory properties.¹ Alpha-1 antitrypsin deficiency (AATD) is a rare hereditary condition that predisposes to premature onset of COPD especially in smokers, liver cirrhosis in children and adults, neutrophilic panniculitis, systemic vasculitis, as well as other possible inflammatory, autoimmune, and neoplastic diseases.²⁻⁴

The AAT gene has two alleles, which are transmitted by autosomal codominant Mendelian inheritance. Normal alleles are present in 85%–90% of individuals and are...
of tobacco smoke, occupational exposure to pollutants and hepatotoxic substances, and information on clinical management, including intravenous augmentation therapy.\(^1\)

Since Laurell and Eriksson discovered AATD in 1963,\(^8\) a large number of original studies and reviews on the genetic epidemiology of the AAT gene have been published.\(^9\)-\(^{22}\) However, most were conducted in Europe, the US, and Canada, with a remarkable lack of data from more than half of the 193 countries in the world. Moreover, there is a great lack of uniformity in the methodology used to achieve the results of some of these studies leading to the presentation of inaccurate data and the subsequent need to consider these results with caution. It is of note that in recent years, the results of other well-designed studies have substantially changed previous findings.\(^23\)-\(^{25}\) All of the above justify the need for an update of the current data available, and therefore, the aim of this study was to develop an updated database of valid epidemiological studies on Pi*Z frequency to make accurate estimates on the prevalence and number of Pi*ZZ genotypes throughout the world.

### Methods

#### Study selection criteria

1. Samples representative of the general population.
2. AAT phenotyping characterized by adequate methods, including IEF, or alternatively, crossed antigen-antibody electrophoresis or polymerase chain reaction.
3. To assess the statistical validity of the results, a coefficient of variation (CV) of the Pi*Z frequencies in each cohort was calculated. This CV is a measure of the precision of results in terms of dispersion of the data around the mean. Its value depends on the number of alleles studied and on the frequencies of Pi*Z actually found, according to the formula:

\[
Z_{cv} = \frac{100 \times (Z_{ul} - Z_{ll})}{4 \times Z_{ll}}
\]

where \(Z_{ul}\) is the 95% confidence interval (95% CI) upper limit of \(Z\), \(Z_{ll}\) is the 95% CI lower limit of \(Z\), and \(Z_{ll}\) is the frequency of \(Z\). The value of this variable was adjusted for each country, according to the methodology used in several previous peer-reviewed publications by our group.\(^{18}\)-\(^{20}\)

The following studies were excluded from the analysis:

1. Cohort studies of patients with AATD-related diseases (eg, COPD, liver cirrhosis, etc.).
2. Screening studies in which phenotypes were determined only in samples with AAT concentrations below any given cut-off point.
Calculation of Pi*Z allele frequency

The Z allelic frequency was obtained by adding the number of Z alleles and expressing this total as a fraction of the total number of Pi alleles in the population (expressed as alleles per 1,000 of all Pi types).

Inverse distance weighted multivariate interpolation maps

Pi*Z gene frequencies were used to develop maps using a geographic information system (GIS) of an inverse distance weighted (IDW) interpolation method. IDW interpolation is based on the assumption that the attributed value of an unsampled point is the weighted average of known values within the neighborhood, and the weights are inversely related to the distances between the prediction location and the locations sampled.28 With this approach, Pi*Z frequencies were represented by scales of 10 colors (with red representing the maximum frequencies; yellow, green, and grey intermediate frequencies, and blue minimum frequencies). The colors of these maps can be converted into quantitative data, providing both graphical and numerical information on their geographical distribution worldwide.

Results

Cohorts selected and number of subjects studied

A total of 224 cohorts (90 from Europe, 42 from America, 17 from Africa, 66 from Asia, 8 from Australia, and 1 from New Zealand; Supplementary material) were selected. The database comprised 65 of the 193 countries in the world: 25 from Europe, 23 from America, 15 from Asia, 1 from Australia, and 1 from New Zealand.

These cohorts included a total of 193,444 subjects, distributed as follows: 75,603 in Europe; 85,118 in America; 5,031 in Africa; 21,353 in Asia; 5,536 in Australia; and 1,303 in New Zealand.

Pi*Z frequency and Pi*ZZ prevalence

Tables 1–5 show the estimated frequency of Pi*Z (expressed per 1,000, with 95% CI), the prevalence of Pi*ZZ (expressed as 1:x), and the calculated number of Pi*ZZ genotypes (with 95% CI) in the countries included in the study.

The Atlantic and sea coast regions of Europe showed the highest frequency and prevalence of Pi*Z and Pi*ZZ (Figure 1), which gradually fell in eastern Europe and practically disappeared in Asia, except in Pakistan and Thailand, which presented moderate values (Figure 2). Some regions of eastern Africa had moderate Pi*Z values, being very low or absent in the remainder of Africa (Figure 3).

In North America, the frequency of Pi*Z is high in the east of Canada and gradually decreases to the west and the north of the country. The same was found in the US with high values in the eastern states, which progressively decreased in the western and south-eastern areas of the country. Pi*Z frequencies were found to be low in Mexico, Central America, and in the Caribbean and South America countries.

South-eastern Australia showed moderate Pi*Z frequencies, being very high in New Zealand.

IDW-interpolation maps of Pi*Z frequencies

Figures 1–3 show the IDW-interpolation maps with the Pi*Z values found in Europe (Figure 1) and worldwide (Figures 2 and 3). Figure 3 shows the distribution of Pi*Z frequencies in the world. The largest number of studies was found in Europe, being scarce and scattered throughout the remainder of the world, making the results from areas with limited data less reliable.

Estimated numbers of Pi*ZZ genotypes

With the data provided by the cohorts selected from the 65 countries studied, a total of 253,404 ZZ genotypes were estimated and distributed as follows: Europe (119,594; Table 1); America (91,490; Table 2); Africa (3,824; Table 3); Asia (32,154; Table 4); and Australia (4,126) and New Zealand (2,216) (Table 5).

Of the almost 120,000 subjects identified as carrying Pi*ZZ genotype in Europe, most may be found in Germany (20,611), France (17,191), the Iberian Peninsula (14,522 in Spain and 4,944 in Portugal), the UK (13,044), Italy (10,652), Poland (6,791), the Netherlands (5,353), Denmark (4,090), Latvia (4,005), Belgium (3,193), and lower numbers in other European countries (Table 1).

In North, Central, and South America and the Caribbean islands ~90,000 individuals carry Pi*ZZ genotypes (Table 2), most being found in the US (~70,000). In addition, it has been estimated that there may be around 7,000 ZZ genotypes in Canada; 4,000 in Mexico; 6,000 in Brazil; 2,000 in Colombia, Venezuela, and Argentina; 1,000 in Peru; 700 in Chile; 600 in Cuba; and lower numbers in other countries from the American continents.

Table 3 shows that among the >1.2 billion inhabitant of Africa, there may be only around 4,000 Pi*ZZ, being mainly distributed in Nigeria and Somalia.

It has been estimated that among the >4 billion people from Asia, ~30,000 have the Pi*ZZ, most of whom may be found in Pakistan, Thailand, and South Korea (Table 4).
Table 1 Estimated Pi*Z allele frequency (per 1,000), Pi*ZZ gene prevalence (1:x), and number of Pi*ZZ genotypes among 25 European countries

<table>
<thead>
<tr>
<th>Geographic region (number of countries)</th>
<th>Country (total population)</th>
<th>Number of cohorts</th>
<th>Number of subjects</th>
<th>Pi*Z frequency (95% CI)</th>
<th>Pi*ZZ prevalence (mean)</th>
<th>Number of Pi*ZZ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern Europe (8)</td>
<td>Denmark (5,594,000)</td>
<td>2</td>
<td>10,096</td>
<td>27 (25–29)</td>
<td>1:3.68</td>
<td>4,090 (3,459–4,834)</td>
</tr>
<tr>
<td></td>
<td>Estonia (1,258,545)</td>
<td>2</td>
<td>1,636</td>
<td>24 (19–30)</td>
<td>1:1.663</td>
<td>752 (481–1,170)</td>
</tr>
<tr>
<td></td>
<td>Finland (5,498,211)</td>
<td>2</td>
<td>8,836</td>
<td>19 (17–21)</td>
<td>1:2.850</td>
<td>1,929 (1,553–2,394)</td>
</tr>
<tr>
<td></td>
<td>Iceland (335,878)</td>
<td>1</td>
<td>94</td>
<td>0 (0–19)</td>
<td>infinite</td>
<td>0 (0–127)</td>
</tr>
<tr>
<td></td>
<td>Latvia (1,965,686)</td>
<td>1</td>
<td>288</td>
<td>45 (30–66)</td>
<td>1:491</td>
<td>4,005 (1,802–8,648)</td>
</tr>
<tr>
<td></td>
<td>Lithuania (2,854,235)</td>
<td>1</td>
<td>1,577</td>
<td>16 (12–21)</td>
<td>1:3.979</td>
<td>717 (405–1,261)</td>
</tr>
<tr>
<td></td>
<td>Norway (5,265,158)</td>
<td>4</td>
<td>4,492</td>
<td>18 (16–21)</td>
<td>1:2.929</td>
<td>1,798 (1,321–2,442)</td>
</tr>
<tr>
<td></td>
<td>Sweden (9,880,604)</td>
<td>11</td>
<td>5,750</td>
<td>15 (13–17)</td>
<td>1:4.368</td>
<td>2,262 (1,674–3,052)</td>
</tr>
<tr>
<td>Total/mean</td>
<td></td>
<td>24</td>
<td>32,769</td>
<td>20</td>
<td>2.041</td>
<td>15,553</td>
</tr>
<tr>
<td>Western Europe (7)</td>
<td>Belgium (11,409,077)</td>
<td>1</td>
<td>1,345</td>
<td>17 (12–22)</td>
<td>1:3.573</td>
<td>3,193 (1,745–5,787)</td>
</tr>
<tr>
<td></td>
<td>France (66,836,154)</td>
<td>8</td>
<td>7,233</td>
<td>16 (14–18)</td>
<td>1:3.888</td>
<td>17,191 (13,255–22,270)</td>
</tr>
<tr>
<td></td>
<td>England (53,012,456)</td>
<td>3</td>
<td>2,832</td>
<td>15 (12–18)</td>
<td>1:4.400</td>
<td>11,939 (7,729–18,360)</td>
</tr>
<tr>
<td></td>
<td>The Netherlands (17,016,967)</td>
<td>4</td>
<td>1,466</td>
<td>18 (13–23)</td>
<td>1:3.179</td>
<td>5,353 (3,057–9,298)</td>
</tr>
<tr>
<td></td>
<td>Northern Ireland (1,810,863)</td>
<td>1</td>
<td>1,000</td>
<td>20 (15–28)</td>
<td>1:2.380</td>
<td>761 (404–1,417)</td>
</tr>
<tr>
<td></td>
<td>Republic of Ireland (4,757,976)</td>
<td>1</td>
<td>1,100</td>
<td>22 (16–29)</td>
<td>1:2.101</td>
<td>2,265 (1,264–4,019)</td>
</tr>
<tr>
<td>Total/mean</td>
<td></td>
<td>21</td>
<td>17,519</td>
<td>17</td>
<td>1:4.992</td>
<td>41,046</td>
</tr>
<tr>
<td>Central Europe (4)</td>
<td>Austria (8,711,770)</td>
<td>1</td>
<td>868</td>
<td>13 (9–20)</td>
<td>1:5.697</td>
<td>1,529 (646–3,536)</td>
</tr>
<tr>
<td></td>
<td>Germany (80,722,792)</td>
<td>5</td>
<td>2,691</td>
<td>16 (13–20)</td>
<td>1:3.916</td>
<td>20,611 (13,380–31,626)</td>
</tr>
<tr>
<td></td>
<td>Poland (38,523,261)</td>
<td>5</td>
<td>3,201</td>
<td>13 (11–16)</td>
<td>1:5.673</td>
<td>6,791 (4,395–10,454)</td>
</tr>
<tr>
<td></td>
<td>Switzerland (8,179,294)</td>
<td>1</td>
<td>1,148</td>
<td>11 (7–17)</td>
<td>1:7.798</td>
<td>972 (434–2,139)</td>
</tr>
<tr>
<td>Total/mean</td>
<td></td>
<td>12</td>
<td>7,908</td>
<td>13</td>
<td>1:5.821</td>
<td>10,652 (7,046–16,049)</td>
</tr>
<tr>
<td>Southern Europe (3)</td>
<td>Italy (62,007,540)</td>
<td>4</td>
<td>3,586</td>
<td>13 (11–16)</td>
<td>1:5.821</td>
<td>10,652 (7,046–16,049)</td>
</tr>
<tr>
<td></td>
<td>Portugal (10,833,816)</td>
<td>4</td>
<td>1,498</td>
<td>21 (15–30)</td>
<td>1:2.191</td>
<td>4,944 (2,403–10,004)</td>
</tr>
<tr>
<td></td>
<td>Spain (48,563,476)</td>
<td>4</td>
<td>2,458</td>
<td>17 (14–21)</td>
<td>1:3.344</td>
<td>14,522 (9,405–22,331)</td>
</tr>
<tr>
<td>Total/mean</td>
<td></td>
<td>12</td>
<td>7,542</td>
<td>17</td>
<td>1:3.785</td>
<td>30,118</td>
</tr>
<tr>
<td>Eastern Europe (3)</td>
<td>Macedonia (2,100,025)</td>
<td>4</td>
<td>441</td>
<td>8 (4–18)</td>
<td>1:4.814</td>
<td>142 (27–654)</td>
</tr>
<tr>
<td></td>
<td>Russia (142,355,415)</td>
<td>16</td>
<td>8,364</td>
<td>3 (2–4)</td>
<td>1:86.127</td>
<td>1,653 (966–2,813)</td>
</tr>
<tr>
<td></td>
<td>Serbia (7,143,921)</td>
<td>1</td>
<td>1,060</td>
<td>12 (8–19)</td>
<td>1:6.165</td>
<td>1,159 (525–2,510)</td>
</tr>
<tr>
<td>Total/mean</td>
<td></td>
<td>21</td>
<td>9,865</td>
<td>8</td>
<td>35,702</td>
<td>2,954</td>
</tr>
</tbody>
</table>

Total Pi*ZZ genotypes in Europe: 119,594

Abbreviation: CI, confidence interval.

In Australia and New Zealand, Pi*Z has been estimated to be present in approximately 6,000 inhabitants, most being descendants of Anglo Saxon settlers (Table 5).

Discussion

A recent study analyzing the database of studies on the genetic epidemiology of AATD estimated that a total of 181,894 Pi*ZZ genotypes could be found in 97 countries in which data were available. In parallel, the first maps of IDW interpolation in AATD were created to graphically express the frequency of the distribution of the S and Z alleles in countries with real data and make estimates for countries in which no data were available. However, subsequent well-designed studies such as that by an Irish group which analyzed 1,100 individuals randomly selected from the general Irish population and a more recent study in 2,482 Finnish individuals reported much higher frequencies of Pi*Z, which were comparable to the highest frequencies described in neighboring countries.

However, this study followed strict scientific rigor in the selection of studies to evaluate with the application of the established selection and exclusion criteria. Indeed, numerous studies carried out in small samples of minority ethnic groups (eg, Australian aboriginal; Maoris of New Zealand; Asian, Amerindian, Alaska natives; Hawaiian and other Pacific Islander natives from the US, etc.) were excluded from the analysis because they are not representative of the inhabitants of a specific country, despite their unquestionable anthropological interest. In addition, several pioneering studies using starch gel electrophoresis for lab diagnosis...
were ruled out, because this technique has some difficulty for detecting the Z bands of the MZ phenotypes, thereby underestimating the true frequency of the Z allele. Screening studies in which phenotypes were determined in samples with AAT concentrations below a low cut-off point were also discarded, since normal phenotypes are omitted with this methodology while deficiency types are overrepresented. For example, the Sveger et al study in 200,000 newborns used an AAT

### Table 2: Estimated Pi*Z allele frequency (per 1,000), Pi*ZZ gene prevalence (Hardy-Weinberg statistics, 1:x), and number of Pi*ZZ genotypes in 23 North, South, and Central American countries

<table>
<thead>
<tr>
<th>Geographic region (number of countries)</th>
<th>Country (total population)</th>
<th>Number of cohorts</th>
<th>Number of subjects</th>
<th>Pi*Z frequency (95% CI)</th>
<th>Pi*ZZ prevalence (mean)</th>
<th>Number of Pi*ZZ genotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>North America (3)</td>
<td>Canada (35,362,905)</td>
<td>5</td>
<td>4,807</td>
<td>14 (12–16)</td>
<td>1:4,925</td>
<td>7,181 (5,109–10,072)</td>
</tr>
<tr>
<td></td>
<td>US (non-Hispanic whites: 259,196,422)</td>
<td>15</td>
<td>17,032</td>
<td>15 (14–17)</td>
<td>1:4,126</td>
<td>62,820</td>
</tr>
<tr>
<td></td>
<td>Mexico (123,166,749)</td>
<td>2</td>
<td>3,279</td>
<td>6 (4–8)</td>
<td>1:3,141</td>
<td>3,921 (2,002–7,599)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total/mean</td>
<td></td>
<td>22</td>
<td>25,118</td>
<td>12</td>
<td>1:13,489</td>
<td>73,922</td>
</tr>
<tr>
<td>Central America and Caribbean (10)</td>
<td>Costa Rica (4,872,543)</td>
<td>1</td>
<td>3,000</td>
<td>8 (6–10)</td>
<td>1:17,013</td>
<td>286 (157–517)</td>
</tr>
<tr>
<td></td>
<td>Cuba (11,179,995)</td>
<td>1</td>
<td>3,000</td>
<td>7 (5–9)</td>
<td>1:20,408</td>
<td>548 (292–1,018)</td>
</tr>
<tr>
<td></td>
<td>Dominican Republic (10,606,865)</td>
<td>1</td>
<td>3,000</td>
<td>6 (4–9)</td>
<td>1:24,931</td>
<td>425 (219–817)</td>
</tr>
<tr>
<td></td>
<td>El Salvador (6,156,670)</td>
<td>1</td>
<td>3,000</td>
<td>7 (5–9)</td>
<td>1:22,500</td>
<td>274 (144–517)</td>
</tr>
<tr>
<td></td>
<td>Guatemala (15,189,958)</td>
<td>1</td>
<td>3,000</td>
<td>6 (4–8)</td>
<td>1:31,142</td>
<td>488 (242–973)</td>
</tr>
<tr>
<td></td>
<td>Haiti (10,485,800)</td>
<td>1</td>
<td>3,000</td>
<td>1 (0.7–3)</td>
<td>1:444,444</td>
<td>24 (6–92)</td>
</tr>
<tr>
<td></td>
<td>Honduras (8,893,259)</td>
<td>1</td>
<td>3,000</td>
<td>6 (4–9)</td>
<td>1:24,931</td>
<td>357 (184–685)</td>
</tr>
<tr>
<td></td>
<td>Jamaica (2,970,340)</td>
<td>1</td>
<td>3,000</td>
<td>1 (0.7–3)</td>
<td>1:444,444</td>
<td>7 (2–26)</td>
</tr>
<tr>
<td></td>
<td>Nicaragua (5,966,798)</td>
<td>1</td>
<td>3,000</td>
<td>6 (4–9)</td>
<td>1:24,931</td>
<td>229 (123–460)</td>
</tr>
<tr>
<td></td>
<td>Panama (3,705,246)</td>
<td>1</td>
<td>3,000</td>
<td>6 (4–8)</td>
<td>1:27,778</td>
<td>133 (67–261)</td>
</tr>
<tr>
<td>Total/mean</td>
<td></td>
<td>10</td>
<td>30,000</td>
<td>5</td>
<td>1:108,252</td>
<td>2,781</td>
</tr>
<tr>
<td>South America (10)</td>
<td>Argentina (43,886,748)</td>
<td>1</td>
<td>3,000</td>
<td>6 (4–8)</td>
<td>1:26,297</td>
<td>1,669 (852–3,234)</td>
</tr>
<tr>
<td></td>
<td>Bolivia (10,969,649)</td>
<td>1</td>
<td>3,000</td>
<td>6 (4–8)</td>
<td>1:30,298</td>
<td>305 (152–606)</td>
</tr>
<tr>
<td></td>
<td>Brazil (205,823,665)</td>
<td>1</td>
<td>3,000</td>
<td>6 (4–8)</td>
<td>1:31,142</td>
<td>6,162 (3,053–12,297)</td>
</tr>
<tr>
<td></td>
<td>Chile (17,650,114)</td>
<td>1</td>
<td>3,000</td>
<td>6 (4–9)</td>
<td>1:24,931</td>
<td>708 (365–1,360)</td>
</tr>
<tr>
<td></td>
<td>Colombia (47,220,856)</td>
<td>1</td>
<td>3,000</td>
<td>6 (5–9)</td>
<td>1:23,669</td>
<td>1,995 (1,038–3,799)</td>
</tr>
<tr>
<td></td>
<td>Ecuador (16,080,778)</td>
<td>1</td>
<td>3,000</td>
<td>6 (4–8)</td>
<td>1:27,778</td>
<td>579 (293–1,132)</td>
</tr>
<tr>
<td></td>
<td>Paraguay (6,862,812)</td>
<td>1</td>
<td>3,000</td>
<td>7 (5–9)</td>
<td>1:22,500</td>
<td>305 (160–576)</td>
</tr>
<tr>
<td></td>
<td>Peru (30,741,062)</td>
<td>1</td>
<td>3,000</td>
<td>6 (4–9)</td>
<td>1:24,931</td>
<td>229 (123–460)</td>
</tr>
<tr>
<td></td>
<td>Uruguay (3,351,016)</td>
<td>1</td>
<td>3,000</td>
<td>6 (4–8)</td>
<td>1:27,778</td>
<td>121 (61–236)</td>
</tr>
<tr>
<td></td>
<td>Venezuela (30,912,302)</td>
<td>1</td>
<td>3,000</td>
<td>8 (6–10)</td>
<td>1:16,287</td>
<td>1,897 (1,049–3,405)</td>
</tr>
<tr>
<td>Total/mean</td>
<td></td>
<td>10</td>
<td>30,000</td>
<td>6</td>
<td>1:26,002</td>
<td>14,787</td>
</tr>
<tr>
<td>Total Pi*ZZ genotypes in America:</td>
<td></td>
<td>91,490</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note:** The numbers reported for the US refer only to the non-Hispanic white population, due to the lack of reliable studies in other ethnic groups.

**Abbreviation:** CI, confidence interval.

### Table 3: Estimated Pi*Z allele frequency (per 1,000), Pi*ZZ gene prevalence (1:x), and number of Pi*ZZ genotypes in 10 African countries

<table>
<thead>
<tr>
<th>Geographic region (number of countries)</th>
<th>Country (total population)</th>
<th>Number of cohorts</th>
<th>Number of subjects</th>
<th>Pi*Z frequency (95% CI)</th>
<th>Pi*ZZ prevalence (mean)</th>
<th>Number of Pi*ZZ genotypes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa (10)</td>
<td>Cameroon (24,360,803)</td>
<td>1</td>
<td>266</td>
<td>0 (0–7)</td>
<td>1:infinite</td>
<td>0 (0–1,163)</td>
</tr>
<tr>
<td></td>
<td>Cabo Verde Islands (553,432)</td>
<td>1</td>
<td>202</td>
<td>2 (0–16)</td>
<td>1:infinite</td>
<td>3 (0–140)</td>
</tr>
<tr>
<td></td>
<td>Democratic Republic of the Congo (81,331,050)</td>
<td>1</td>
<td>132</td>
<td>0 (0–14)</td>
<td>1:infinite</td>
<td>0 (0–15,659)</td>
</tr>
<tr>
<td></td>
<td>Morocco (33,655,786)</td>
<td>1</td>
<td>439</td>
<td>1 (0–7)</td>
<td>1:770,884</td>
<td>44 (0–1,825)</td>
</tr>
<tr>
<td></td>
<td>Mozambique (25,930,150)</td>
<td>1</td>
<td>274</td>
<td>0 (0–7)</td>
<td>1:infinite</td>
<td>0 (0–1,167)</td>
</tr>
<tr>
<td></td>
<td>Nigeria (186,053,386)</td>
<td>3</td>
<td>564</td>
<td>3 (1–10)</td>
<td>1:79,524</td>
<td>2,340 (240–17,552)</td>
</tr>
<tr>
<td></td>
<td>Republic of the Congo (4,852,412)</td>
<td>2</td>
<td>237</td>
<td>0 (0–8)</td>
<td>1:infinite</td>
<td>0 (0–292)</td>
</tr>
<tr>
<td></td>
<td>Somalia (10,817,350)</td>
<td>1</td>
<td>347</td>
<td>11 (5–23)</td>
<td>1:7,526</td>
<td>1,437 (312–5,986)</td>
</tr>
<tr>
<td></td>
<td>South Africa (54,300,704)</td>
<td>1</td>
<td>732</td>
<td>0 (0–2)</td>
<td>1:infinite</td>
<td>0 (0–344)</td>
</tr>
<tr>
<td></td>
<td>Tunisia (10,937,521)</td>
<td>5</td>
<td>1,838</td>
<td>0 (0–2)</td>
<td>1:infinite</td>
<td>0 (0–11)</td>
</tr>
<tr>
<td>Total/mean</td>
<td></td>
<td>17</td>
<td>5,031</td>
<td>2</td>
<td>–</td>
<td>3,824</td>
</tr>
</tbody>
</table>

**Abbreviation:** CI, confidence interval.
Table 4 Estimated Pi*Z allele frequency (per 1,000), Pi*ZZ gene prevalence (1:x), and number of Pi*ZZ genotypes in 15 Asian countries

<table>
<thead>
<tr>
<th>Geographic region</th>
<th>Country (total population)</th>
<th>Number of cohorts</th>
<th>Number of subjects</th>
<th>Pi*Z frequency (95% CI)</th>
<th>Pi*ZZ prevalence (mean)</th>
<th>Number of Pi*ZZ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle East (1)</td>
<td>Iran (82,801,633)</td>
<td>6</td>
<td>1,348</td>
<td>7 (4–11)</td>
<td>1:22,433</td>
<td>6,391 (1,381–9,579)</td>
</tr>
<tr>
<td>Central Asia (1)</td>
<td>Kazakhstan (18,360,353)</td>
<td>4</td>
<td>417</td>
<td>2 (0–9)</td>
<td>1:173,889</td>
<td>106 (3–1,700)</td>
</tr>
<tr>
<td>East and Southeast</td>
<td>China (1,373,541,278)</td>
<td>19</td>
<td>6,594</td>
<td>0 (0–0.3)</td>
<td>infinite</td>
<td>0 (0–107)</td>
</tr>
<tr>
<td>Asia (10)</td>
<td>Indonesia (258,316,051)</td>
<td>5</td>
<td>724</td>
<td>0 (0–2)</td>
<td>infinite</td>
<td>0 (0–1,672)</td>
</tr>
<tr>
<td></td>
<td>Japan (126,702,133)</td>
<td>7</td>
<td>6,010</td>
<td>0 (0–6)</td>
<td>1:361,120.100</td>
<td>4 (0–57)</td>
</tr>
<tr>
<td></td>
<td>Malaysia (30,949,962)</td>
<td>2</td>
<td>1,279</td>
<td>0 (0–1)</td>
<td>infinite</td>
<td>0 (0–64)</td>
</tr>
<tr>
<td></td>
<td>Mongolia (3,031,330)</td>
<td>1</td>
<td>505</td>
<td>0 (0–4)</td>
<td>infinite</td>
<td>0 (0–40)</td>
</tr>
<tr>
<td></td>
<td>Papua New Guinea (6,791,317)</td>
<td>2</td>
<td>337</td>
<td>0 (0–5)</td>
<td>infinite</td>
<td>0 (0–202)</td>
</tr>
<tr>
<td></td>
<td>Philippines (102,624,209)</td>
<td>1</td>
<td>243</td>
<td>0 (0–7)</td>
<td>infinite</td>
<td>0 (0–5,868)</td>
</tr>
<tr>
<td></td>
<td>Singapore (5,781,728)</td>
<td>2</td>
<td>545</td>
<td>0 (0–3)</td>
<td>infinite</td>
<td>0 (0–66)</td>
</tr>
<tr>
<td></td>
<td>South Korea (50,924,172)</td>
<td>2</td>
<td>487</td>
<td>6 (2–14)</td>
<td>1:26,352</td>
<td>1,932 (320–10,082)</td>
</tr>
<tr>
<td></td>
<td>Thailand (68,200,824)</td>
<td>2</td>
<td>852</td>
<td>11 (7–18)</td>
<td>1:8,043</td>
<td>8,479 (3,266–21,380)</td>
</tr>
</tbody>
</table>

Total/mean

<table>
<thead>
<tr>
<th>Country (total population)</th>
<th>Number of cohorts</th>
<th>Number of subjects</th>
<th>Pi*Z frequency (95% CI)</th>
<th>Pi*ZZ prevalence (mean)</th>
<th>Number of Pi*ZZ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Asia (3)</td>
<td>Nepal (29,033,914)</td>
<td>1</td>
<td>144</td>
<td>0 (0–12)</td>
<td>infinite</td>
</tr>
<tr>
<td>Pakistan (201,995,540)</td>
<td>1</td>
<td>269</td>
<td>9 (3–23)</td>
<td>1:11,578</td>
<td>17,447 (2,372–105,463)</td>
</tr>
<tr>
<td>India (1,266,883,598)</td>
<td>11</td>
<td>1,599</td>
<td>0.6 (0.1–2)</td>
<td>1:2,556,801</td>
<td>495 (15–8,040)</td>
</tr>
</tbody>
</table>

Total/mean

<table>
<thead>
<tr>
<th>Country (total population)</th>
<th>Number of cohorts</th>
<th>Number of subjects</th>
<th>Pi*Z frequency (95% CI)</th>
<th>Pi*ZZ prevalence (mean)</th>
<th>Number of Pi*ZZ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total/mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Pi*ZZ genotypes in Asia</td>
<td>32,154</td>
<td>43</td>
<td>17,576</td>
<td>2 (–)</td>
<td>10,415</td>
</tr>
</tbody>
</table>

Abbreviation: CI, confidence interval.

Table 5 Estimated Pi*Z allele frequency (per 1,000), Pi*ZZ gene prevalence (1:x), and number of Pi*ZZ genotypes in Australia and New Zealand

<table>
<thead>
<tr>
<th>Geographic region</th>
<th>Country (total population)</th>
<th>Number of cohorts</th>
<th>Number of subjects</th>
<th>Pi*Z frequency (95% CI)</th>
<th>Pi*ZZ prevalence (mean)</th>
<th>Number of Pi*ZZ (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia and</td>
<td>Australia (22,992,654)</td>
<td>8</td>
<td>5,536</td>
<td>13 (11–16)</td>
<td>1:5,532</td>
<td>4,126 (2,894–5,695)</td>
</tr>
<tr>
<td>New Zealand</td>
<td>New Zealand (4,474,549)</td>
<td>1</td>
<td>1,303</td>
<td>22 (17–29)</td>
<td>1:2,019</td>
<td>2,216 (1,307–3,731)</td>
</tr>
</tbody>
</table>

Total Pi*ZZ genotypes in Australia and New Zealand: 6,342

Abbreviation: CI, confidence interval.

cut-off of <40%, and although this strategy was useful to detect all the Pi*Z phenotypes, it is not able to detect all the MM and MS, most MZ and SS, and a considerable number of SZ. Finally, several studies conducted with small sample cohorts and very broad limits of 95% CI (such as those performed in the Greek islands, the Balkans, Romania, etc.) were discarded due to the inaccuracy of their results.

Taking all of the above into account, the number of cohorts included was reduced to 224 and the number of countries studied to 65. Nonetheless, the number of estimated ZZ patients in the world (>250,000) substantially increased and the quality of the IDW-interpolation maps significantly improved. This new analysis confirms that the vast majority of Pi*ZZ individuals are found in Europe (~120,000), and in Caucasians of European descent living in the US (~70,000). We also found that there may be a smaller, albeit significant, number of homozygous Z in Canada, Australia, New Zealand, and in some South American countries with populations of European origin.

Despite careful selection of the studies finally included, our estimates may be biased because of the composition of the samples studied, mainly due to newborns as well as several of the inclusion of subjects labeled under the generic title of “healthy unrelated people”, who may include blood donors, school or college students, soldiers, hospital staff, persons undergoing pre-employment or routine medical examinations, active workers, indigenous subjects, and so on. Indeed, only a few samples from a minority of studies were representative of the general population such as a study carried out in a randomly sampled Spanish population of 1,116 subjects enrolled in the municipal population register,77 or the aforementioned studies conducted in Ireland and Finland 23,24. Nevertheless, with our approach, some of the classically reported results notably improved, providing values of Pi*Z frequencies which are significantly higher than previously estimated for Caucasian populations living in the US, Australia, and in northern, western, and central European countries.
It was of note that high Pi*Z frequencies were obtained for several regions in which AATD has not previously been studied (e.g., the Canary Islands), and these regions might be considered as targets for further targeted high-risk case finding screening studies, to identify and provide adequate clinical management to individuals at risk for AATD-associated diseases.

Although some authors have suggested multiple origins of the Pi*Z mutation, it seems very likely that the mutation occurred in Sweden and was spread into the European
continent through the maritime and inland waterways via the Baltic countries. The Pi*Z mutation has a relatively recent origin in Caucasians living in northern Europe, having occurred approximately 2000 years or 66 generations ago. Along the following centuries, emigrants from Europe distributed the gene throughout non-European countries to which they emigrated. Interestingly, the IDW-interpolation maps of Pi*Z distribution developed in this study agree with this hypothesis.

Currently, massive screenings in newborns and healthy individuals from the general population have become obsolete and are only justified in regions with a high prevalence of AATD and significant under-diagnosis. In addition, before ordering the genetic test, the physician should weigh the pros (eg, early diagnosis, preventive measures, genetic counseling, specific replacement therapy) and cons (eg, psychological trauma and family, social, and negative occupational impact, including possible stigmatization) that the detection of this inherited condition could induce in a healthy person.

Conversely, a targeted high-risk case finding approach is highly recommended. The joint American Thoracic Society and European Respiratory Society statement and most scientific societies worldwide recommend targeted testing for AATD in individuals included in groups at risk. This initiative is a relatively feasible and cost-effective tool to increase the detection of AATD in individuals with obstructive lung diseases, chronic liver disease at any age, granulomatosis with polyangiitis, neutrophilic panniculitis, and family members of patients with confirmed AATD.

In conclusion, the updating of the data regarding the prevalence of the Pi*ZZ genotype worldwide provided in this study may be useful for future research and to develop diagnostic approaches in targeted risk groups and optimize the therapeutic resources currently available (this information is available from the authors upon request).

**Authors’ contribution**

This study was designed by IB, PB, and ID. IB provided the database and wrote the draft document. PB, SP-H, FC, CE, and MM collaborated in the redaction and revision of the manuscript. Professor ID designed the figures summarizing the database. All authors contributed toward data analysis, drafting and revising the paper and agree to be accountable for all aspects of the work.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**


