Geographical distribution of COPD prevalence in Europe, estimated by an inverse distance weighting interpolation technique

Ignacio Blanco¹
Isidro Diego²
Patricia Bueno³
Eloy Fernández⁴
Francisco Casas-Maldonado⁵
Cristina Esquinas⁶
Joan B Soriano⁷
Marc Miravitlles⁸

¹Alpha1-Antitrypsin Deficiency Spanish Registry, Lung Foundation Breathe, Spanish Society of Pneumology, Barcelona, ²Materials and Energy Department, School of Mining Engineering, Oviedo University, ³Internal Medicine Department, County Hospital of Jarrio, ⁴Clinical Analysis Laboratory, University Hospital of Cabueñes, Principality of Asturias, ⁵Pneumology Department, University Hospital San Cecilio, Granada, ⁶Pneumology Department, Hospital Universitari Vall d’Hebron, CIBER de Enfermedades Respiratorias (CIBERES), Barcelona, ⁷Instituto de Investigación Hospital Universitario de la Princesa, Universidad Autónoma de Madrid, Madrid, Spain

Abstract: Existing data on COPD prevalence are limited or totally lacking in many regions of Europe. The geographic information system inverse distance weighted (IDW) interpolation technique has proved to be an effective tool in spatial distribution estimation of epidemiological variables, when real data are few and widely separated. Therefore, in order to represent cartographically the prevalence of COPD in Europe, an IDW interpolation mapping was performed. The point prevalence data provided by 62 studies from 19 countries (21 from 5 Northern European countries, 11 from 3 Western European countries, 14 from 5 Central European countries, and 16 from 6 Southern European countries) were identified using validated spirometric criteria. Despite the lack of data in many areas (including all regions of the eastern part of the continent), the IDW mapping predicted the COPD prevalence in the whole territory, even in extensive areas lacking real data. Although the quality of the data obtained from some studies may have some limitations related to different confounding factors, this methodology may be a suitable tool for obtaining epidemiological estimates that can enable us to better address this major public health problem.

Keywords: epidemiology, geographic information system, GIS, Geographic Resources Analysis Support System

Introduction

COPD is a large, growing public health problem. According to the World Health Organization, its expected burden will increase in the coming decades, mostly due to continued exposure to risk factors, population growth and aging, to become the third leading cause of death by 2030.¹

Despite being a major health problem, existing data on COPD prevalence are limited in many countries, and universally high COPD underdiagnosis² or misdiagnosis³ deprives patients and health authorities of the implementation of adequate preventive and therapeutic measures, in order to avoid its potential serious effects and high costs.⁴ In fact, of the 50 sovereign European countries, only 19 (38%) of them have available reliable data on COPD prevalence.⁵

The geographic information system (GIS) is considered a useful tool for reporting the distribution of health-related states, in particular diseases.⁶,⁷ Specifically, the GIS inverse distance weighting (IDW) interpolation (or spatial analysis) technique (an informatics mathematical approach of manipulating spatial information to extract new information and meaning from the original data, using points with known values to estimate values at other unknown points) has proved to be an effective tool in spatial...
distribution estimation of epidemiological variables, when real data are few and widely separated, as might be the case with the COPD prevalence in Europe.5–10

The objective of the present study was to apply the IDW mapping to integrate the existing data from different European areas, in order to represent cartographically the mean percentage of the population affected by COPD in each geographical region of the entire continent, both in studied areas with known data and in many other areas in which no studies have been conducted and consequently with blank data.

Methods

Source of epidemiological studies for COPD prevalence

The vast majority of the data entered in the database used in the IDW interpolation software were taken from the systematic review and meta-analysis recently published by Adeloye et al on the prevalence of COPD across the world from January 1990 to December 2014, based on a confirmed diagnosis of COPD with different formally acknowledged spirometric criteria.5 For the current study, only the variables referring to the place of origin of the samples and the mean COPD prevalence (in percent) of the samples composed of subjects aged ≥40 years were used.11–63 To the initially selected 61 studies, a later one on the prevalence of COPD, as defined by the Global Initiative for Obstructive Lung Disease (GOLD), in a representative sample of northeastern Italy general population was added.64

The application of an age selection criterion of ≥40 years excluded, from the present study, an international study that included young subjects from several cities in Europe aged between 20 and 44 years,65 2 studies performed in 12 Russian regions,66,67 and another study of the Turkish region of Elazig all of them carried out in subjects from the age of 18.68

In addition, to facilitate visual comparisons between Europe and the neighboring regions of Asia and Africa, the available COPD prevalence data from several Middle East and North Africa regions were also included.69–77

IDW multivariate interpolation method

To elaborate colored geographical maps of COPD prevalence, an IDW interpolation process was started through the freely available software QGIS 18.9 in order to link it with the 7.3 64-bit wxPython 3 Geographic Resources Analysis Support System (GRASS).

Then, the geographical longitude and latitude coordinates of the different places were automatically obtained for each point by translating the location of the place of origin of the samples included. The translation was made through the Google Maps application programming interface using the web page of GPS Visualizer (Carleton University Library, Ottawa, ON, Canada).

Then, these geographical coordinates were imported into the database and then exported to a “csv” (comma separated value) text file, which was opened and visualized in QGIS. A shape file containing the contours of all world countries was also loaded.

Finally, the interpolation process was carried out by applying the “v.surf.idw” library included in the GRASS GIS version 7, another free and open source GIS software suite used for geospatial data management and analysis, spatial modeling, image processing, graphics and maps production, and visualization.

The v.surf.idw library provided surface interpolation from vector point data by filling out a raster matrix with interpolated values generated from a set of irregularly spaced data points, using numerical approximation (weighted averaging) techniques. The interpolated value of a cell was determined by the values of the nearby data points and the distance of the cell from those input points.

An implementation of an IDW technique, as defined by Shepard,7 is as follows: the way to find an interpolated value $u$ at a given (arbitrary) point $x$ based on $N$ known samples $u_i = u(x_i)$ for $i=1, 2, \ldots, N$ is:

$$
    u(x) = \begin{cases} 
    \frac{\sum_{i=1}^{N} w_i(x) u_i}{\sum_{i=1}^{N} w_i(x)} & \text{if } d(x, x_i) \neq 0 \text{ for all } i \\
    u_i & \text{if } d(x, x_i) = 0 \text{ for some } i 
    \end{cases}
$$

where $w_i(x) = \frac{1}{d(x, x_i)^p}$

The “$N$ known samples” is the $N$ closest ones to the interpolation point. In our study, the usual $N$ value of 4 ($N=4$) was chosen. This means that a sample point takes the 4 closest pixel centers and linearly interpolates their color values according to their distance from the sample point; $p$ is called the power parameter, which has a default value of 2, which was the value used in this study.

To express the range of values in the maps, a diverging low–mid–high color gradient, with dark blue tones representing the lowest values, yellow and green representing the intermediate ones, and brown and dark red representing the highest ones, was used.
Geographical regions of Europe

Although there is not a universally accepted definition of the boundaries and countries of the European regions, to facilitate the description of findings, we used a simple and conventional classification of the regions of Europe into 5 major subregions, namely Northern Europe, Western Europe, Central Europe, Southern Europe, and Eastern Europe.79

Northern Europe consists approximately of the Europe regions above the 52nd parallel north, and in this geographical area, the following countries have been included for this study: Iceland, Denmark, Norway, Sweden, Finland, Estonia, Latvia, and Lithuania.

In Western Europe, the following countries have been included: the Netherlands, Belgium, Luxembourg, France, Ireland, and UK.

In Central Europe, the following countries have been included: Germany, Austria, Poland, Slovakia, Czech Republic, Hungary, Slovenia, and Switzerland.

In Southern Europe, the following countries have been included: Spain, Portugal, Italy, Greece, Malta, Cyprus, and Turkey.

Although there is no consensus on the precise area covered by Eastern Europe, because the term has a wide range of geopolitical, geographical, cultural, and socioeconomic connotations, the following countries have been included in this region: the Balkan Peninsula countries (Albania, Croatia, Bosnia and Herzegovina, Montenegro, Serbia, and Macedonia), Romania, Moldova, Bulgaria, Ukraine, Belarus, and European Russia.

Statistical analysis

Descriptive statistics were used to summarize study sites. Quantitative variables were expressed as mean and SD, 95% CI for the mean, and range, and they were analyzed using the analysis of variance. A p<0.05 was considered statistically significant. Statistical calculations were performed using the analytical software SPSS 19 (IBM Corp., Armonk, NY, USA).

Results

The cohorts

A total of 62 studies from 19 countries (21 from 5 Northern European countries, 11 from 3 Western European countries, 14 from 5 Central European countries, and 16 from 6 Southern European countries) were selected (Tables 1–4).

As mentioned previously, no prevalence data were available for any of the 12 countries from Eastern Europe.

Numbers of point estimates by individual countries (from high to less) were as follows: Sweden, 8; United Kingdom, 7; Norway, 5; Finland, Germany, and Turkey, 4 each; the Netherlands, Austria, Poland, Portugal, and Spain, 3 each; Italy and Greece, 2; and other remaining countries, 1.

Table 1 Mean prevalence of COPD in 21 studies from 5 Northern Europe countries

<table>
<thead>
<tr>
<th>Country</th>
<th>Source of samples</th>
<th>Setting</th>
<th>Age (years)</th>
<th>Diagnosis criteria</th>
<th>COPD prevalence (%)</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denmark</td>
<td>Copenhagen</td>
<td>Urban</td>
<td>63</td>
<td>GOLD</td>
<td>17</td>
<td>Fabricius et al11</td>
</tr>
<tr>
<td>Denmark</td>
<td>Copenhagen</td>
<td>Urban</td>
<td>50</td>
<td>GOLD</td>
<td>25</td>
<td>Hansen et al12</td>
</tr>
<tr>
<td>Denmark</td>
<td>Aarhus and Aalborg</td>
<td>Mixed</td>
<td>65</td>
<td>GOLD</td>
<td>12</td>
<td>Hansen et al13</td>
</tr>
<tr>
<td>Finland</td>
<td>Helsinki</td>
<td>Urban</td>
<td>49</td>
<td>GOLD/LLN</td>
<td>6</td>
<td>Kainu et al14</td>
</tr>
<tr>
<td>Finland</td>
<td>Lahti</td>
<td>Rural</td>
<td>48</td>
<td>GOLD</td>
<td>5</td>
<td>Jyrki-Tapani et al15</td>
</tr>
<tr>
<td>Finland</td>
<td>Lieto</td>
<td>Rural</td>
<td>73</td>
<td>FEV/FVC = 65%</td>
<td>8</td>
<td>Isoaho et al16</td>
</tr>
<tr>
<td>Finland</td>
<td>Ylöjärvi</td>
<td>Mixed</td>
<td>55</td>
<td>GOLD</td>
<td>5</td>
<td>Kanervisto et al17</td>
</tr>
<tr>
<td>Iceland</td>
<td>Reykjavik</td>
<td>Urban</td>
<td>57</td>
<td>GOLD</td>
<td>18</td>
<td>Benediktsdottir et al18</td>
</tr>
<tr>
<td>Norway</td>
<td>Hordaland county</td>
<td>Mixed</td>
<td>55</td>
<td>GOLD</td>
<td>9</td>
<td>Vaahtevik et al19</td>
</tr>
<tr>
<td>Norway</td>
<td>Bergen municipality</td>
<td>Mixed</td>
<td>61</td>
<td>GOLD</td>
<td>9</td>
<td>Hvidsten et al20</td>
</tr>
<tr>
<td>Norway</td>
<td>Bergen</td>
<td>Urban</td>
<td>58</td>
<td>GOLD</td>
<td>19</td>
<td>Buist et al21</td>
</tr>
<tr>
<td>Norway</td>
<td>Hordaland county</td>
<td>Mixed</td>
<td>42</td>
<td>GOLD</td>
<td>5</td>
<td>Bakke et al22</td>
</tr>
<tr>
<td>Norway</td>
<td>Bergen municipality</td>
<td>Mixed</td>
<td>46</td>
<td>GOLD</td>
<td>5</td>
<td>Johannessen et al23</td>
</tr>
<tr>
<td>Sweden</td>
<td>Uppsala</td>
<td>Urban</td>
<td>59</td>
<td>LLN</td>
<td>10</td>
<td>Danielsson et al24</td>
</tr>
<tr>
<td>Sweden</td>
<td>Lulea</td>
<td>Rural</td>
<td>61</td>
<td>GOLD</td>
<td>14</td>
<td>Lindberg et al25</td>
</tr>
<tr>
<td>Sweden</td>
<td>Lulea</td>
<td>Rural</td>
<td>49</td>
<td>GOLD</td>
<td>14</td>
<td>Lindberg et al26</td>
</tr>
<tr>
<td>Sweden</td>
<td>Norrbotten county</td>
<td>Rural</td>
<td>61</td>
<td>GOLD</td>
<td>11</td>
<td>Lindström et al27</td>
</tr>
<tr>
<td>Sweden</td>
<td>Northern Sweden</td>
<td>Rural</td>
<td>50</td>
<td>BTS</td>
<td>2</td>
<td>Maio et al28</td>
</tr>
<tr>
<td>Sweden</td>
<td>Stockholm</td>
<td>Urban</td>
<td>53</td>
<td>GOLD</td>
<td>10</td>
<td>Hasselgren et al29</td>
</tr>
<tr>
<td>Sweden</td>
<td>Värmland county</td>
<td>Rural</td>
<td>45</td>
<td>BTS</td>
<td>0</td>
<td>Larsson29</td>
</tr>
</tbody>
</table>

Note: The numerical values of age and prevalence have been rounded.

Abbreviations: ATS, American Thoracic Society; BTS, British Thoracic Society; GOLD, Global Initiative for Chronic Obstructive Lung Disease; LLN, lower limit of normal.
Accordingly, a total of 329,413 subjects from the 4 regions of Europe with published data were studied; they were distributed as follows: 39,836 in Northern Europe; 229,083 in Western Europe; 25,553 in Central Europe; and 34,941 in Southern Europe.

The size of the selected samples was significantly different in the 4 European regions, with a mean (minimum and maximal values in parentheses) of 1,897 (548–6,525) subjects in Northern Europe; 20,751 (246–6,126) in Central Europe; and 4 regions of Europe with published data were studied; they were distributed as follows: 39,836 in Northern Europe; 229,083 in Western Europe; 25,553 in Central Europe; and 34,941 in Southern Europe.

The size of the selected samples was significantly different in the 4 European regions, with a mean (minimum and maximal values in parentheses) of 1,897 (548–6,525) subjects in Northern Europe; 20,751 (246–6,126) in Central Europe; and 2,329 (500–9,061) in Southern Europe.

Most of the studies were conducted in isolated cities or localized regions of a particular country, while only a few were carried out on samples representative of the adult general population of an entire country, such as those performed in the Netherlands,\textsuperscript{33} UK,\textsuperscript{39} Switzerland,\textsuperscript{50} Portugal,\textsuperscript{52} Spain,\textsuperscript{54,55} and Cyprus.\textsuperscript{60}

According to the study base of the samples, 23 of them were classified as urban (7 from Northern Europe, 4 from Western Europe, 7 from Central Europe, and 5 from Southern Europe), 9 (6 from Northern Europe, 1 from Central Europe, and 2 from Southern Europe) were classified as rural, and the remaining 30 were catalogued as mixed.

The mean age of the subjects from the 62 selected samples was 55.9 years (SD = 6.1), without significant differences between regions (Table 5).

In 80% of the studies, the diagnosis of COPD was made by using the GOLD spirometric criteria of the

Table 2 Mean prevalence of COPD in 11 studies from 3 Western Europe countries

<table>
<thead>
<tr>
<th>Country</th>
<th>Source of samples</th>
<th>Setting</th>
<th>Age (years)</th>
<th>Diagnosis criteria</th>
<th>COPD prevalence (%)</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td>Paris</td>
<td>Urban</td>
<td>63</td>
<td>GOLD</td>
<td>7</td>
<td>Roche et al\textsuperscript{34}</td>
</tr>
<tr>
<td>the Netherlands</td>
<td>Maastricht</td>
<td>Urban</td>
<td>59</td>
<td>GOLD/LLN</td>
<td>19</td>
<td>Vanzetereen et al\textsuperscript{31}</td>
</tr>
<tr>
<td>the Netherlands</td>
<td>Rotterdam</td>
<td>Urban</td>
<td>69</td>
<td>GOLD</td>
<td>12</td>
<td>van Durme et al\textsuperscript{32}</td>
</tr>
<tr>
<td>the Netherlands</td>
<td>General population of the whole country</td>
<td>Mixed</td>
<td>58</td>
<td>GOLD</td>
<td>4</td>
<td>Afonso et al\textsuperscript{33}</td>
</tr>
<tr>
<td>UK</td>
<td>Birmingham</td>
<td>Mixed</td>
<td>58</td>
<td>GOLD</td>
<td>9</td>
<td>Jordan et al\textsuperscript{34}</td>
</tr>
<tr>
<td>UK</td>
<td>Ashington</td>
<td>Urban</td>
<td>59</td>
<td>GOLD</td>
<td>25</td>
<td>Melville et al\textsuperscript{35}</td>
</tr>
<tr>
<td>UK</td>
<td>Greater Belfast area</td>
<td>Mixed</td>
<td>55</td>
<td>ATS</td>
<td>7</td>
<td>Murtagh et al\textsuperscript{36}</td>
</tr>
<tr>
<td>UK</td>
<td>Glasgow</td>
<td>Urban</td>
<td>51</td>
<td>GOLD</td>
<td>24</td>
<td>Maio et al\textsuperscript{12}</td>
</tr>
<tr>
<td>UK</td>
<td>Barton-upon-Humber</td>
<td>Mixed</td>
<td>68</td>
<td>GOLD</td>
<td>10</td>
<td>Dickinson et al\textsuperscript{37}</td>
</tr>
<tr>
<td>UK</td>
<td>Manchester</td>
<td>Mixed</td>
<td>64</td>
<td>GOLD</td>
<td>26</td>
<td>Renwick and Connolly\textsuperscript{38}</td>
</tr>
<tr>
<td>UK</td>
<td>General population of the whole country</td>
<td>Mixed</td>
<td>56</td>
<td>GOLD</td>
<td>13</td>
<td>Shah et al\textsuperscript{39}</td>
</tr>
</tbody>
</table>

Table 3 Mean prevalence of COPD in 14 studies from 5 Central Europe countries

<table>
<thead>
<tr>
<th>Country</th>
<th>Source of samples</th>
<th>Setting</th>
<th>Age (years)</th>
<th>Diagnosis criteria</th>
<th>COPD prevalence (%)</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>Salzburg</td>
<td>Mixed</td>
<td>58</td>
<td>GOLD</td>
<td>17</td>
<td>Weiss et al\textsuperscript{40}</td>
</tr>
<tr>
<td>Austria</td>
<td>Salzburg</td>
<td>Urban</td>
<td>59</td>
<td>GOLD</td>
<td>26</td>
<td>Schirhoffer et al\textsuperscript{41}</td>
</tr>
<tr>
<td>Austria</td>
<td>Vienna</td>
<td>Urban</td>
<td>56</td>
<td>GOLD</td>
<td>20</td>
<td>Maio et al\textsuperscript{12}</td>
</tr>
<tr>
<td>Germany</td>
<td>Hannover</td>
<td>Mixed</td>
<td>58</td>
<td>GOLD</td>
<td>13</td>
<td>Geldmacher et al\textsuperscript{43}</td>
</tr>
<tr>
<td>Germany</td>
<td>Dusseldorf</td>
<td>Urban</td>
<td>59</td>
<td>GOLD</td>
<td>9</td>
<td>Ginger et al\textsuperscript{42}</td>
</tr>
<tr>
<td>Germany</td>
<td>Munich</td>
<td>Urban</td>
<td>59</td>
<td>GOLD</td>
<td>21</td>
<td>Maio et al\textsuperscript{12}</td>
</tr>
<tr>
<td>Germany</td>
<td>Berlin</td>
<td>Urban</td>
<td>41</td>
<td>GOLD</td>
<td>16</td>
<td>Maio et al\textsuperscript{12}</td>
</tr>
<tr>
<td>Poland</td>
<td>Warsaw</td>
<td>Mixed</td>
<td>57</td>
<td>LLN</td>
<td>9</td>
<td>Bedaraike et al\textsuperscript{44}</td>
</tr>
<tr>
<td>Poland</td>
<td>Malopolska province</td>
<td>Mixed</td>
<td>56</td>
<td>GOLD</td>
<td>22</td>
<td>Niwankowska-Moglinicka et al\textsuperscript{45}</td>
</tr>
<tr>
<td>Poland</td>
<td>Warsaw</td>
<td>Urban</td>
<td>49</td>
<td>GOLD</td>
<td>11</td>
<td>Pityczukiewicz et al\textsuperscript{46}</td>
</tr>
<tr>
<td>Poland</td>
<td>Lublin region</td>
<td>Rural</td>
<td>50</td>
<td>GOLD</td>
<td>11</td>
<td>Paprzycki et al\textsuperscript{47}</td>
</tr>
<tr>
<td>Poland</td>
<td>Bytom</td>
<td>Mixed</td>
<td>59</td>
<td>GOLD</td>
<td>6</td>
<td>Siatkowska et al\textsuperscript{48}</td>
</tr>
<tr>
<td>Slovenia</td>
<td>Sezana (Carso)</td>
<td>Urban</td>
<td>56</td>
<td>GOLD</td>
<td>10</td>
<td>Grzetic-Romcevic et al\textsuperscript{49}</td>
</tr>
<tr>
<td>Switzerland</td>
<td>General population of the whole country</td>
<td>Mixed</td>
<td>54</td>
<td>GOLD</td>
<td>5</td>
<td>Bridevaux et al\textsuperscript{50}</td>
</tr>
</tbody>
</table>

Note: The numerical values of age and prevalence have been rounded.
Abbreviations: ATS, American Thoracic Society; GOLD, Global Initiative for Chronic Obstructive Lung Disease; LLN, lower limit of normal.
postbronchodilator (post-BD) fixed ratio $FEV_1/FVC <0.70$; in the remaining 12 studies, other spirometric criteria (not post-BD) endorsed by different scientific societies were used, such as the lower limit of normal, the British Thoracic Society, the American Thoracic Society, and European Respiratory Society criteria.

## COPD prevalence

The global mean prevalence of COPD (expressed in percent) was $12.38$ (SD $6.2$) without significant differences among the 4 European regions with available data (Table 4).

The highest values of prevalence were found in 2 samples from the cities of Manchester and Salzburg (26% each), followed by those of Ashington and Copenhagen (with 25% each); the COPD subjects from Glasgow, the Po Delta area of Italy, and the southern Poland region of Malopolska (with 24%, 23%, and 22%, respectively); and the COPD subjects from the metropolitan areas of Stockholm and Vienna (with 20% each).

Prevalence rates between $\geq 15\%$ and $\leq 20\%$ were found in the cities of Bergen and Maastricht and in the Turkish province of Adana (all with 19%), in the metropolitan areas of Reykjavik and Thessaly (both 18%), and in Berlin (with 16%).

Prevalence ranges between $\geq 10\%$ and $\leq 15\%$ were found in a sample from Luleå (in the northernmost county of Sweden) with 14%, followed by 2 samples from Barton-upon-Humber (England) and Hannover (Germany) both with 13% and the patients’ samples from the Danish cities of Aarhus and Aalborg and the Dutch city of Rotterdam (with 12%). Prevalence value of 11% was found in Warsaw, in the Polish region of Lublin, and in the Swedish counties of Norrbotten and Örebro. Prevalence values slightly $>10\%$ were found in Slovenia, Spain, and Portugal and in the general population of the UK.

In practically all of the remaining places with available data, prevalence ranged from 5 to $<10\%$. Anecdotally, a young rural cohort, with a mean age of 45 years, from the

## Table 4 Mean prevalence of COPD in 15 studies from 6 Southern Europe countries

<table>
<thead>
<tr>
<th>Country</th>
<th>Source of samples</th>
<th>Setting</th>
<th>Age (years)</th>
<th>Diagnosis criteria</th>
<th>COPD prevalence (%)</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Portugal</td>
<td>Lisbon</td>
<td>Urban</td>
<td>63</td>
<td>GOLD</td>
<td>14</td>
<td>Bárbara et al$^{12}$</td>
</tr>
<tr>
<td>Portugal</td>
<td>General population of the whole country</td>
<td>Mixed</td>
<td>56</td>
<td>GOLD</td>
<td>9</td>
<td>Cardoso et al$^{13}$</td>
</tr>
<tr>
<td>Portugal</td>
<td>Porto</td>
<td>Urban</td>
<td>58</td>
<td>GOLD</td>
<td>11</td>
<td>Mascarenhas et al$^{14}$</td>
</tr>
<tr>
<td>Spain</td>
<td>Multicenter study on 11 cities</td>
<td>Mixed</td>
<td>57</td>
<td>GOLD</td>
<td>10</td>
<td>Miravitlès et al$^{15}$</td>
</tr>
<tr>
<td>Spain</td>
<td>Multicenter study on 7 geographical areas</td>
<td>Mixed</td>
<td>54</td>
<td>ERS</td>
<td>9</td>
<td>Peña et al$^{16}$</td>
</tr>
<tr>
<td>Spain</td>
<td>Canary Islands</td>
<td>Urban</td>
<td>55</td>
<td>GOLD</td>
<td>7</td>
<td>Cabrera López et al$^{17}$</td>
</tr>
<tr>
<td>Italy</td>
<td>Po Delta area</td>
<td>Rural</td>
<td>57</td>
<td>GOLD</td>
<td>23</td>
<td>Viegi et al$^{18}$</td>
</tr>
<tr>
<td>Italy</td>
<td>Verona</td>
<td>Mixed</td>
<td>56</td>
<td>GOLD/LLN</td>
<td>12</td>
<td>Guerriero et al$^{19}$</td>
</tr>
<tr>
<td>Greece</td>
<td>Exoichi, Thessaloniki</td>
<td>Mixed</td>
<td>52</td>
<td>GOLD</td>
<td>6</td>
<td>Sichletidis et al$^{20}$</td>
</tr>
<tr>
<td>Greece</td>
<td>Thessaly region</td>
<td>Rural</td>
<td>54</td>
<td>GOLD</td>
<td>18</td>
<td>Minas et al$^{21}$</td>
</tr>
<tr>
<td>Cyprus</td>
<td>General population of the whole island</td>
<td>Mixed</td>
<td>54</td>
<td>GOLD</td>
<td>5</td>
<td>Zachariades et al$^{22}$</td>
</tr>
<tr>
<td>Turkey</td>
<td>Kocaeli city</td>
<td>Urban</td>
<td>56</td>
<td>GOLD</td>
<td>13</td>
<td>Arslan et al$^{23}$</td>
</tr>
<tr>
<td>Turkey</td>
<td>Erzurum</td>
<td>Rural</td>
<td>55</td>
<td>GOLD</td>
<td>5</td>
<td>Erdogan et al$^{24}$</td>
</tr>
<tr>
<td>Turkey</td>
<td>Malaya region</td>
<td>Mixed</td>
<td>47</td>
<td>ATS/ERS</td>
<td>7</td>
<td>Gunen et al$^{25}$</td>
</tr>
<tr>
<td>Turkey</td>
<td>Adana</td>
<td>Mixed</td>
<td>54</td>
<td>GOLD</td>
<td>19</td>
<td>Buist et al$^{26}$</td>
</tr>
</tbody>
</table>

Notes: The numerical values of age and prevalence have been rounded.

Abbreviations: ATS, American Thoracic Society; ERS, European Respiratory Society; GOLD, Global Initiative for Chronic Obstructive Lung Disease; LLN, lower limit of normal.

## Table 5 Ages of selected samples, descriptive statistics

<table>
<thead>
<tr>
<th>Region</th>
<th>Number of studies</th>
<th>Mean (years)</th>
<th>SD</th>
<th>95% CI (lower limit)</th>
<th>95% CI (upper limit)</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern Europe</td>
<td>21</td>
<td>54.8</td>
<td>7.7</td>
<td>51.3</td>
<td>58.3</td>
<td>42.0</td>
<td>73.0</td>
</tr>
<tr>
<td>Western Europe</td>
<td>11</td>
<td>60.3</td>
<td>5.6</td>
<td>56.5</td>
<td>64.0</td>
<td>50.9</td>
<td>69.5</td>
</tr>
<tr>
<td>Central Europe</td>
<td>14</td>
<td>55.0</td>
<td>5.2</td>
<td>52.0</td>
<td>58.1</td>
<td>40.1</td>
<td>59.5</td>
</tr>
<tr>
<td>Southern Europe</td>
<td>16</td>
<td>55.2</td>
<td>3.2</td>
<td>53.5</td>
<td>56.9</td>
<td>47.5</td>
<td>62.8</td>
</tr>
<tr>
<td>Global</td>
<td>62</td>
<td>55.9</td>
<td>6.1</td>
<td>54.4</td>
<td>57.5</td>
<td>40.8</td>
<td>73.0</td>
</tr>
</tbody>
</table>

Notes: Interregions: p=0.071, no significant differences (ANOVA).

Abbreviation: ANOVA, analysis of variance.
The greatest numbers of black spots are located in the Iberian Peninsula, Great Britain, Central Europe, and the coastal regions bordering the North and the Baltic Seas. In contrast, spots are in general scattered and scanty in most of the remaining parts of the continent and absolutely absent in the Mediterranean islands of Corsica, Sardinia, Sicily and Greek islands, southern Italy, Balkan Peninsula, and regions belonging to Eastern Europe.

Some black spots can also be seen in some nearby countries of North Africa and Middle East (for guidance).

In Figure 2, the numerical values of the corresponding prevalence rates have been added. The numerical values appear with a decimal and are the same reproduced (rounded) in Tables 1–4.

In the map of Figure 3, both black spots and numerical figures have been removed to allow a clearer view of the colors with which the different regions were colored.

### Table 6 COPD prevalence in Europe, descriptive statistics

<table>
<thead>
<tr>
<th>Region</th>
<th>Number of studies</th>
<th>Mean prevalence (%)</th>
<th>SD</th>
<th>95% CI (lower limit)</th>
<th>95% CI (upper limit)</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern Europe</td>
<td>21</td>
<td>11.5</td>
<td>5.9</td>
<td>8.8</td>
<td>14.1</td>
<td>2.1</td>
<td>25.4</td>
</tr>
<tr>
<td>Western Europe</td>
<td>11</td>
<td>14.2</td>
<td>8.0</td>
<td>8.8</td>
<td>19.6</td>
<td>3.9</td>
<td>26.4</td>
</tr>
<tr>
<td>Central Europe</td>
<td>14</td>
<td>14.1</td>
<td>6.5</td>
<td>10.4</td>
<td>17.9</td>
<td>5.0</td>
<td>26.1</td>
</tr>
<tr>
<td>Southern Europe</td>
<td>16</td>
<td>10.8</td>
<td>5.6</td>
<td>7.8</td>
<td>13.8</td>
<td>4.9</td>
<td>23.4</td>
</tr>
<tr>
<td>Global</td>
<td>62</td>
<td>12.4</td>
<td>6.2</td>
<td>10.8</td>
<td>14.0</td>
<td>2.1</td>
<td>26.4</td>
</tr>
</tbody>
</table>

Note: Interregions: p=0.3460, no significant differences (ANOVA).

Abbreviation: ANOVA, analysis of variance.

Swedish county of Värmland showed the lowest prevalence of COPD in Europe, with a percentage slightly >2%.

**IDW interpolation maps**

In the IDW interpolation map of Figure 1, the prevalence rates of COPD of all European countries are represented by a red–blue scale. Black spots indicate the places of origin of the samples.

In a simplified way, since there is no exact equivalence between the qualitative data provided by the color scale and the quantitative ones provided by the numerical values, it can be interpreted with reasonable certainty that red tones represent the maximum values ≥20%; brown and orange colors symbolize high but lower values of ~15%–20%; shades of yellow and green tones represent intermediate values of ~10%–15%; and blue tones represent minimal values of 0%–5%.

In Figure 1, the COPD mean prevalence in Europe is shown using an IDW interpolation map. The COPD prevalence is represented by a red–blue scale. Black spots indicate the places of origin of the data obtained from 62 selected studies. Some black spots are also shown in some nearby countries of North Africa and Middle East for guidance.

Abbreviation: IDW, inverse distance weighting.
A very high prevalence represented by red tones (≈20%-27%) appears in practically all Great Britain, except in the south of the island that presents brownish colors in the west and yellow in the center and east consistent with an intermediate prevalence. Reddish colors, indicative of a very high prevalence, are also shown in the southeast of the Scandinavian Peninsula, the Island of Amager (Denmark), southeast of Germany, western and southern Poland, Austria, 

Figure 2 COPD mean prevalence in Europe.

Note: The IDW interpolation map shows the prevalence values of COPD at the points of origin of the patients, both in Europe and in some neighboring countries in North Africa and the Middle East.

Abbreviation: IDW, inverse distance weighting.

Figure 3 COPD mean prevalence in Europe.

Notes: In this IDW interpolation map, spots and numerical values have been removed to allow a clearer view of the colors with which the different regions appear colored. Although there is not an exact equivalence between the qualitative data provided by the color scale and the numerical values obtained from published studies, there is an approximate correlation between them, which allows interpreting findings of the mapping with reasonable certainty.

Abbreviation: IDW, inverse distance weighting.
Czech Republic, and Slovakia (these last 2 countries without studies and therefore without real numerical values).

Brown and orange tones (approximately equivalent to a range of numerical values between 15% and 20%) appear in Iceland, Hungary, large areas of Italy (except the regions of Lombardy, Piedmont, Liguria, and Calabria, which appear in yellow color), Croatia, Bosnia and Herzegovina, Serbia, and Greece, as well as in the province of Adana (southern Turkey). Of note, there are no real numerical values for the Balkan countries that were colored by means of the IDW mapping.

Most of the remaining regions of Europe are stained in yellow and green colors, indicative of an intermediate prevalence of ~10%–15%; the Danish peninsula of Jutland, most of Sweden, Norway, and northern Finland, the Baltic seacoast countries and the Baltic islands of Gotland and Aland, Belorussia, most of the lands of Poland, the western half of Germany, the coastal regions of Western Europe and Great Britain bordering the North Sea and the English Channel, Ireland, Portugal, Spain, the Mediterranean major islands, Slovenia, Ukraine, Romania, Bulgaria, European Russia, Montenegro, Albania, Macedonia, and the Anatolian Peninsula are included in this group. Remarkably, many of these regions were stained by IDW interpolation in the absence of real numerical data.

Blue areas (indicative of the prevalence of around or <5%) are scarce and isolated; Switzerland, the island of Cyprus, some cities of Spain, such as Vigo, Burgos, Cáceres, and Seville, scattered areas of the Scandinavian peninsula and Finland, and the Erzurum Province of eastern Turkey are among them.

Discussion
Existing data on COPD prevalence are in general insufficient and irregularly distributed in many regions of Europe, in some of which there is a striking scarcity or even an absolute lack of epidemiological information about this important health issue.

To assess the COPD prevalence throughout the European continent and its surrounding islands, an efficient method of proven utility, known as IDW multivariate interpolation (or spatial analysis), which is usually applied to various disciplines concerned with the earth’s surface, such as cartography, geography, hydrology, climatology, and agriculture, was used in the present study in order to generate a continuous map of COPD prevalence.

Basically, IDW interpolation uses data points with known values to estimate other values at unknown locations, in order to create a raster surface consisting of an array of cells (pixels) arranged in rows and columns (or a grid) in which each cell contains a value representing some prevalence information, finally covering a whole area. The IDW method is based on the assumption that the attribute value of an unsampled point is the weighted average of known values within the 4 nearest data points. IDW assumes that each measured point has a local influence that diminishes with distance, weighting the points closer to the prediction location greater than those located farther away, which means that weights are inversely related to the distances between the sampled locations and the predicted ones.7

The IDW interpolation is a simple and intuitive method, easy to handle and interpret, relatively fast for calculations, and easy to compute, which also has the advantage of allowing to easily renewing the database, with a flexibility of the inclusion of studies newly identified, that are more recent, or with a better study design. Similar flexibility applies to the exclusion of others of less scientific quality to automatically reconstructing new, updated maps. In addition, IDW has a mathematical precision that makes it always exact, but this advantage can be a disadvantage if the actual data were erroneous or inaccurate. Since it works automatically creating new points from others supposedly accurate, it is crucial to introduce reliable data; otherwise the final results will also be inaccurate or erroneous.

For the present study, the numerical data of prevalence rates were obtained with the best available evidence in publicly accessible scientific databases, and only surveys constituted by patients of at least ≥40 years of age and a confirmed diagnosis of COPD with spirometry were selected for the analysis.

Limitations
However, the authors of the present study are aware that the analysis may have some limitations due to a number of confounding factors frequently associated with cross-sectional studies, and sometimes these confounding variables could have led to spurious associations. For example, the sampling approach varied from study to study or even among sites in the same study. In addition, the size of the samples was highly variable, with significant differences between regions, and some of them were not proportional to the general population of the studied region. Remarkably, the lack or scarcity of data in extensive regions of Europe makes the scientific hypotheses from which some conclusions were drawn less reliable.

There are also some (though not statistically significant) differences in the age distribution between regions of
Europe and even within the same region of the same country, and since in spirometry-based studies the prevalence of COPD increases with age, the lack of homogeneity in this variable may be the reason for some of the erratic values.

On the other hand, when modeling relative risks or ORs from data obtained from a given population, results may not be applicable to populations with different levels of exposure, as the relations may not be linear across all exposure levels. Regrettably, in our study, there was no proportionality between the number of urban and rural cohorts, and thus, studies performed in populations with high levels of smoking or occupational air pollutants exposure may not be generalizable to those of populations with lower levels of exposure.

Another limitation may be related to the differences in case definitions. The definition based on post-BD FEV1/FVC <0.70 was used in the large majority (80%) of studies, but this still does not address all possible sources of variation in case definition, since this fixed ratio criterion may potentially overdiagnose COPD in the elderly and it may underdiagnose COPD in younger patients. Besides, a number of technical issues could have affected the estimates, such as the quality of the commercial models of the spirometers used, the level of training of the operators, and the process of collection and storage of spirometry measurements.

Finally, an intrinsic limitation of GIS is that geographic proximity does not take into account borders, which often within Europe mirror a different language or dialect. Common language is a well-established factor for unevenly distributing genes, related to or not to health, and many other social determinants.

Nevertheless, in spite of all these criticisms, it should be emphasized that GIS unquestionably offers a useful new tool in epidemiology, which makes it ever easier to connect spatially referenced physical and social phenomena to population patterns of health, disease, and well-being.

Therefore, in our opinion, the easy-to-interpret and eye-catching impacting results of this technique may contribute to enhancing public health surveillance as well as delineating the prevalence of COPD in Europe and in other countries, to increase the number of doctors who treat this type of patients and the health authorities and national governments’ interest in strengthening regulations to address occupational and environmental risk factors, regulate tobacco use, and improve public awareness; to all together increase the rate of diagnoses and to implement the application of preventive and therapeutic measures to avoid potential side effects and unnecessary costs; and to reduce the high global COPD burden in coming years.

Disclosure
The authors report no conflicts of interest in this work.

References


