



Contents lists available at ScienceDirect

## Journal of Infection

journal homepage: [www.elsevier.com/locate/jinf](http://www.elsevier.com/locate/jinf)

# Geographical distribution of cystic fibrosis carriers as population genetic determinant of COVID-19 spread and fatality in 37 countries

Chiara Gabbi<sup>a,\*</sup>, Alessandra Renieri<sup>b,c,d</sup>, Birgitta Strandvik<sup>a</sup><sup>a</sup> Department of Biosciences and Nutrition, Karolinska Institutet, Neo, SE-141 83, Huddinge, Sweden<sup>b</sup> Medical Genetics, University of Siena, Siena, Italy<sup>c</sup> Med Biotech Hub and Competence Center, Department of Medical Biotechnologies, University of Siena, Italy<sup>d</sup> Genetica Medica, Azienda Ospedaliero-Universitaria Senese, Italy

## ARTICLE INFO

## Article history:

Accepted 7 June 2022

Available online 11 June 2022

## Keywords:

SARS-CoV-2

COVID-19

CFTR

Cystic fibrosis

CF carriers

## SUMMARY

COVID-19 has shown a relevant heterogeneity in spread and fatality among countries together with a significant variability in its clinical presentation, indicating that host genetic factors may influence COVID-19 pathogenicity. Indeed, subjects carrying single pathogenic variants of the Cystic Fibrosis (CF) Transmembrane Conductance Regulator (*CFTR*) gene – i.e. CF carriers – are more susceptible to respiratory tract infections and are more likely to undergo severe COVID-19 with higher risk of 14-day mortality.

Given that CF carrier prevalence varies among ethnicities and nations, an ecological study in 37 countries was conducted, in order to determine to what extent the diverse CF carrier geographical distribution may have affected COVID-19 spread and fatality during the first pandemic wave.

The CF prevalence in countries, as indicator of the geographical distribution of CF carriers, significantly correlated in a direct manner with both COVID-19 prevalence and its Case Fatality Rate (CFR). In a regression study weighted for the number of tests performed, COVID-19 prevalence positively correlated with CF prevalence, while CFR correlated with population percentage older than 65-year, cancer and CF prevalence. Multivariate regression model also confirmed COVID-19 CFR to be associated with CF prevalence, after adjusting for elderly, cancer prevalence, and weighting for the number of tests performed.

This study suggests a putative contribution of population genetics of *CFTR* in understanding the spatial distribution of COVID-19 spread and fatality.

© 2022 The Author(s). Published by Elsevier Ltd on behalf of The British Infection Association.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

## Introduction

The COVID-19 pandemic has shown a significant variability in spread and fatality among geographical areas.<sup>1</sup> While demographic, socioeconomic and environmental factors explain to some extent such geo-epidemiological variability,<sup>2–4</sup> it's conceivable that the genetic profile of populations may also influence COVID-19 spatial spread and fatality.<sup>5</sup>

At individual level, several genetic variants have been identified in association with severe COVID-19.<sup>6–10</sup> In particular, we have shown with the GEN-COVID consortium,<sup>11</sup> that individuals carrying single pathogenic variants of the Cystic Fibrosis Transmembrane Conductance Regulator (*CFTR*) gene – i.e. CF carriers – are more likely to undergo the severe form of COVID-19 and have higher risk of 14-day mortality.<sup>11</sup> In addition, ultra-rare variants of *CFTR* have been identified in association with COVID-19 severity.<sup>9</sup> The *CFTR*

gene encodes for a chloride and bicarbonate channel expressed on the apical membrane of epithelial cell.<sup>12</sup> If both copies of *CFTR* exhibit pathogenic variants, individuals develop Cystic Fibrosis (CF), a severe congenital disease characterized by high viscosity of secreted fluids and by an abnormal inflammatory cascade.<sup>12</sup>

CF prevalence varies significantly among ethnicities – and therefore nations – being more frequent in Caucasian (1/2500) and Ashkenazi Jews (1/2300) compared to African Americans (1/15,000) and Asians (1/35,000).<sup>13</sup> Importantly, given that CF is an autosomal recessive inherited disorder, CF prevalence in a specific country is considered an indicator of the prevalence of CF carriers among those inhabitants.<sup>12</sup> CF carriers are more likely to undergo respiratory tract infections<sup>14,15</sup> and have higher risk of the severe form of COVID-19.<sup>11</sup> Therefore, our hypothesis is that the CF prevalence of different populations could be an additional factor that may contribute to explain the high variability of COVID-19 spread and CFR among countries.<sup>1</sup>

Here we present an ecological study<sup>16</sup> in 37 countries on the prevalence of officially registered CF patients<sup>12</sup> in relation to

\* Corresponding author.

E-mail address: [chiaragabbi@virgilio.it](mailto:chiaragabbi@virgilio.it) (C. Gabbi).

COVID-19 prevalence and CFR during the first pandemic wave. Such time frame is chosen in order to avoid possible confounding factors from the immune protection generated by prior exposure, as occurs in subsequent waves.<sup>17</sup> The final aim is to investigate a putative scaled impact of *CFTR* population genetics on COVID-19 geoepidemiological variability.

## Methods

Data on the crude CF cases were collected from 37 countries through the official registry reports, available online, of the Cystic Fibrosis Foundation (<https://www.cff.org>), the European Cystic Fibrosis Society ([www.ecfs.eu](http://www.ecfs.eu)), and the Cystic Fibrosis Canada ([www.cysticfibrosis.ca](http://www.cysticfibrosis.ca)), all recording CF patients for over 20 years. The prevalence of COVID-19 cases was obtained from the World Health Organization at the end of the first pandemic wave, i.e. on June 1, 2020. Population data were taken from the Population Reference Bureau (PRB) ([www.prb.org](http://www.prb.org)). Data on population density, age, and cancer prevalence were obtained from the datasets of the World Bank ([www.data.worldbank.org](http://www.data.worldbank.org)). Data on the tests performed and the stringency index (SI) – a composite measure of 9 indicators of national responses, including school/workplaces closure, travel bans – were taken from the datasets of Our World in Data (<https://ourworldindata.org/>). The SI of the first pandemic wave for each country was considered as the average of the daily index from January 21 to June 1, 2020. The relationship between variables was explored using Pearson's correlation methods after logarithmic transformation. Multiple linear regression method was used to study the relationship between COVID-19 prevalence/CFR and explanatory factors like the population age. When indicated, data were weighted for the number of tests performed (per 1000 inhabitants) or by the number of inhabitants in each country. Statistical analysis was performed with the aid of STATA/IC 15.0 software on a Mac workstation.

## Results

### COVID-19 prevalence correlates with CF prevalence

In the 37 studied countries COVID-19 prevalence directly and significantly correlated with the prevalence of CF in each country (Fig. 1A;  $r = 0.45$ ;  $p = 0.0052$ ). Ireland, United Kingdom, Belgium were the countries with the highest prevalence of both CF patient (2.64, 1.56, 1.29 /10,000 inhabitants respectively), and COVID-19 cases.

Considering that substantial variation in national responses to the pandemic occurred and that the reported numbers of COVID-19 cases were influenced by the number of diagnostic tests executed and by the population size, further linear regression studies, weighted for each of those values, and adjusted for the SI, were performed. A significant direct correlation was confirmed between COVID-19 and CF prevalences when data were weighted for the number of diagnostic tests performed (Fig. 1B;  $p = 0.0000$ ) also after adjusting for the SI ( $R^2 = 0.32$ ;  $\beta = 0.52$ ;  $CI = 0.39–0.65$ ;  $p = 0.000$ ). Weighting for the number of inhabitants (Fig. 1C;  $p = 0.0003$ ), and additionally adjusting for the SI confirmed the finding ( $R^2 = 0.39$ ;  $\beta = 0.51$ ;  $CI = 0.26–0.75$ ;  $p = 0.000$ ).

In addition, in a regression model weighted for the number of tests performed, no significant correlation was found between COVID-19 prevalence and the percentage of population older than 65 years ( $\beta = -0.9$ ;  $CI = -2.38–0.58$ ;  $p = 0.225$ ).

### COVID-19 CFR correlates with CF prevalence

The CFR of COVID-19 significantly correlated in a direct manner with CF prevalence (Fig. 1D;  $r = 0.47$ ,  $p = 0.0032$ ) also

in a regression model weighted for the number of inhabitants (Fig. 1E;  $p = 0.0068$ ) or performed tests ( $\beta = 0.33$ ;  $CI = 0.05–0.6$ ;  $p = 0.022$ ). In addition, COVID-19 CFR correlated with the percentage of population older than 65 years (Fig. 1F;  $r = 0.47$ ;  $p = 0.036$ ) and the total cancer prevalence ( $\beta = 1.1$ ;  $CI = 0.35–1.85$   $p = 0.005$ ) in a regression weighted for the number of tests performed.

In a multiple linear regression model the correlation between CF prevalence and COVID-19 CFR was confirmed to be significant after adjusting for the prevalence of people older than 65 years, the cancer prevalence and weighting the analysis for the number of tests performed ( $R^2 = 0.40$ ;  $\beta = 0.32$ ;  $CI = 0.01–0.64$ ;  $p = 0.043$ ).

## Discussion

This report shows that the country-specific CFR and prevalence of COVID-19 during the first pandemic wave directly correlated with the CF carrier prevalence in 37 countries.

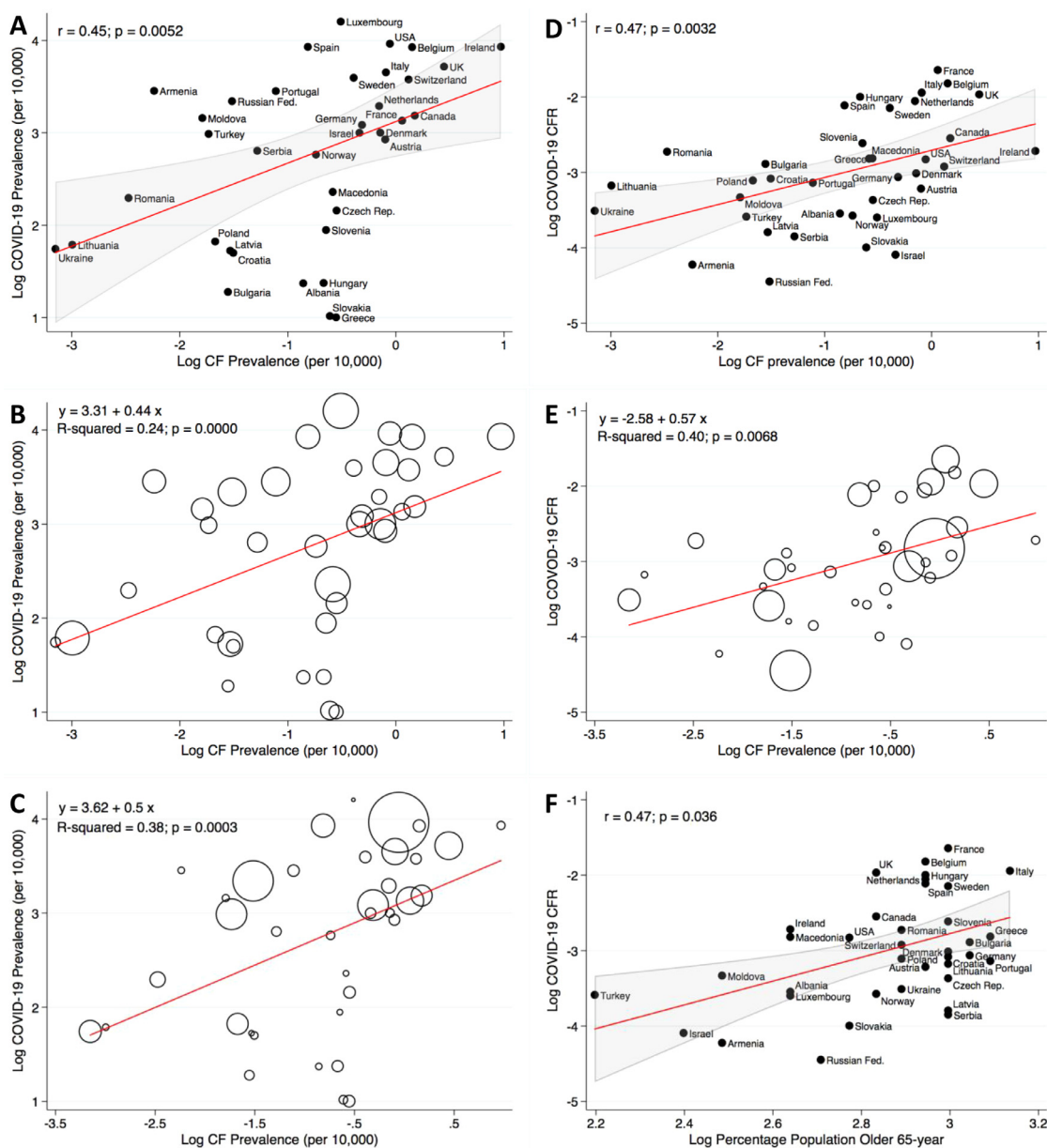
In our previous report on a cohort of 874 COVID-19 patients from Italy, CF carriers had a significant Hazard Ratio for 14-day mortality of 3.1 and deaths of CF carriers accounted for 7.27% of the deaths of the entire cohort.<sup>11</sup> Considering that Italy reported 33,415 COVID-19 deaths during the first pandemic wave<sup>1</sup> about 2429 of them would be potentially related to CF carriers. In addition, the prevalence of CF carriers in Italy is estimated to be around 1 out of 31 inhabitants,<sup>18</sup> resulting in a projected population of about 2 million Italian and, theoretically, up to 24 million European CF carriers. As they are more likely to undergo respiratory tract infections<sup>14,15</sup> and severe COVID-19,<sup>11</sup> the burden of COVID-19 morbidity and mortality attributable to the CF-carrier status may be relevant.

Many factors are certainly involved in the etiology of severe COVID-19, but preclinical evidences regarding coronavirus attacking the cell enlighten a putative pathogenic role of *CFTR*. *In-vitro* studies have shown that the Spike protein of SARS-CoV-2 binds to *CFTR* itself inhibiting its activity<sup>19</sup> and cleavage sites of the coronavirus 3CL<sup>pro</sup> proteinase have been predicted in the intracellular region of *CFTR* by computational models.<sup>20</sup> Accordingly, when WT animals are infected by SARS-CoV-2, a reduced mRNA expression of *CFTR* is found in lungs,<sup>21</sup> together with an increased expression of the epithelial sodium channel (ENaC) that facilitates sodium absorption,<sup>21</sup> a pattern similar to the drastic *CFTR* impairment seen in CF.<sup>22</sup>

In this context it's important to note that while CF carriers have been reported to have a COVID-19 related mortality of 10%,<sup>11</sup> individuals affected by CF do not undergo severe COVID-19 and their highest reported mortality is 3.87%<sup>23</sup> raising the hypothesis that CF patients may have intrinsic protective factors.<sup>23</sup>

Our study has some limitations. First, being based on country-level data, the described associations don't demonstrate a causal link between outcomes and predictors. However, the previous report on severe COVID-19 and higher death risk in CF carriers<sup>11</sup> together with the described preclinical evidences,<sup>19–21</sup> enlighten a possible contribution of *CFTR* in COVID-19 pathogenesis. Detailed laboratory studies in cells expressing single pathogenic *CFTR* variants are required to understand the eventual molecular mechanism behind our findings. Second, testing and data reporting practices might differ substantially among countries – especially during the first pandemic wave and an underestimation of COVID-19 deaths may have occurred. However, COVID-19 data were taken from the WHO<sup>1</sup> and other official registries, ensuring the highest reliability and uniformity possible.

In conclusion, this report suggests a possible contribution of the *CFTR* genetic profile of populations in understanding the country variability of COVID-19 spread and fatality.



**Fig. 1.** CF Prevalence Correlation with COVID-19 Prevalence and CFR. **Panel A** shows a scatterplot with line of best fit (red) and the 95% confidence interval (gray), illustrating a significant direct Pearson's correlation between the prevalence of CF and COVID-19 prevalence in 37 countries.  $r = 0.45$ ,  $p = 0.0052$ . In **panel B** the liner regression model is weighted for the number of tests performed per 1000 inhabitants in each country. A positive linear correlation between CF and COVID-19 prevalences is confirmed ( $R^2 = 0.24$ ;  $p = 0.0000$ ). The size of the marker represents the number of tests performed per 1000 inhabitants. **Panel C** illustrates a linear regression model weighted for the countries' inhabitants. A positive correlation between CF and COVID-19 prevalences is seen ( $R^2 = 0.38$ ;  $p = 0.0003$ ). The size of the marker represents the number of inhabitants in each country. **Panel D** shows a scatterplot with line of best fit (red) and the 95% confidence interval (gray), illustrating a significant direct Pearson's correlation between the prevalence of CF and the COVID-19 CFR.  $r = 0.47$ ,  $p = 0.0032$ . **Panel E** illustrates a liner regression model weighted for the countries' inhabitants. A positive correlation between CF prevalence and COVID-19 CFR is seen ( $R^2 = 0.40$ ;  $p = 0.0068$ ). The size of the marker represents the number of inhabitants in each country. In **panel F** is a scatterplot with line of best fit (red) and the 95% confidence interval (gray), illustrating a significant direct Pearson's correlation between COVID-19 CFR and the percentage of inhabitants older than 65-year.  $r = 0.47$ ,  $p = 0.036$ . Prevalence is expressed per 10,000 inhabitants and all data are in logarithmic scale.

**Declaration of Competing Interest**

Authors declare no conflict of interest.

**Acknowledgments**

This work was partially supported by the Swedish Cystic Fibrosis Association (CG). Authors are grateful to the European Cystic Fibrosis Society, the Cystic Fibrosis Foundation and the Cystic Fibrosis Canada for their work in registering CF patients.

**References**

1. World Health Organization. Coronavirus disease (COVID-19) Weekly Epidemiological Update and Weekly Operational Update. WHO; 2020. [www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports](http://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports).
2. Gaudart J, Landier J, Huiart L, Legendre E, Lehot L, Bendiane MC, et al. Factors associated with the spatial heterogeneity of the first wave of COVID-19 in France: a nationwide geo-epidemiological study. *Lancet Public Health* 2021;6(4):e222–31. doi:10.1016/S2468-2667(21)00006-2.
3. Cao Y, Hiyoshi A, Montgomery S. COVID-19 case-fatality rate and demographic and socioeconomic influencers: worldwide spatial regression analysis based on country-level data. *BMJ Open* 2020;10(11):43560. doi:10.1136/bmjopen-2020-043560.
4. Piovani D, Christodoulou MN, Hadjidemetriou A, Pantavou K, Zaza P, Bagos BP,

- et al. Effect of early application of social distancing interventions on COVID-19 mortality over the first pandemic wave: an analysis of longitudinal data from 37 countries. *J Infect* 2021;**82**(1):133–42. doi:10.1016/j.jinf.2020.11.033.
5. Aung AK, Aitken T, Teh BM, Yu C, Ofori-Asenso R, Chin KL, et al. Angiotensin converting enzyme genotypes and mortality from COVID-19: an ecological study. *J Infect* 2020;**81**(6):961–5. doi:10.1016/j.jinf.2020.11.012.
  6. Pairo-Castineira E, Clohisey S, Klaric L, Bretherick AD, Rawlik K, Pasko D, et al. Genetic mechanisms of critical illness in COVID-19. *Nature* 2020. doi:10.1038/s41586-020-03065-y.
  7. Benetti E, Tita R, Spiga O, Ciolfi A, Birolo G, Bruselles A, et al. ACE2 gene variants may underlie interindividual variability and susceptibility to COVID-19 in the Italian population. *Eur J Hum Genet* 2020;**28**(11):1602–14. doi:10.1038/s41431-020-0691-z.
  8. Baldassarri M, Picchiotti N, Fava F, Fallerini C, Benetti E, Daga S, et al. Shorter androgen receptor polyQ alleles protect against life-threatening COVID-19 disease in European males. *EBioMedicine* 2021;**65**. doi:10.1016/j.ebiom.2021.103246.
  9. Fallerini C, Picchiotti N, Baldassarri M, Zguro K, Daga S, Fava F, et al. Common, low-frequency, rare, and ultra-rare coding variants contribute to COVID-19 severity. *Hum Genet* 2022;**141**(1):147–73. doi:10.1007/s00439-021-02397-7.
  10. Croci S, Venneri MA, Mantovani S, Fallerini C, Benetti E, et al., Gen-Covid Multicenter Study. The polymorphism L412F in *TLR3* inhibits autophagy and is a marker of severe COVID-19 in males. *Autophagy* 2021:1–11. doi:10.1080/15548627.2021.1995152.
  11. Baldassarri M, Fava F, Fallerini C, Daga S, Benetti E, Zguro K, et al. Severe COVID-19 in hospitalized carriers of single CFTR pathogenic variants. *J Pers Med* 2021;**11**(6):558. doi:10.3390/jpm11060558.
  12. Elborn JS. Cystic fibrosis. *Lancet* 2016;2519–31. doi:10.1016/S0140-6736(16)00576-6.
  13. Palomaki GE, Fitzsimmons SC, Haddow JE. Clinical sensitivity of prenatal screening for cystic fibrosis via CFTR carrier testing in a United States panethnic population. *Genet Med* 2004;**6**(5):405–14. doi:10.1097/01.GIM.0000139505.06194.39.
  14. Miller AC, Comellas AP, Hornick DB, Stoltz DA, Cavanaugh JE, Gerke AK, et al. Cystic fibrosis carriers are at increased risk for a wide range of cystic fibrosis-related conditions. *Proc Natl Acad Sci U S A* 2020;**117**(3):1621–7. doi:10.1073/pnas.1914912117.
  15. Polgreen PM, Brown GD, Hornick DB, Ahmad F, London B, Stoltz DA, et al. CFTR heterozygotes are at increased risk of respiratory infections: a population-based study. *Open Forum Infect Dis* 2018;**5**(11):1–7. doi:10.1093/ofid/ofy219.
  16. Morgenstern H. Ecologic studies in epidemiology: concepts, principles, and methods. *Rev Public Health* 1995;**16**:61–81. doi:10.1146/annurev.pu.16.050195.000425.
  17. Mathews JD, McBryde ES, McVernon J, Pallaghy PK, McCaw JM. Prior immunity helps to explain wave-like behaviour of pandemic influenza in 1918–9. *BMC Infect Dis* 2010;**10**(1):128. doi:10.1186/1471-2334-10-128.
  18. Picci L, Cameran M, Marangon O, Marzenta D, Ferrari S, Frigo AC, et al. A 10-year large-scale cystic fibrosis carrier screening in the Italian population. *J Cyst Fibros* 2010;**9**(1):29–35. doi:10.1016/j.jcf.2009.10.003.
  19. Caohuy H, Eidelman O, Chen T, Yang Q, Walton NW, Pollard HB. Inflammation in the COVID-19 airway is due to inhibition of CFTR signaling by the SARS-CoV-2 Spike protein. *bioRxiv* 2022. doi:10.1101/2022.01.18.476803v2.
  20. Lars K, Kiemer L, Lund O, Brunak S, Blom N. Coronavirus 3CLpro proteinase cleavage sites: possible relevance to SARS virus pathology. *BMC Bioinform* 2004;**5**:1–9. doi:10.1186/1471-2105-5-72.
  21. Morales L, Oliveros JC, Enjuanes L, Sola I. Contribution of host miRNA-223-3p to SARS-CoV-induced lung inflammatory pathology. *mBio* 2022. doi:10.1128/mbio.03135-21.
  22. Strandvik B. Is the ENaC dysregulation in CF an effect of protein-lipid interaction in the membranes? *Int J Mol Sci* 2021;**22**(5):2739. doi:10.3390/ijms22052739.
  23. Mathew HR, Choi MY, Parkins MD, Fritzlner MJ. Systematic review: cystic fibrosis in the SARS-CoV-2/COVID-19 pandemic. *BMC Pulm Med* 2021;**21**(1):173. doi:10.1186/s12890-021-01528-0.