COVID-19 severity is associated with population-level gut microbiome variations

Eva Lymberopoulos[1,2†], Giorgia Isabella Gentili[1†],
Sanjay Budhdeo[1,3,4] and Nikhil Sharma[1*]

[1]The Sharma Lab, Department of Clinical and Movement Neurosciences, Queen Square Institute of
Neurology, University College London, London, England, [2]Centre for Doctoral Training in
AI-London enabled Healthcare Systems, Institute of Health Informatics, University College London,
London, England, [3]National Hospital for Neurology and Neurosurgery, Queen Square, London,
England, [4]School of Biomedical Engineering & Imaging Sciences, Faculty of Life Sciences &
Medicine, King's College London, London, England

The human gut microbiome interacts with many diseases, with recent small
studies suggesting a link with COVID-19 severity. Exploring this association at
the population-level may provide novel insights and help to explain differences
in COVID-19 severity between countries. We explore whether there is an
association between the gut microbiome of people within different countries
and the severity of COVID-19, measured as hospitalisation rate. We use data
from the large (n = 3,055) open-access gut microbiome repository
curatedMetagenomicData, as well as demographic and country-level
metadata. Twelve countries were placed into two groups (high/low)
according to COVID-19 hospitalisation rate before December 2020
(ourworldindata.com). We use an unsupervised machine learning method,
Topological Data Analysis (TDA). This method analyses both the local
geometry and global topology of a high-dimensional dataset, making it
particularly suitable for population-level microbiome data. We report an
association of distinct baseline population-level gut microbiome signatures
with COVID-19 severity. This was found both with a PERMANOVA, as well as
with TDA. Specifically, it suggests an association of anti-inflammatory bacteria,
including *Bifidobacteria* species and *Eubacterium rectale*, with lower severity,
and pro-inflammatory bacteria such as *Prevotella copri* with higher severity.
This study also reports a significant impact of country-level confounders,
specifically of the proportion of over 70-year-olds in the population, GDP,
and human development index. Further interventional studies should examine
whether these relationships are causal, as well as considering the contribution
of other variables such as genetics, lifestyle, policy, and healthcare system. The
results of this study support the value of a population-level association design
in microbiome research in general and for the microbiome-COVID-19
relationship, in particular. Finally, this research underscores the potential of
TDA for microbiome studies, and in identifying novel associations.

KEYWORDS

gut microbiome, COVID-19, population health, topological data analysis, probiotics

1 Introduction

Despite the vast advances made in researching COVID-19 over the first two years of the pandemic, many of the mechanisms underpinning its severity and inter-individual variations in disease course are incompletely understood. This made it difficult to predict the likely severity within a given population and to plan patient flow at a regional and national level. Here, we explore the link between variations in the gut microbiome between countries and COVID severity. This is important as the gut microbiome is amenable to manipulation at the population level, potentially providing a novel intervention in future outbreaks and pandemics.

There are significant variations in the response to COVID-19. For example, while the disease predominantly affects the respiratory tract, it has been found to also cause a range of other symptoms, including non-specific (such as fever and headaches), neurological (such as anosmia and dysgeusia) or gastrointestinal (GI) symptoms (such as abdominal pain, vomiting, and diarrhoea) (Bostancıklıoğlu, 2020; Molina-Mora et al., 2022). Additionally, patients can experience rapid worsening associated with the so-called cytokine storm (Ragab et al., 2020) which is treated during a hospital stay (Ramiro et al., 2020; Cron et al., 2021). The factors and their combination that lead to such severe courses of COVID-19 are not understood yet. Several risk factors have been discovered, spanning multiple axes. One of these axes is socioeconomic factors. These include gender, ethnicity, BMI, household income, and especially the intersection of such vulnerabilities (Booth et al., 2021; Chen and Krieger, 2021; Pijls et al., 2021). For example, people of minority ethnic status in the most deprived neighbourhoods have the worst prognosis when infected and are most likely to get infected in the first place (Berkowitz et al., 2020). The second axis relates to weakness of the immune system, for example in the elderly, chronically ill, disabled, and immunocompromised (Mohammed et al., 2020; Baek et al., 2021; Booth et al., 2021). The microbiome is a further variable that appears to affect the severity of a COVID-19 infection but has received less attention so far (Segal et al., 2020).

The gut microbiome is a collection of an estimated 38 trillion microbes (Sender et al., 2016) colonising mostly the intestines, which have vast effects on host health. By metabolising food, they produce many key metabolites for physiological functioning across several organs and body systems. There is a particular interaction of the microbiome with immune pathways (Thaiss et al., 2016; Levy et al., 2017). As such, it has become the focus of research on inflammatory diseases, including chronic infections, autoimmune diseases, metabolic, and even psychiatric and neurodegenerative diseases (Wang et al., 2017). The microbiome has a key translational advantage over other research targets in these diseases as it is easily measured and modifiable by diet.

The gut microbiome is an intriguing candidate for a unifying risk factor, or even mediator, of COVID-19. When considering the epidemiological triad, it is part of both host factors, and the effect of environmental factors on the host.

For one, there is a complex relationship with social risk factors. For example, the microbiome varies by ethnicity (Gupta et al., 2017; Deschasaux et al., 2018), socioeconomic status (Bowyer et al., 2019), and BMI (Falony et al., 2016). It also changes with age (Yatsunenko et al., 2012; Zapata and Quagliarello, 2015; Badal et al., 2020). Several medical conditions which increase the risk of COVID-19 are also associated with changes to gut microbiome composition, leading to the hypothesis that the microbiome is a mediator for the disease-risk association (Kim, 2021). Another piece of evidence of the involvement of the gut microbiome is the reported GI symptomatology in COVID-19 (Villapol, 2020), likely through gut epithelial cells which express ACE-2 receptors, a major entryway for SARS-CoV-2 (Wang et al., 2020; Guney and Akar, 2021). Further, SARS-CoV-2 is found in the stool of COVID-19 patients (Wu et al., 2020). The gut microbiome is a major modulator of the innate immune system, with metabolites of the gut affecting intestinal barrier integrity, as well as the immune system itself (Jiao et al., 2020). These metabolites have been found to directly up- or downregulate cytokines, which lends some bacteria the labels anti-inflammatory or pro-inflammatory (Wang et al., 2020). In physiological homeostasis, these bacteria are thought to regulate each other through competition and quorum sensing, resulting in normal functioning. However, in an imbalanced state, these control mechanisms are reduced, and pro-inflammatory bacteria and pathogens can increase exponentially, a phenomenon called 'blooming'. An altered microbiome is also often associated with a reduction in the diversity of bacteria in the gut (Belizário and Faintuch, 2018), though this is likely an oversimplification.

Immune-modulatory effects of the gut microbiome have been found to play a role in small studies of COVID-19 patients. A study from the USA compared patients of various levels of severity and controls (n = 48), reporting a significant decrease in anti-inflammatory and increase of pro-inflammatory species in symptomatic patients. Specifically, they found a reduction of *Bifidobacterium* and *Faecalibacterium* in patients that was inversely predictive of disease severity, as well as an increase in *Bacteroides* which was also predictive of severity (Hazan et al., 2022). In hospitalised patients, a similar decrease in commensal symbionts and an increase in opportunistic pathogens (compared to healthy controls) was reported. Of note, it did not resolve when the infection cleared (Zuo et al., 2021). Additionally, these patients show an increase in gastrointestinal symptoms (Bozkurt and Bilen, 2021a). While such reports show that the microbiome differs in more severely ill patients, they cannot explain the

causality of this association. Does COVID-19 cause microbiome changes, or does variation in the microbiome at baseline influence severity?

Some evidence suggests that gut bacteria have a direct influence on COVID-19 severity through regulating cytokines. This is a well-known function of the gut microbiome, mostly by regulating the T-helper 17 (Th17) and T-regulatory (Treg) cell balance Omenetti and Pizarro, 2015. For example, a retrospective study of a probiotic booster administered to hospitalised patients reduced the number of inpatient days and the mortality rate, and improved radiological findings Bozkurt and Bilen, 2021b. This study also observed reduced levels of Interleukin-6 (IL6), a key pro-inflammatory cytokine of the innate immune response that has been implicated in increased severity of COVID-19 Coomes and Haghbayan, 2020 and the COVID-19 associated cytokine storm Ragab et al., 2020. Such evidence suggests that the microbiome could in fact have a causal relationship with COVID-19 severity through regulating the immune response.

Small sample size is a significant limitation to the current literature. Studies have typically been conducted at the individual level with small sample sizes (n<30). This introduces major issues as there is considerable inter-subject variability driven in part by the huge scale of the microbiome, as well as its complex nonlinear relationships. This is exacerbated by microbiome analyses that focus on differences in abundance between groups. Additionally, these studies are underpowered for detecting key confounding factors such as demographic, socioeconomic, or medical factors.

Examining the gut microbiome at a population level may provide an elegant solution to address the limitations of small populations and could provide a target for public health interventions. For one, this approach allows sample sizes of several thousand participants. Additionally, by shifting the focus from individuals to populations, we can consider how the microbiome might affect the differences in COVID-19 severity that were observed across countries. Indeed, it is known that the microbiome of the population varies between countries. This is influenced by differences in local water and soil microbiomes, diet and other lifestyle factors, antibiotic prescription patterns, hygiene, and pollution (Karlsson et al., 2014; Gupta et al., 2019; Lymberopoulos et al., 2021). By considering the link between the gut microbiome and COVID-19 severity, it is possible that the microbiome of a population influences the level of COVID-19 severity experienced in that population. In fact, the impact of the COVID-19 pandemic varied strongly between countries. While the rate of infection was influenced by governmental response and strictness of public health measures such as lockdowns (Huang et al., 2021; Arnold et al., 2022), as well as other population metrics (Sorci et al., 2020), there remains a difference in experienced severity across countries with similar rates of infection.

While there is now evidence to suggest a link between the gut microbiome and COVID-19 severity, current evidence is limited by the size and design of individual-level studies which can be addressed with a population-based approach. We aim to explore whether country-level differences in the microbiome before the pandemic can explain a degree of the country-level differences in COVID-19 severity, as measured by hospitalisation rate. In addition to standard microbiome analysis, we employ an unsupervised machine learning tool - Topological Data Analysis - which is particularly appropriate for large microbiome datasets and can detect even subtle non-linear and mixed effects (Liao et al., 2019).

We hypothesise that there will be significant differences in microbiome composition between countries based on COVID-19 severity grouping that will be evident with standard analysis as well as TDA. We expect to find specific microbiome signatures to be associated with such groups of low and high severity. Specifically, we hypothesise that TDA will find more complex associations above and beyond those results from standard analysis. Finally, we expect to be able to discern the effects of relevant confounders.

2 Methods

2.1 Data

2.1.1 Microbiome data

This study used an open microbiome data source, curatedMetagenomicData (cMD) accessed through BioConductor (Pasolli et al., 2017). This dataset used 16S rRNA amplicon sequencing. Because of its highly conserved nature, this technique now represents the standard application to the investigation of the microbiome (Gupta et al., 2019).

For this study, healthy participants over the age of 2 years old were included. In the first 2 years of life, the gut microbiome undergoes drastic shaping to its structure and functionality, particularly between 6 months and 2 years when solid foods are introduced, resulting in the final maturation stages of the microbiome (Robertson et al., 2019). Considering this, we excluded participants under 2 years old as their microbiome is still in the maturation phases and may therefore report misleading conclusions. Children over the age of 2 were not excluded as their microbiome is already more stable. Additionally, they are counted as part of hospitalisation data, and we matched the population of each dataset accordingly. The curat dataset comprised both age data and age category, which included newborns, school-age, children, and adults. In the case where subjects had missing data for both age and age category, they were excluded. The age group comprising newborns were fully excluded, as well as participants under the age of two. For ease of analysis, all participants' age group was collected as

metadata. Additionally, sample-level demographic data of gender and BMI were used.

We further excluded samples from countries with less than 100 samples, to ensure a level of representativeness of the microbiome data.

This filtering yielded 3,055 healthy gut microbiome samples from 12 countries. These were Canada, Denmark, Estonia, Finland, France, Great Britain, Israel, Italy, Netherlands, Spain, Sweden, United States of America.

2.1.2 COVID-19 severity data

For each of these countries, COVID-19 severity data was obtained from Our World in Data (https://github.com/owid/covid-19-data), using hospitalisations as the preferred metric representing the number of COVID-19 patients in hospital on a given day. We chose severity to be operationalised by hospitalisations as it is the measure most likely to reflect the hypothesised immune-modulating effects of the microbiome, as opposed to factors relating to public policy or the healthcare system. These would have more influence on metrics of cases, test-positivity-rate, ICU beds, or fatality rates.

We restricted the data to comprise the hospitalisations before the initiation of the vaccination programme (before December 2020), to avoid potential confounding effects of vaccine roll-out. For each country, the number of absolute hospitalisations was summed over the whole period and then calculated as hospitalisations per 100,000 population.

Using a weighted average, included countries were grouped into 'high' and 'low' COVID-19 severity groups. These include Canada, Denmark, Estonia, Finland, the Netherlands, and the USA in the high severity group, and France, Great Britain, Italy, Israel, Spain, and Sweden in the low severity group. This grouping is reflective of the data distribution as indicated in Figure 1.

Finally, country-level confounding variables were collected that could present relevant confounders in the association between the microbiome and COVID-19 severity. The first of these is the proportion of over 70-year-olds in the population, as the older population is more at risk of severe COVID-19 (Damayanthi et al., 2021). The second includes two measures of wealth and industrialisation level: the Gross Domestic Product (GDP) and the Human Development Index (HDI). These three measures were chosen as there is previous evidence that they are associated with COVID-19 fatality differences between countries which suggests potential relevance to COVID-19 severity in general (Sorci et al., 2020). Additionally, both age and wealth influence the microbiome (Bowyer et al., 2019; Badal et al., 2020), meaning there is potential for influence on the relationship between the microbiome and COVID-19, as well.

2.2 Analysis

Standard microbiome analysis was performed in R version 4.1.2 (2021-11-01) using the *curatedMetagenomicData* (3.2.3), *phyloseq* (1.38.0), and *microbiome* (1.16.0) packages. Unsupervised Machine Learning with TDA was performed in *Python 3.1.4* in a Jupyter Notebook environment, using mainly the *pandas* (1.2.4), *NumPy* (1.19.5), and *tmap* (1.0) *libraries and seaborne* (0.11.1) and *matplotlib* (3.3.4) for visualisations. All code is available on GitHub (*via* thesharmalab-team/microbiome_covid).

2.2.1 Standard microbiome analysis

The composition of the gut microbiome by phylum was investigated using bar plots to observe differences in the overall relative abundance of phyla between the COVID-19 groups.

Alpha diversity compromises a mathematical measure of diversity within the microbiome population. This used the Shannon index to account for both species abundance and evenness ($p < 0.05$ considered significant). This metric was compared within all individual included countries and by Covid-19 grouped countries.

Beta diversity analysis was run to measure the difference in microbiome composition between the high and low severity groups. This was analysed with the Bray-Curtis dissimilarity matrix and visualised with Principal Coordinate Analysis (PCoA).

A permutational multivariate analysis of variance (PERMANOVA) was adopted to identify the top genus-level taxa which significantly differ between the COVID-19 groups. This further allows for the overall significance testing between the microbiome composition and the COVID-19 groups (permutations = 999).

2.2.2 Tmap

Topological Data Analysis (TDA) is an unsupervised machine learning tool based on algebraic topology and differential geometry (Carlsson, 2009). This means that it analyses the topological structure, or shape, of high-dimensional data without sacrificing complexity. TDA has many attributes that make it attractive for microbiome research, particularly its robustness to noise and sensitivity to signal. This enables it to detect even small and non-linear effects, as demonstrated in a previous study which showed that it outperformed standard microbiome tools (Liao et al., 2019). That study combined the TDA algorithm Mapper (Singh et al., 2007) with a modified version of the Spatial Analysis of Functional Enrichment algorithm (SAFE) (Baryshnikova, 2016) into a single library, *tmap*, for integrated stratification

and association analysis of population microbiome data. This library is also used in this study.

The Mapper algorithm is a clustering method based on simplicial complexes that represent high-dimensional data, such as the microbiome data of this study, as a network graph. The input is a point cloud, with each microbiome sample constituting one data point. Through filtering, covering, and clustering in the high-dimensional space, data points are converted into nodes and edges. Nodes represent groups of samples with similar microbiomes which are close in the high-dimensional space, while edges represent overlap between these groups.

The SAFE algorithm, as used here, maps variables such as host metadata or bacterial taxa, onto the microbiome network as node attributes and generates subnetwork scores. These so-called SAFE scores exist for each node but can be transformed to a network-wide score called SAFE enriched score which enables comparison between variables, as well as co-enrichment in which relationships between variables are analysed.

The procedure used to implement tmap in this study is as follows. The first step is the optimisation of the cover ratio, which includes hand-tuning the parameters of overlap and resolution, as well as the minimum number of samples required to form a node. The cover ratio determines how sparse versus dense the graph network representation of the microbiome data will be. Based on these, the final Mapper network graph is calculated. Here, hand-tuning resulted in using a minimum number of samples of 7, resolution of 100 and overlap of 1.1. Then, all microbiome taxa and metadata variables are enriched using the SAFE algorithm. This is then used for metadata stratification, meaning the comparison of

enrichment for groups. The final step is co-enrichment analysis, in which pairwise co-enrichment is calculated and a strict 0.5th percentile significance cut-off applied.

3 Results

3.1 COVID-19 severity

Table 1 shows the severity of COVID-19 by total hospitalisations and hospitalisation per 100,000 before the commencing of the vaccination programme.

Based on the data reported in Table 1, we allocated the 12 countries into two groups: high and low Covid-19 hospitalisation per 100,000. The separating limit used for the Covid-19 groups is 1842.75 hospitalisations per 100,000. As displayed in Figure 1, the low severity group includes Canada, Denmark, Estonia, Finland, the Netherlands, and the United States of America. The high severity group includes Spain, France, Great Britain, Israel, Italy, and Sweden. While the tiles are coloured according to the observed variables, the area of each tile represents the microbiome sample size for each country.

Presented in Table 2 are the demographic factors for the cMD dataset (Gender, Age Group, and BMI) observed for the high and low Covid-19 grouping.

3.2 Standard microbiome analysis

In assessing the relative abundances of phyla, an increase in Bacteroidetes and Proteobacteria and a decrease in Firmicutes

TABLE 1   Displayed in the table are the country-level variables for the 12 countries included in the analysis.

Country-level data

Location	Proportion over 70 years old	GDP	HDI	Hospitalisations per 100,000	Frequency
Canada	10.80	44,017.59	0.93	803.68	75
Denmark	12.33	46,682.51	0.94	517.23	494
Spain	13.80	34,272.36	0.90	1,968.95	210
Estonia	13.49	29,481.25	0.89	868.86	24
Finland	13.26	40,585.72	0.94	258.09	61
France	13.08	38,605.67	0.90	5,621.29	61
Great Britain	12.53	39,753.24	0.93	2,830.04	250
Israel	7.36	33,132.32	0.92	2,086.07	900
Italy	16.24	35,220.08	0.89	5,364.83	59
Netherlands	11.88	48,472.54	0.94	1,401.75	470
Sweden	13.43	46,949.28	0.95	2,366.49	257
USA	9.73	54,225.45	0.93	1,816.21	194

This includes hospitalisation per 100,000 and the number of microbiome samples (Frequency) which are the basis of the grouping into low and high severity. Confounding factors are also included for each country: proportion over 70, GDP and HDI. These were obtained from the COVID-19 Dataset from Our World in Data.
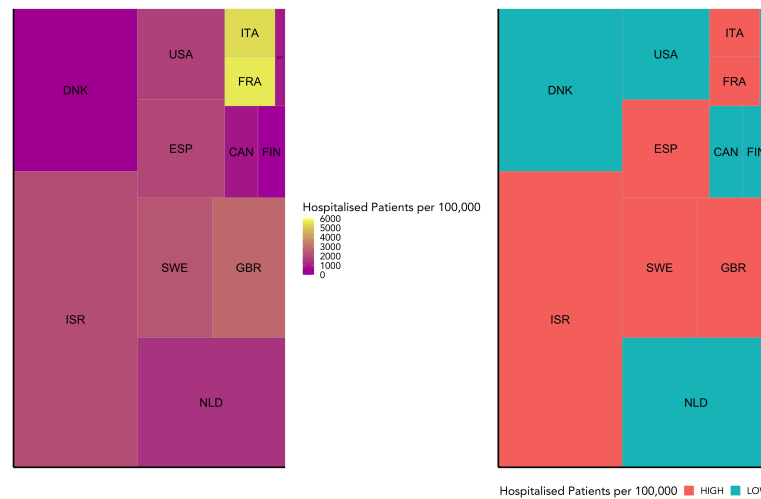
FIGURE 1
This figure visualises the weighted grouping of countries into COVID-19 severity groups. It shows the hospitalisation rate for each country included in this study relative to the number of microbiome samples available for that country. While the hospitalisation rate is indicated by colour, the sample size is represented by the area of each tile. On the left is the raw data, and on the right the grouping into high or low COVID-19 severity groups (high in red, low in blue) based on a weighted mean into even groups.

and Actinobacteria is shown in the high group in comparison to the low group (Figure 2).

The alpha diversity represented by violin plots in Figure 3, reveals significant differences, *via* the Wilcoxon Sign test ($p<0.05$), in microbiome diversity and distribution between the high and low groups. France presents with the greatest bacterial diversity and the lowest is associated with the United States of America.

TABLE 2 The demographics table reported shows the number of samples for each variable by COVID-19 severity group.

Demographic Data by COVID-19 Severity Groups.

Variable	Low Severity Group	High Severity group	Total
Gender			
Female	528	487	1015
Male	476	89	565
Unknown	314	1161	1475
Age Group			
Adult	1130	1546	2676
Child	88	92	180
Schoolage	66	3	69
Senior	34	96	130
BMI			
Lean (BMI 18.5+ to 25)	508	209	717
Overweight (BMI 25+ to 30)	230	114	344
Obese (BMI 30+ to 35)	146	43	189
Morbid obese (BMI 35+ to 40)	9	2	11
Severe Obese (BMI 40+ to 45)	25	11	36
Super Obese (BMI 45+)	16	11	27
Unknown	384	1347	1731

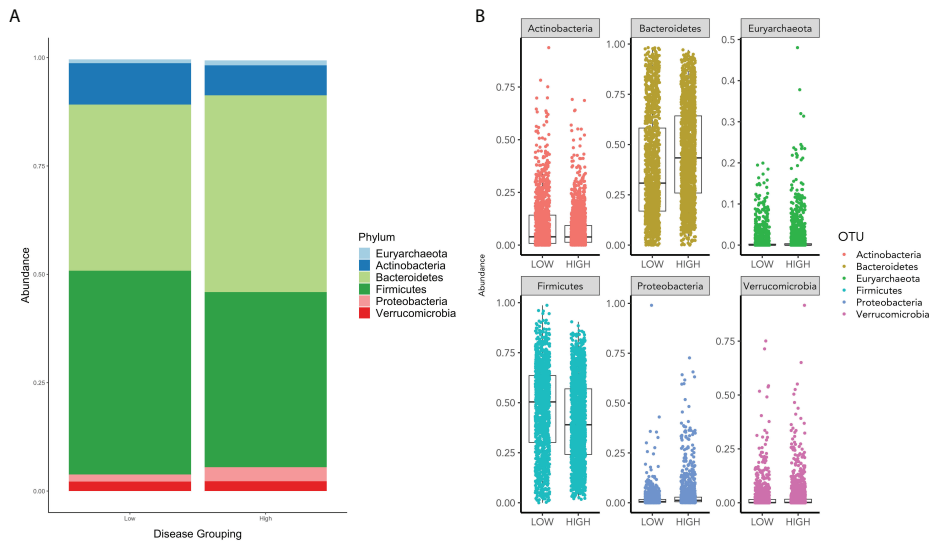These are gender, age group, and BMI. These values were obtained from the cMD Dataset.

FIGURE 2

(A) displays the relative phyla abundance for the Low and High COVID-19 groups. **(B)** displays the relative abundance for each phylum for High and Low High COVID-19 groups.

Beta diversity visually presents microbiome composition variations and differences between Covid-19 groups (Figure 4). Samples are represented as points on the graph, their closeness implies their similarity in sequence composition. The clustering of the high and low groups appears to occupy the same region.

Extending from the overall differences in phyla abundances between the high and low groups (Figure 2), differences in top genus-level taxa were explored and observed in Figure 5. A noticeable higher proportion of *Eubacterium rectale* dominates the low group contrarily to the *Prevotella copri* in the high group.
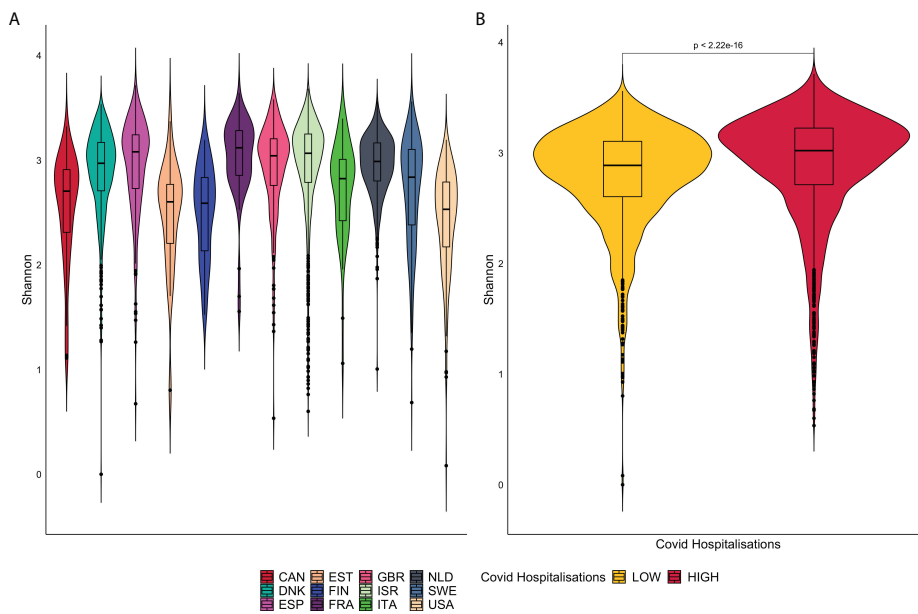

FIGURE 3

(A) is a representation of violin plots based on the Shannon alpha diversity index for each country. **(B)** displays violin plots based on Low and High COVID-19 groups with corresponding p-values.

FIGURE 4
Represents the Bray-Curtis dissimilarity distance metric applied to PCoA graph representing the difference between samples grouped by COVID-19 hospitalisations.
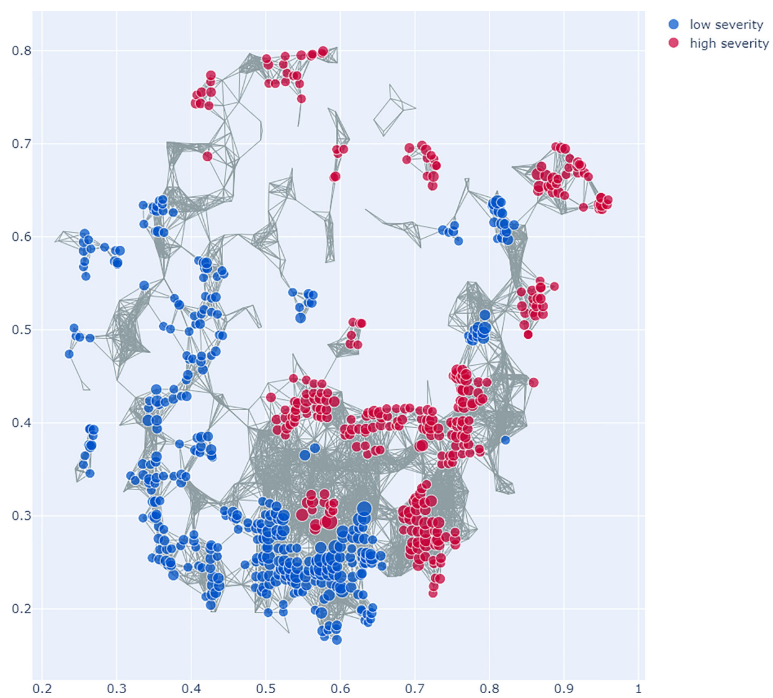

FIGURE 5
Top genus-level bacteria separating low and high COVID-19 groups. The bars towards the right represent higher abundance in the low COVID-19 group whilst the bars towards the left indicate higher abundance in the high COVID-19 groups. The x-axis in this figure displays the logarithmic change of bacteria which differed significantly.
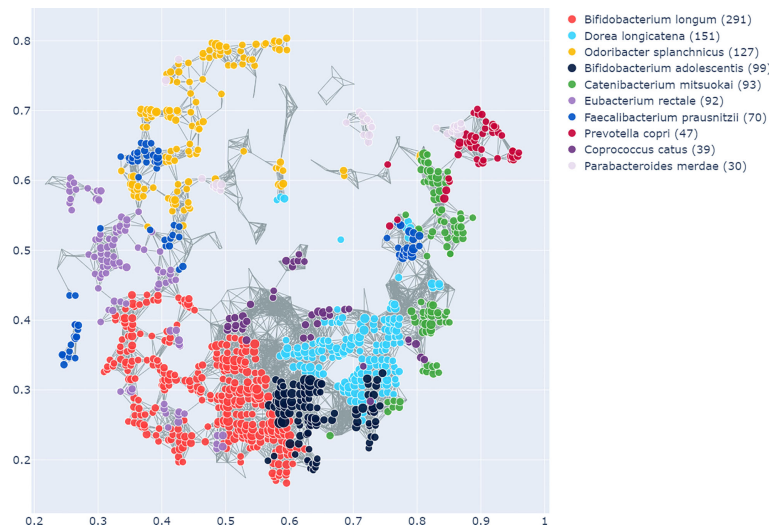
FIGURE 6
Topological Data Analysis (TDA) with Mapper: This figure shows the network graph made from the microbiome data. Each node represents a group of samples with similar microbiome profiles, each edge represents overlap in the groups. The colours represent the COVID-19 severity groups, enriched with the SAFE algorithm on the network. Specifically, which of the groups has the higher SAFE score at a given node. The figure shows that there is clear clustering of the groups across the network, with the low severity group mostly enriched across the lower left of the network graph. The high severity group is enriched in the right and top of the network graph. This suggests different microbiome profiles associated with the groups.

Overall multiple species of the *Ruminococcus* genus prevail in the low group. Both groups include the genus *Bacteroides* although distinct species are associated with them. The second most abundant genus-level taxa associated with the high and low group are *Escherichia coli* and *Bifidobacterium adolescentis,* respectively.

3.3 Topological data analysis

The analysis with TDA Mapper yields a graph network of the microbiome data that contains 1,323 nodes and 11,258 edges, with 1,125 (63.18%) samples lost.

SAFE enrichment of the COVID-19 severity groups on the Mapper graph shows clustering of the groups in distinct areas of the network: while the low severity group is enriched in the lower left part of the network, the high severity group is enriched in several clusters in the right and top parts of the network (see Figure 6). This suggests that different microbiome profiles are associated with low and high severity.

The enrichment graph comparing the bacterial taxa with the most enriched nodes across the network (Figure 7) also shows distinct clusters of enrichment. Most of these bacteria are highly common in the human gut. The genus with the highest number of enriched nodes is *Bifidobacterium longum* which also has the highest SAFE enriched score. It is enriched in a large cluster in

the bottom left of the network, as is for example *Eubacterium rectale* (second highest SAFE enriched score). *Dorea longicatena* on the other hand, the taxon with the second most enriched nodes and fifth highest SAFE score, is enriched in a cluster on the bottom right. While this graph visually suggests co-enrichment between the severity groups and some taxa, this can only be shown with pairwise co-enrichment analysis. The results of this are presented in the association matrix in Figure 8, with asterisks indicating statistically significant associations. While the high severity group only significantly co-enriched with the adult age category, the low severity group has several associations. It is associated firstly with three of the most highly enriched taxa, namely *Bifidobacteria longum* and *bifidum*, and *Eubacterium rectale*. Further, it is associated with many metadata of interest, including the individual-level variables male gender and the schoolchildren age category, as well as the population-level variables proportion of over 70-year-olds in the population, GDP, and HDI. Of the taxa associated with the low group, the *Bifidobacteria* are associated with normal BMI, and all three taxa co-enrich with each other and taxa such as *Eubacterium hallii*, *Bifidobacteria bifidum* and *adolescentis*, and *Ruminococcus bromii*. The three population-level covariates all significantly co-enrich with each other. Proportion over 70 years old and GDP are also each associated with the key taxon *Bifidobacterium longum* and other anti-inflammatory bacteria.
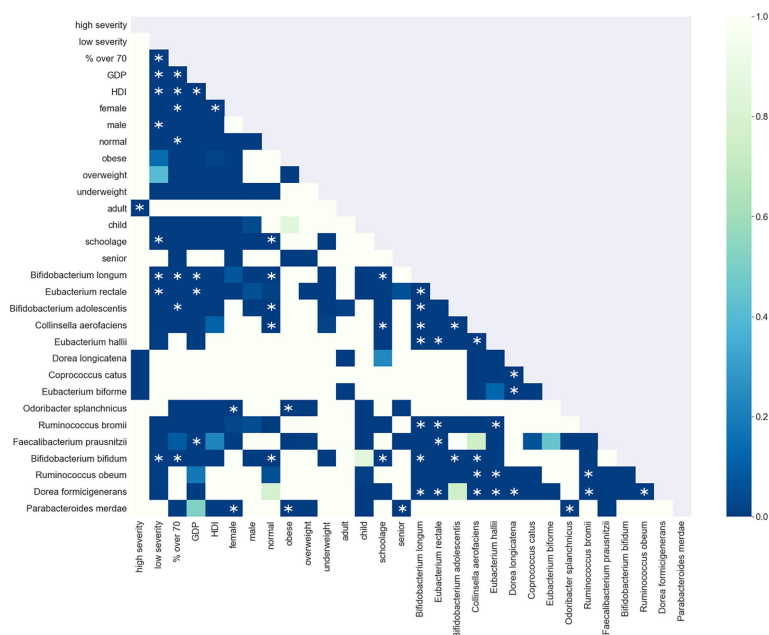
FIGURE 7
As in Figure 6, this shows the graph network constructed from the microbiome data with the TDA Mapper algorithm. This graph shows the bacterial taxa with the highest number of enriched nodes and the nodes in which a given taxon has the highest SAFE score are coloured accordingly.

Further, the co-enrichment analysis shows clusters of association between several anti-inflammatory bacteria, and between pro-inflammatory bacteria. One interesting association is between *Collinsella aerofaciens* and *Bifidobacteria longum* and *bifidum*, which, as mentioned above, are both associated with the low severity group.

Looking at the SAFE enriched scores of the metadata, it shows that both country-level metadata, namely proportion over
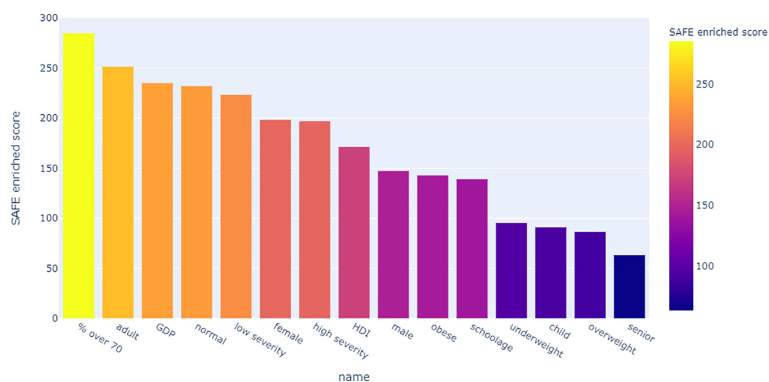

FIGURE 8
Shown is the association matrix of pairwise co-enrichment of all metadata and the 15 top taxa. The top taxa are a combination of the 10 taxa with the highest SAFE enriched scores (network-wide) and the 10 taxa with the highest number of enriched nodes. The colours represent the p-value, asterisks indicate significance at the 0.5th percentile of all pairwise associations (not all represented in the matrix). Of note, the low severity group is associated with the taxa *Eubacterium rectale, Bifidobacterium longum* and *Bifidobacterium bifidum*, and the population-level confounders GDP, HDI, and the proportion of over 70s in the population. These bacteria are also associated with each other, the same holds true for the confounders. Additionally, the bacteria and confounders are associated with each other and other anti-inflammatory bacteria such as *Bifidobacterium adolescentis* or *Faecalibacterium prausnitzii*.
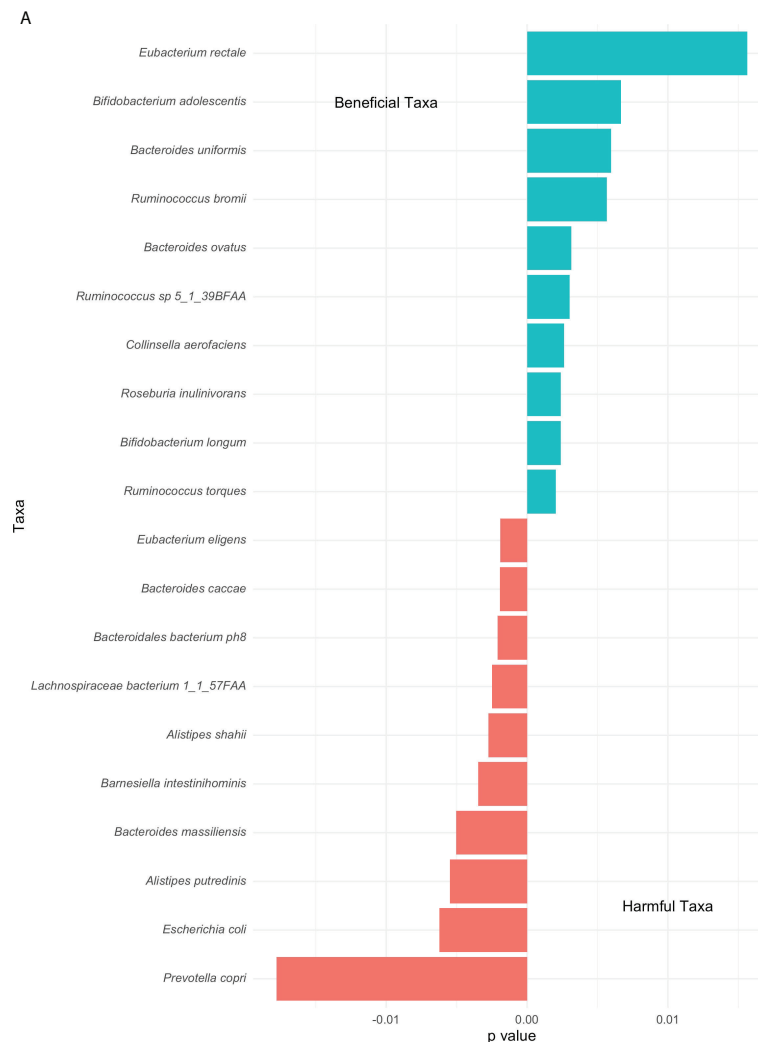
FIGURE 9
Bar chart of the network-wide SAFE enriched score of all metadata. This score is an indicator of effect size, allowing comparison across variables.

70 and GDP, as well as individual-level metadata, namely adult age and normal BMI were more highly enriched than the covid severity grouping (Figure 9).

4 Discussion

This study is the first to apply a population-based indirect association approach to the association between the gut microbiome and COVID-19 severity, supported by the use of unsupervised machine learning with Topological Data Analysis. The study successfully elucidated significant differences between countries with low and high COVID-19 severity. We have

detected microbiome signatures at the genus level which dominate the severity groups, as well as the influence of various confounding factors influencing this relationship. Microbiome composition varies significantly with geographical location, which is associated with a diverse range of diseases (Sun et al., 2020; Chen et al., 2021; Romano et al., 2021). Overall, our results support the complex association between the gut microbiome and COVID-19 severity.

This study comprises a large sample size of 3,055 gut microbiome samples which is vastly greater in comparison to that used by similar studies conducted at the individual level. Another major strength lies in the combination of two analysis approaches, both conventional and machine learning. This helps

to strengthen results and conclusions, as well as expand conventional linear analysis with high-dimensional analysis that can account for the complexities of the microbiome.

We observed both similarities and discrepancies with individual-level studies. Our results are consistent with the study from Yeoh et al., with an increased proportion of Bacteroides and decreased Actinobacteria in COVID-19 patients compared to controls (Yeoh et al., 2021). A small cross-sectional study (n=48) reported a decreased diversity and an increase in the Bacteroidetes phylum associated with increased severity of COVID-19 which is consistent with our findings, as well (Hazan et al., 2022). On the other hand, a study comparing controls with patients of different COVID-19 severity reported an increase in Actinobacteria in Covid-19 patients. They report a decrease in Firmicutes in line with our results, as well as a decrease in microbial diversity as Covid-19 severity increases which we did not find (Wu et al., 2021). The direct comparison (PERMANOVA) identified taxa specifically associated with COVID-19. Yeoh et al. (2021) found COVID-19 patients to have lower quantities of *Eubacterium rectale* and *Bifidobacterium adolescentis*. We identified *Eubacterium rectale* and Bifidobacterium, more specifically the genus *adolescentis*, as the two most beneficial taxa. This perhaps accounts for the lack of such bacterial genera in diseased patients. These genera in particular are also known to influence the immune response, further supporting our finding (Zheng et al., 2020; Baradaran Ghavami et al., 2021).

Considering the harmful taxa as identified by PERMANOVA, *Prevotella copri*, which belongs to the Bacteroidetes phylum, is also known to play a role in the development of rheumatoid arthritis (Alpizar-Rodriguez et al., 2019; Drago, 2019). In a small study comparing COVID-19 patients to controls, a decrease in *Prevotella copri* abundance was noted, however, a positive association with viral load of SARS-CoV-2 in the upper respiratory tract was also identified. Moreover, the same study found an increase in *E.coli* in COVID-19 patients, and a positive correlation with COVID-19 severity (Zhou et al., 2022), which is supported by the PERMANOVA in this study. Though *E.coli* are a commensal member of the gut microbiota, there are multiple well known pathogenic subspecies which may alter different cellular processes (Kaper et al., 2004). In line with other studies, these bacteria suggest a potential, detrimental role in COVID-19 severity, however, neither bacterium was found to be a significant association in the TDA analysis. *Prevotella copri* was identified as a highly enriched taxon and enriched in a similar area of the network graph as the high severity group, but this association did not pass the strict significance threshold.

The unsupervised machine learning with TDA provided additional insights above and beyond those from the PERMANOVA. The bacteria significantly associated with the low group are *Eubacterium rectale*, *Bifidobacterium longum*, and *bifidum*. These are thought to be anti-inflammatory bacteria and

the latter two of which are key strains in many probiotic therapeutics with a host of beneficial effects (Sharma et al., 2021). Strains from these and other probiotics species have previously been shown to have beneficial effects on the incidence and course of other viral respiratory infections, potentially by acting on the gut-lung axis and the innate immune system which makes it a target to reduce inflammation in COVID-19 (Baud et al., 2020; Stavropoulou and Bezirtzoglou, 2020). Indeed, strains of these two Bifidobacteria were used in a probiotic 'cocktail' to treat 99 hospitalised COVID-19 patients in a randomised controlled trial (Ivashkin et al., 2021). However, there was no effect of the probiotic on clinical outcomes. The same was found in a smaller case-control study (N=30), although neither the strains nor the dose of probiotics administered was recorded, making that study difficult to interpret (Veterini et al., 2021). While these results may appear to suggest a reduced influence of the microbiome on COVID-19 severity, it is actually in line with the results reported here. Since we take a population-level approach, we consider the 'baseline' microbiome of the population. The results thus speak to the potential preventative resilience of the microbiome and our results point towards potential preventative treatments rather than interventional treatments for COVID-19. This underscores the value of population-level approaches as taken in this study for treatment translation. This point is strengthened by the results reported in a recent preprint of an RCT which observed that exposed people who took a daily Lactobacillus-based probiotic were less likely to develop COVID-19 and if they did, to develop it later than controls (Wischmeyer et al., 2022). Together, this suggests that the microbiome might be an accessible and promising target for prophylaxis.

The depleted levels of the Bifidobacteria genus are consistent with other studies (Bozkurt and Bilen, 2021a). Additionally, hospitalised patients that received a *B.animalis* probiotic were reported to have reduced mortality and improved symptomatology, as well as a reduction in IL-6 (Bozkurt and Bilen, 2021b). Such findings are in line with the proposed mechanism of action of Bifidobacteria species on the immune system through affecting dendritic cells of the intestinal mucosa (López et al., 2011). In fact, it has been shown that different Bifidobacteria strains have specific effects on T cell differentiation. For example, there are four *B.bifidum* strains that are Th17 inducing (López et al., 2011). This is particularly interesting as there is now evidence to suggest that Th17, and the Th17 inducer cytokine IL-6, could be a driver of the COVID-19 cytokine storm, although mechanistic insight so far is inconclusive (Paiva et al., 2021). This is supported by previous evidence that suggests the microbiome as a regulator of the Treg/Th17 axis, which is responsible for appropriate protection from pathogens without excessive or autoimmune response (Omenetti and Pizarro, 2015). Together, this evidence is thus highly suggestive of a link between Bifidobacterium levels and COVID-19 severity.

Nevertheless, there are some discrepancies between our findings and the literature. For example, *Ruminococcus torques* had been found to be more abundant in Covid-19 patients whereas our PERMANOVA analysis includes this specific species in the beneficial taxa. Furthermore, the bacteria *Collinsella aerofacis* has been detected in a recent study published in 2020 in Covid-19 hospitalised patients, therefore going against our findings from the PERMANOVA (Zuo et al., 2021). In the TDA co-enrichment analysis, however, this species was associated with the probiotic species associated with the low severity group - *Bifidobacterium bifidum* - and *Eubacterium hallii*, *Ruminoccus obeum*, and *Dorea formicigenerans*. This could suggest an indirect mechanism of action on COVID-19 severity by reducing the abundance of those probiotic species that future studies should explore further. In line with the TDA finding and the Zuo *et al.* study, but not the PERMANOVA finding, the literature suggests a more pro-inflammatory profile of *Collinsella aerofacis*, as it seems to play a role in regulating immunity: it is associated with immunotoxicity in immune checkpoint inhibitor therapy in cancer (Kageyama et al., 2020), with type 2 diabetes (Lambeth et al., 2015), and with rheumatoid arthritis (Chen et al., 2016), although other studies report the opposite (Jeong et al., 2019). In overweight and obese pregnant women, a study found a positive association of the genus *Collinsella* with circulating insulin and a negative association with dietary fibre intake (Gomez-Arango et al., 2018). In our study, the bacterium was associated with normal BMI, as were the taxa it co-enriched with. This discrepancy could be due to different effects of species within a genus. In addition to pointing toward potentially novel mechanisms of action, these findings also support the superior performance of TDA over conventional analysis in picking up subtle and complex relationships, as suggested by Liao et al. (2019).

In addition to these bacterial associations, TDA suggested an association with both individual- and country-level confounders. Our results show associations of the low severity group with the percentage of over 70-year-olds in the population, GDP, and HDI, which were all also associated with each other, and the three anti-inflammatory bacteria associated with low severity. This suggests a confounding effect of these variables on the microbiome-disease severity association. This finding is in line with a previous study looking at between-country variation of the case fatality rate (CFR) which found positive associations between CFR and percentage over 70-year-olds, GDP, and level of democracy (Sorci et al., 2020). Their results suggest that there are highly complex interactions between different facets of healthcare spending, policy, and social or cultural factors that impact the severity of COVID-19. However, they can impact the microbiome as well. The diet of a population changes with population wealth, becoming more diverse and nutrient-rich with rising GDP (Ritchie and Roser, 2017). Diet in turn massively impacts the composition of the microbiome (David

et al., 2014). Further, socioeconomic status at the individual level is also associated with changes in microbiome composition (Bowyer et al., 2019). Similarly, the composition of the microbiome changes at an individual level with increasing age (Yatsunenko et al., 2012; Zapata and Quagliarello, 2015; Badal et al., 2020), which could arguably be reflected in the composition of the population microbiome.

Further, both the age distribution of a country, as well as its wealth, impact the number of comorbidities present in that population that could impact COVID-19 severity and the microbiome. This again could be mitigated by the performance of different healthcare systems. There are several different ways in which these three variables and the microbiome-COVID-19 relationship can be related. To disentangle these effects, future research needs to account for further variables such as diet, healthcare spending, healthcare performance, comorbidities, and public health policy.

Additionally, our study supports an association between male gender and low severity, which is contrary to the widely reported increased risk of men for worse COVID-19 outcomes and death (Jin et al., 2020). A large meta-analysis reports that male patients have much higher odds of both requiring intensive care (OR 2.84) and mortality (OR 1.39), which occurs globally (Peckham et al., 2020). This result is likely caused by a gender imbalance in the microbiome dataset: the dataset includes twice as many females as males and the high severity group specifically only contains 5% males. Similarly, low severity was significantly associated with the school-age age category, which is also underrepresented in the high severity group and has a low sample number in general. For variables with few samples which are highly distributed across the network, enrichment is difficult to interpret as permutation has a low probability.

There are several limitations of this study.

The study assesses association and not causality, which would require prospective and interventional studies. Additionally, the range of confounding factors included in our study has been limited due to the availability of data. For example, we do not have extensive metadata on ethnicity, or diet, antibiotics use, and other lifestyle habits of individual samples in our dataset – while this data exists for around 10-15% of subjects in the cMD dataset, it is not broad enough to be used in our analysis. These factors are all known to affect the microbiome (Dethlefsen et al., 2008; Monda et al., 2017; Vaughn et al., 2017; Huang and Shi, 2019), and in the case of ethnicity is also known to affect COVID-19 severity (Berkowitz et al., 2020). Further, by excluding samples with known diseases to avoid confounding in the microbiome dataset, we also limit how representative the included microbiome data is of its population, potentially producing sampling error.

There are similar issues with population-level data: while the observed co-enrichment of the country-level confounders and both the low severity group and significant taxa suggests a considerable confounding or mediating effect, more fine-

grained effects were not able to be detected due to a lack of appropriate variables. However, these three variables were chosen based on the existing literature and an *a posteriori* inclusion of further interesting variables to disentangle the observed effects is beyond the scope of this study. Further confounders that we would suggest being included in future studies can be grouped into four archetypes that affect both the microbiome as well as COVID-19 severity. First, confounders reflecting underlying population genetics such as the prevalence of specific COVID-19 prognostic risk alleles (Fricke-Galindo and Falfán-Valencia, 2021) that could have systematic effects on microbiome composition (Goodrich et al., 2014). Second, confounders reflecting individual lifestyle, such as smoking, which are known to affect both COVID-19 severity (Gao et al., 2021) and the microbiome (Savin et al., 2018). Third, confounders reflecting the healthcare system as a proxy measure of a country's ability to treat COVID-19 such as the rate of ICU beds or ECMO capabilities, but also as a proxy measure of the general health of a country as reflected by healthcare expenses per person. Finally, population health measures and policy during the pandemic should be included as they have an effect both on COVID-19 prevalence and severity (Huang et al., 2021; Arnold et al., 2022), as well as on the microbiome due to lifestyle changes induced by lockdowns and increased hygiene (Burchill et al., 2021; Finlay et al., 2021).

Further, due to the scarcity of samples and confounding variables, multiple countries ought to be excluded. The study may have benefitted from the inclusion of more countries to observe a full geographical variation. On the same note, we have noted an uneven distribution of microbiome samples across the countries, with an overall range of gut microbiome samples from 24 in Estonia to 900 in Israel, although using the weighted mean as a threshold for the low and high groups of disease severity allowed the groups to be even. Of course, a key assumption of this study is that the observed samples are representative of their populations. Even though a larger sample size and more evenness between countries would increase the likelihood of that assumption holding true.

Another limitation lies in the metric chosen to assess COVID-19 severity. The hospitalisation rate can be influenced by factors such as the healthcare system itself and the pressure on that system at any moment during the pandemic. Other potentially influential factors can revolve around public policy and in particular public health measures such as social distancing and lockdowns, as infection rates and thereby hospitalisations could have been affected. In light of these limitations, future studies should therefore aim to investigate the causal role of the gut microbiome in COVID-19 severity and further examine the long-term effects of the gut microbiome.

The findings of this study, together with similar findings in the literature, have translational potential. Considering the strong associations between the gut microbiome and the severity of COVID-19, future studies should look further into specific mechanisms with a particular emphasis on the possible beneficial role of prebiotics and probiotics and targeted dietary interventions.

Additionally, as previous studies have demonstrated antiviral activity of specific probiotic strains in several respiratory viruses, including other coronaviruses and COVID-19, special attention should be paid to more targeted approaches (Baud et al., 2020; Bozkurt and Bilen, 2021b). The need for strain-level studies is supported by the finding that different strains of the same species can have differential effects on T cell differentiation López et al., 2011. This includes individual- and population-level studies at the strain level, as well as targeted case-control studies and RCTs. However, the largest effect of probiotics may be exerted when delivered as a public health intervention. As this study suggests, a more anti-inflammatory microbiome profile may be a contributor to fewer hospitalisations from COVID-19 - with effects not only for the individual but for the whole country. Less pressure on the healthcare system could mean potentially fewer excess deaths due to cancelled routine appointments and planned surgeries as has been observed especially for cancer patients during the pandemic (Maringe et al., 2020; Riera et al., 2021). The microbiome is a particularly accessible treatment target for population-level interventions, as it can be modulated through oral supplements or diet. Similar to how foods, such as flour and salt, and other products, such as toothpaste, have been fortified to successfully prevent birth defects (Brown et al., 2011), iodine deficiency (Andersson et al., 2010), and caries (Twetman et al., 2003), respectively, such everyday products could be fortified to improve the microbiome.

4.1 Conclusion

The findings of this study suggest an association between the microbiome of country populations at baseline and the severity of COVID-19 experienced during the pandemic, as measured by hospitalisations. Specifically, anti-inflammatory bacterial taxa including *Bifidobacteria* and *Eubacterium rectale* were identified as protective. We also find evidence that characteristics of countries, such as the proportion of over 70-year-olds, as well as GDP and Human Development Index, influence this association between the microbiome and COVID-19 severity. Future studies should aim to disentangle the direct and indirect effects of wealth, policy, and population characteristics on the microbiome and on hospitalisations. These findings were made possible by a unique study design of indirect association using open-access population-level data. As our findings both support and expand upon individual-level findings, this study highlights the unique capabilities of such population-level studies for understanding diseases and finding novel treatment avenues, especially when considering the

microbiome which is a particularly accessible target. Finally, the use of an unsupervised machine learning tool, Topological Data Analysis, strengthened and expanded findings from conventional microbiome analysis, particularly through its ability to account for nonlinear relationships and the effect of confounders. This validates TDA as a valuable tool for microbiome analysis, particularly for microbiome-disease associations.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://bioconductor.org/packages/release/data/experiment/html/curatedMetagenomicData.html and https://github.com/owid/covid-19-data.

Author contributions

EL and GG: Conceptualization, methodology, formal analysis, writing – original draft, writing – review & editing, visualization, project administration. SB: Writing – review & editing. NS: Conceptualization, methodology, formal analysis, writing – review & editing, visualization, supervision, project administration, funding acquisition. All authors contributed to the article and approved the submitted version.

Conflict of interest

Author NS is CEO and founder of BioCorteX Ltd. Author SB owns equity in Owkin Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alpizar-Rodriguez, D., Lesker, T. R., Gronow, A., Gilbert, B., Raemy, E., Lamacchia, C., et al. (2019). Prevotella copri in individuals at risk for rheumatoid arthritis. *Ann. Rheumatic Dis. BMJ Publishing Group Ltd* 78 (5), 590–593. doi: 10.1136/annrheumdis-2018-214514

Andersson, M., de Benoist, B., and Rogers, L. (2010). 'Epidemiology of iodine deficiency: Salt iodisation and iodine status', Best practice & research clinical endocrinology & metabolism. *Baillière Tindall* 24 (1), 1–11. doi: 10.1016/J.BEEM.2009.08.005

Arnold, K. F., Gilthorpe, M. S., Alwan, N. A., Heppenstall, A. J., Tomova, G. D., McKee, M., et al. (2022). Estimating the effects of lockdown timing on COVID-19 cases and deaths in England: A counterfactual modelling study. *PLOSone* 17. doi: 10.1371/journal.pone.0263432

Badal, V. D., Vaccariello, E. D., Murray, E. R., Yu, K. E., Knight, R., Jeste, D. V., et al. (2020). The gut microbiome, aging, and longevity: A systematic review. *Nutrients* 12 (12), E3759. doi: 10.3390/nu12123759

Baek, M. S., Lee, M., Kim, W., Choi, J., and Jung, S. (2021). COVID-19-related outcomes in immunocompromised patients: A nationwide study in Korea. *PLOSone* 16 doi: 10.1371/journal.pone.0257641

Baryshnikova, A. (2016). Systematic functional annotation and visualization of biological networks. *Cell Syst.* 2, 412–421. doi: 10.1016/j.cels.2016.04.014

Baud, D., Dimopoulou Agri, V., Gibson, G., Reid, G., and Giannoni, E. (2020). Using probiotics to flatten the curve of coronavirus disease COVID-2019 pandemic. *Front. Public Health* 8. doi: 10.3389/fpubh.2020.00186

Belizário, J. E., and Faintuch, J. (2018). Microbiome and gut dysbiosis. *Experientia Supplementum 2012* 109, 459–476. doi: 10.1007/978-3-319-74932-7_13

Berkowitz, R. L., Gao, X., Michaels, E., and Mujahid, M. (2020). Structurally vulnerable neighbourhood environments and racial/ethnic COVID-19 inequities. *Cities Health* 0 (0), 1–4. doi: 10.1080/23748834.2020.1792069

Booth, A., Reed, A., Ponzo, S., Yassaee, A., Aral, M., Plans, D., Mohan, D., et al. (2021). Population risk factors for severe disease and mortality in COVID-19: A global systematic review and meta-analysis. *PLOSone* 16. doi: 10.1371/journal.pone.0247461

Bostancıklıoğlu, M. (2020). Temporal correlation between neurological and gastrointestinal symptoms of SARS-CoV-2. *Inflamm. Bowel Dis.* 26 (8), izaa131. doi: 10.1093/ibd/izaa131

Bowyer, R. C. E., Jackson, M. A., Le Roy, C. I., Ni Lochlainn, M., Spector, T. D., Dowd, J. B., et al. (2019). Socioeconomic status and the gut microbiome: A TwinsUK cohort study. *Microorganisms* 7 (1), 17. doi: 10.3390/microorganisms7010017

Bozkurt, H. S., and Bilen, Ö. (2021a). Gastrointestinal symptoms in COVID-19 could be associated with severe lung involvement and increased readmission rates. *Eur. J. Inflammation* 19, 20587392211048260. doi: 10.1177/20587392211048259

Bozkurt, H. S., and Bilen, Ö. (2021b). Oral booster probiotic bifidobacteria in SARS-COV-2 patients. *Int. J. Immunopathol. Pharmacol.* 35, 20587384211059676. doi: 10.1177/20587384211059677

Brown, R. D., Langshaw, M., Uhr, E., Gibson, J., and Joshua, D. (2011). The impact of mandatory fortification of flour with folic acid on the blood folate levels of an Australian population. *Med. J. Australia. John Wiley Sons Ltd* 194 (2), 65–67. doi: 10.5694/j.1326-5377.2011.tb04169.x

Burchill, E., Lymberopoulos, E., Menozzi, E., Budhdeo, S., McIlroy, J., Macnaughtan, J., et al. (2021). The unique impact of COVID-19 on human gut microbiome research. *Front. Med.* 8. doi: 10.3389/FMED.2021.652464

Carlsson, G. (2009). Topology and data. *Bull. Am. Math. Soc.* 46 (2), 255–308. doi: 10.1090/S0273-0979-09-01249-X

Chen, J. T., and Krieger, N. (2021). Revealing the unequal burden of COVID-19 by income, Race/Ethnicity, and household crowding: US county versus zip code analyses. *J. Public Health Manage. Pract.* 27, S43. doi: 10.1097/PHH.0000000000001263

Chen, Z., Radjabzadeh, D., Chen, L., Kurilshikov, A., Kavousi, M., Ahmadizar, F., et al. (2021). Association of insulin resistance and type 2 diabetes with gut microbial diversity: A microbiome-wide analysis from population studies. *JAMA Network Open* 4 (7), e2118811. doi: 10.1001/jamanetworkopen.2021.18811

Chen, J., Wright, K., Davis, J. M., Jeraldo, P., Marietta, E. V., Murray, J., et al. (2016). An expansion of rare lineage intestinal microbes characterizes rheumatoid arthritis. *Genome Med.* 8 (1), 43. doi: 10.1186/s13073-016-0299-7

Coomes, E. A., and Haghbayan, H. (2020). Interleukin-6 in covid-19: A systematic review and meta-analysis. *Rev. Med. Virol.* 30 (6), e2141. doi: 10.1002/rmv.2141

Cron, R. Q., Caricchio, R., and Chatham, W. W. (2021). Calming the cytokine storm in COVID-19. *Nat. Med.* 27 (10), 1674–1675. doi: 10.1038/s41591-021-01500-9

Damayanthi, H. D. W. T., Prabani, K. I. P., and Weerasekara, I. (2021). Factors associated for mortality of older people with COVID 19: A systematic review and meta-analysis. *Gerontol. Geriatric Med.* 7, 23337214211057390. doi: 10.1177/23337214211057392

David, L. A., Maurice, C. F., Carmody, R. N., Gootenberg, D. B., Button, J. E., Wolfe, B. E., et al. (2014). Diet rapidly and reproducibly alters the human gut microbiome. *Nature.* 505 (7484), 559–563. doi: 10.1038/nature12820

Deschasaux, M., Bouter, K. E., Prodan, A., Levin, E., Groen, A. K., Herrema, H., Nieuwdorp, M., et al. (2018). Depicting the composition of gut microbiota in a population with varied ethnic origins but shared geography. *Nat. Med.* 24, 10. doi: 10.1038/s41591-018-0160-1

Dethlefsen, L., Huse, S., Sogin, M., and Relman, D. (2008). The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16s rRNA sequencing. *PLoS Biol.* 6 (11), 2383–2400. doi: 10.1371/journal.pbio.0060280

Drago, L. (2019). Prevotella copri and microbiota in rheumatoid arthritis: Fully convincing evidence? *J. Clin. Med.* 8 (11), 1837. doi: 10.3390/jcm8111837

Falony, G., Joossens, M., Viera-Silva, S., Wang, J., DarziI, Y., Faust, K., et al. (2016). Population-level analysis of gut microbiome variation. *Science* 352, 560-564. doi: 10.1126/science.aad3503

Finlay, B., Amato, K., Azad, M., Blaser, M., Bosch, T., Chu, H., et al. (2021). The hygiene hypothesis, the COVID pandemic, and consequences for the human microbiome. *Proc. Natl. Acad. Sci.* 118 (6), e2010217118. doi: 10.1073/pnas.2010217118

Fortner, A., and Schumacher, D. (2021). First COVID-19 vaccines receiving the US FDA and EMA emergency use authorization. *Discoveries (Craiova Romania)* 9 (1), e122. doi: 10.15190/d.2021.1

Fricke-Galindo, I., and Falfán-Valencia, R. (2021). Genetics insight for COVID-19 susceptibility and severity: A review. *Front. Immunol.* 12. doi: 10.3389/fimmu.2021.622176

Gao, Y., Ding, M., Dong, J-j, Zhang, X., Azkur, A. K., Azkur, D., et al. (2021). Risk factors for severe and critically ill COVID-19 patients: A review. *Allergy* 76 (2), 428–455. doi: 10.1111/all.14657

Baradaran Ghavami, S., Pourhamzeh, M., Farmani, M., Keshavarz, H., Shahrokh, S., Shpichka, A., et al. (2021). Cross-talk between immune system and microbiota in COVID-19. *Expert Rev. Gastroenterol. Hepatol.* pp, 1–14. doi: 10.1080/17474124.2021.1991311

Gomez-Arango, L. F., Barrett, H. L., Wilkinson, S. A., Callaway, L. K., Mclntyre, H. D., Morrison, M., et al. (2018). Low dietary fiber intake increases collinsella abundance in the gut microbiota of overweight and obese pregnant women. *Gut Microbes* 9 (3), 189–201. doi: 10.1080/19490976.2017.1406584

Goodrich, J. K., Waters, J. L., Poole, A. C., Sutter, J. L., Koren, O., Blekham, R., et al. (2014). Human genetics shape the gut microbiome. *Cell* 159 (4), 789–799. doi: 10.1016/j.cell.2014.09.053

Guney, C., and Akar, F. (2021). Epithelial and endothelial expressions of ACE2: SARS-CoV-2 entry routes. *J. Pharm. Pharm. Sci.* 24, 84–93. doi: 10.18433/jpps31455

Gupta, S., Mortensen, M. S., Schjorring, S., Trivedi, U., Vestergaard, G., Stockholm, J., et al. (2019). Amplicon sequencing provides more accurate microbiome information in healthy children compared to culturing. *Commun. Biol. Nat. Publishing Group* 2 (1), 1–7. doi: 10.1038/s42003-019-0540-1

Gupta, V. K., Paul, S., and Dutta, C. (2017). Geography, ethnicity or subsistence-specific variations in human microbiome composition and diversity. *Front. Microbiol.* 8. doi: 10.3389/fmicb.2017.01162

Hazan, S., Stollman, N., Bozkurt, H. S., Dave, S., Papoutsis, A. J., Daniels, J., et al. (2022). Lost microbes of COVID-19: Bifidobacterium, faecalibacterium depletion and decreased microbiome diversity associated with SARS-CoV-2 infection severity. *BMJ Open Gastroenterol.* 9 (1), e000871. doi: 10.1136/bmjgast-2022-000871

Huang, X., Shao, X., Xing, L., Hu, Y., Sin, D. D., and Zhang, X. (2021).The impact of lockdown timing on COVID-19 transmission across US counties 38 *EClinicalMedicine*, 101035. doi: 10.1016/j.eclinm.2021.101035

Huang, C., and Shi, G. (2019). Smoking and microbiome in oral, airway, gut and some systemic diseases. *J. Trans. Med.* 17 (1), 225. doi: 10.1186/s12967-019-1971-7

Ivashkin, V., Fomin, V., Moiseev, S., Brovko, M., Maslennikov, R., Ulyanin, A., et al. (2021). Efficacy of a probiotic consisting of lacticaseibacillus rhamnosus PDV 1705, bifidobacterium bifidum PDV 0903, bifidobacterium longum subsp. infantis PDV 1911, and bifidobacterium longum subsp. longum PDV 2301 in the treatment of hospitalized patients with COVID-19: a randomized controlled trial. *Probiotics Antimicrobial Proteins* 1, 1–9. doi: 10.1007/S12602-021-09858-5/TABLES/4

Jeong, Y., Kim, J., You, H. J., Park, S., Lee, L., Ju, J. H., et al. (2019). Gut microbial composition and function are altered in patients with early rheumatoid arthritis. *J. Clin. Med. Multidiscip. Digital Publishing Inst.* 8 (5), 693. doi: 10.3390/jcm8050693

Jiao, Y., Wu, L., Huntington, N. D., and Zhang, X. (2020)Crosstalk between gut microbiota and innate immunity and its implication in autoimmune diseases. Available at: https://www.frontiersin.org/article/10.3389/fimmu.2020.00282 (Accessed 18 May 2022).

Jin, J.-M., Bai, P., He, W., Wu, F., Liu, X.-F., Han, D.-M., et al. (2020)Gender differences in patients with COVID-19: Focus on severity and mortality. Available at: https://www.frontiersin.org/article/10.3389/fpubh.2020.00152 (Accessed 2 March 2022).

Kageyama, T., Matsuo, T., Kurakake, R., Sano, T., and Sun, J. (2020).Chapter four - relationship between T cells and microbiota in health and disease. *Progress in molecular biology and translational science*, 95–129. doi: 10.1016/bs.pmbts.2020.03.007

Kaper, J. B., Nataro, J. P., and Mobley, H. L. T. (2004). Pathogenic escherichia coli. *Nat. Rev. Microbiol.* 2 (2), 123–140. doi: 10.1038/nrmicro818

Karlsson, F. H., Nookaew, I., and Nielsen, J. (2014). Metagenomic data utilization and analysis (MEDUSA) and construction of a global gut microbial gene catalogue. *PLoS Comput. Biol.* 10 (7), e1003706. doi: 10.1371/journal.pcbi.1003706

Kim, H. S. (2021). Do an altered gut microbiota and an associated leaky gut affect COVID-19 severity? *mBio. Am. Soc. Microbiol.* 12 (1), e03022–e03020. doi: 10.1128/mBio.03022-20

Lambeth, S. M., Carson, T., Lowe, J., Ramaraj, T., Leff, J. W., Luo, L., et al. (2015). Composition, diversity and abundance of gut microbiome in prediabetes and type 2 diabetes. *J. Diabetes Obes.* 2 (3), 1–7. doi: 10.15436/2376-0949.15.031

Levy, M., Kolodziejczyk, A. A., Thaiss, C. A., and Elinav, E. (2017). Dysbiosis and the immune system. *Nat. Rev. Immunol.* 17 (4), 219–232. doi: 10.1038/nri.2017.7

Liao, T., Wei, Y., Luo, M., Zhao, G.-P., and Zhou, H. (2019). Tmap: An integrative framework based on topological data analysis for population-scale microbiome stratification and association studies. *Genome Biol.* 20 (1), 1–19. doi: 10.1186/s13059-019-1871-4

López, P., González-Rodríguez, I, Gueimonde, M, Margolles, A, and Suárez, A. (2011). Immune response to bifidobacterium bifidum strains support Treg/Th17 plasticity 6. *PLoS ONE*, e24776. doi: 10.1371/journal.pone.0024776

Lymberopoulos, E., Gentili, G. I., Alomari, M., and Sharma, N. (2021) Topological data analysis highlights novel geographical signatures of the human gut microbiome. Available at: https://www.frontiersin.org/article/10.3389/frai.2021.680564 (Accessed 27 April 2022).

Maringe, C., Spicer, J., Morris, M., Purushotham, A., Nolte, E., Sullivan, R., et al. (2020). The impact of the COVID-19 pandemic on cancer deaths due to delays in diagnosis in England, UK: a national, population-based, modelling study. *Lancet Oncol.* 21 (8), 1023–1034. doi: 10.1016/S1470-2045(20)30388-0

Mohammed, A. H., Blebil, A., Dujaili, J., and Rasool-Hassan, B. A. (2020). The risk and impact of COVID-19 pandemic on immunosuppressed patients: Cancer, HIV, and solid organ transplant recipients. *AIDS Rev.* 22 (3), 151–157. doi: 10.24875/AIDSRev.20000052

Molina-Mora, J. A., González, A, Jiménez-Morgan, S, Cordero-Laurent, E, Brenes, H, et al. (2022). Clinical profiles at the time of diagnosis of SARS-CoV-2 infection in Costa Rica during the pre-vaccination period using a machine learning approach. *Phenomics* 7, 1–11. doi: 10.1007/s43657-022-00058-x

Monda, V., Villano, I., Messina, A., Valenzano, A., Esposito, T., Moscatelli, F., et al. (2017). Exercise modifies the gut microbiota with positive health effects. *Oxid. Med. Cell. Longevity* (2017) 1–8. doi: 10.1155/2017/3831972

Omenetti, S., and Pizarro, T. T. (2015). The Treg/Th17 axis: A dynamic balance regulated by the gut microbiome. *Front. Immunol.* 6. doi: 10.3389/fimmu.2015.00639

Paiva, I. A., Badolato-Corrêa, J, Familiar-Macedo, D, and de-Oliveira-Pinto, LM. (2021). Th17 cells in viral infections–friend or foe? *Cells* 10 (5), 1159. doi: 10.3390/cells10051159

Pasolli, E., Schiffer, L., Manghi, P., Renson, A., Obenchain, V., Truong, D., et al. (2017). Accessible, curated metagenomic data through ExperimentHub'. *Nature Methods* 14 (11), 1023–1024. doi: 10.1038/nmeth.4468

Peckham, H., de Gruijter, N. M., Raine, C., Radziszewska, A., Ciurtin, C., Wedderburn, L. R., et al. (2020). Male Sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission. *Nat. Commun.* 11 (1), 6317. doi: 10.1038/s41467-020-19741-6

Pijls, B. G., Jolani, S., Artherley, A., Derckx, R. T., Dijkstra, J. I. R., Franssen, G. H. L., et al. (2021). Demographic risk factors for COVID-19 infection, severity, ICU admission and death: a meta-analysis of 59 studies. *BMJ Open* 11 (1), e044640. doi: 10.1136/bmjopen-2020-044640

Ragab, D., Ragab, D., Salah Eldin, H., Taeimah, M., Khattab, R., and Salem, R. (2020)The COVID-19 cytokine storm; what we know so far. Available at: https://www.frontiersin.org/article/10.3389/fimmu.2020.01446 (Accessed 18 May 2022).

Ramiro, S., Mustard, R. L. M., Magro-Checa, C., van Dongen, C. M. P., Dormans, T., Buijs, J., et al. (2020). Historically controlled comparison of glucocorticoids with or without tocilizumab versus supportive care only in patients with COVID-19-associated cytokine storm syndrome: results of the CHIC study. *Ann. Rheumatic Dis.* 79 (9), 1143–1151. doi: 10.1136/annrheumdis-2020-218479

Riccaboni, M., and Verginer, L. (2022). The impact of the COVID-19 pandemic on scientific research in the life sciences. *PLOSone* 17. doi: 10.1371/journal.pone.0263001

Riera, R., Bagattini, A. M., Pacheco, R. L., Pachito, D. V., Roitberg, F., and Ilbawi, A. (2021). Delays and disruptions in cancer health care due to COVID-19 pandemic: Systematic review. *JCO Global Oncol.* 7), 311–323. doi: 10.1200/GO.20.00639

Ritchie, H., and Roser, M. (2017)Diet compositions. Available at: https://ourworldindata.org/diet-compositions (Accessed 27 May 2022).

Robertson, R. C., Manges, A. R., Finlay, B. B., and Prendergast, A. J. (2019). The human microbiome and child growth – first 1000 days and beyond. *Trends Microbiol.* 27 (2), 131–147. doi: 10.1016/j.tim.2018.09.008

Romano, S., Savva, G. M., Bedarf, J. R., Charles, I. G., Hildebrand, F., and Narbad, A. (2021). Meta-analysis of the parkinson's disease gut microbiome suggests alterations linked to intestinal inflammation. *NPJ Parkinson's Dis.* 7 (1), 1–13. doi: 10.1038/s41531-021-00156-z

Savin, Z., Kivity, S, Yonath, H, and Yehuda, S. (2018). Smoking and the intestinal microbiome. *Arch. Microbiol.*, 200 (5), 677–684. doi: 10.1007/s00203-018-1506-2

Segal, J. P., Mak, J. W. Y., Mullish, B. H., Alexander, J. L., Ng, S. C., and Marchesi, J. R. (2020). The gut microbiome: an under-recognised contributor to the COVID-19 pandemic? *Ther. Adv. Gastroenterol.* 13, 1756284820974914. doi: 10.1177/1756284820974914

Sender, R., Fuchs, S., and Milo, R. (2016). Revised estimates for the number of human and bacteria cells in the body. *PloS Biol.* 14 (8), e1002533. doi: 10.1371/journal.pbio.1002533

Sharma, M., Wasan, A., and Sharma, R. K. (2021). Recent developments in probiotics: An emphasis on bifidobacterium. *Food Bioscience* 41, 100993. doi: 10.1016/j.fbio.2021.100993

Singh, G., Facundo, M, and Gunnar, C. (2007). Topological methods for the analysis of high dimensional data sets and 3D object recognition. *Eurographics symposium on point-based graphics*. Eds. M. Botsch, et al (The Eurographics Association), 91–100. doi: 10.2312/SPBG/SPBG07/091-100

Sorci, G., Faivre, B., and Morand, S. (2020). Explaining among-country variation in COVID-19 case fatality rate. *Sci. Rep.* 10 (1), 18909. doi: 10.1038/s41598-020-75848-2

Stavropoulou, E., and Bezirtzoglou, E. (2020)Probiotics as a weapon in the fight against COVID-19. Available at: https://www.frontiersin.org/article/10.3389/fnut.2020.614986 (Accessed 6 June 2022).

Sun, S., Wang, H., Tsilimigras, M. C. B., Howard, A. G., Sha, W., Zhang, J., et al. (2020). Does geographical variation confound the relationship between host factors

and the human gut microbiota: a population-based study in China. *BMJ Open* 10 (11), e038163. doi: 10.1136/bmjopen-2020-038163

Thaiss, C. A., Zmora, N., Levy, M., and Elinav, E. (2016). The microbiome and innate immunity. *Nature* 535 (7610), 65–74. doi: 10.1038/nature18847

Twetman, S., Axelsson, S., Dahlgren, H., Holm, A.-K., Kallestal, C., Langerlof, F., et al. (2003). Caries-preventive effect of fluoride toothpaste: a systematic review. *Acta Odontologica Scandinavica* 61 (6), 347–355. doi: 10.1080/00016350310007590

Vaughn, A., M. Cooper, E., M. DiLorenzo, P., J. O'Loughlin, L., E. Konkel, M., H. Peters, J., et al. (2017). Energy-dense diet triggers changes in gut microbiota, reorganization of gut-brain vagal communication and increases body fat accumulation. *Acta Neurobiologiae Experimentalis* 77 (1), 18–30. doi: 10.21307/ane-2017-033

Veterini, A. S., Kurniawati, E. M., Soedarmo, S. M., Hamzah,, Prakoeswa, C. R. S., and Tinduh, D. (2021). Probiotics intake as adjunct therapy for infected health-care with SARS COV-2. *Indian J. Forensic Med. Toxicol.* 15 (2), 4021–4025. doi: 10.37506/ijfmt.v15i2.15003

Villapol, S. (2020). Gastrointestinal symptoms associated with COVID-19: impact on the gut microbiome. *Trans. Res.* 226, 57–69. doi: 10.1016/j.trsl.2020.08.004

Wang, J., Chen, W.-D., and Wang, Y.-D. (2020)The relationship between gut microbiota and inflammatory diseases: The role of macrophages. Available at: https://www.frontiersin.org/article/10.3389/fmicb.2020.01065 (Accessed 18 May 2022).

Wang, B., Yao, M., Lv, L., Ling, Z., and Li, L. (2017). The human microbiota in health and disease. *Engineering* 3 (1), 71–82. doi: 10.1016/J.ENG.2017.01.008

Wang, J., Zhao, S., Liu, M., Zhao, Z., Xu, Y., Wang, P., et al. (2020). ACE2 expression by colonic epithelial cells is associated with viral infection, immunity and energy metabolism. *Allergy Immunol* 96, 19–24. doi: 10.1101/2020.02.05.20020545

Wischmeyer, P. E., Tang, H., Ren, Y., Bohannon, L., Ramirez, Z. E., Andermann, T. M., et al. (2022). Daily lactobacillus probiotic versus placebo in COVID-19-Exposed household contacts (PROTECT-EHC): A randomized clinical trial. *medRxiv*, 2022.01.04.21268275. doi: 10.1101/2022.01.04.21268275

Wu, Y., Cheng, X., Jiang, G., Tang, H., Ming, S., Tang, L., et al. (2021). Altered oral and gut microbiota and its association with SARS-CoV-2 viral load in COVID-19 patients during hospitalization. *NPJ Biofilms Microbiomes* 7, 61. doi: 10.1038/s41522-021-00232-5

Wu, Y., Guo, C., Tang, L., Hong, Z., Zhou, J., Dong, X., et al. (2020). Prolonged presence of SARS-CoV-2 viral RNA in faecal samples. *Lancet Gastroenterol. Hepatol.* 5 (5), 434–435. doi: 10.1016/S2468-1253(20)30083-2

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreas, M., et al. (2012). Human gut microbiome viewed across age and geography. *Nature* 486, 2012, 222–228. doi: 10.1038/nature11053

Yeoh, Y. K., Zuo, T., Lui, G. C.-Y., Zhang, F., Liu, Q., Li, A. Y. L., et al. (2021). Gut microbiota composition reflects disease severity and dysfunctional immune responses in patients with COVID-19. *Gut* 70 (4), 698–706. doi: 10.1136/gutjnl-2020-323020

Zapata, H. J., and Quagliarello, V. J. (2015). The microbiota and microbiome in aging: Potential implications in health and age-related diseases. *J. Am. Geriatrics Soc.* 63 (4), 776–781. doi: 10.1111/jgs.13310

Zhang, H., Wu, Y., He, Y., Liu, X., Liu, M., Tang, Y., et al. (2022)Age-related risk factors and complications of patients with COVID-19: A population-based retrospective study. Available at: https://www.frontiersin.org/article/10.3389/fmed.2021.757459 (Accessed 18 May 2022).

Zheng, D., Liwinski, T., and Elinav, E. (2020). Interaction between microbiota and immunity in health and disease. *Cell Res.*, 30 (6), 492–506. doi: 10.1038/s41422-020-0332-7

Zhou, T., Wu, J., Zeng, Y., Li, J., Yan, J., Meng, W., et al. (2022). SARS-CoV-2 triggered oxidative stress and abnormal energy metabolism in gut microbiota. *MedComm* 3 (1), e112. doi: 10.1002/mco2.112

Zuo, T., Liu, Q., Zhang, F., Lui, G. C.-Y., Tso, E. Y. K., Yeoh, Y. K., et al. (2021). Depicting SARS-CoV-2 faecal viral activity in association with gut microbiota composition in patients with COVID-19. *Gut* 70 (2), 276–284. doi: 10.1136/gutjnl-2020-322294