University of Dundee

The sputum microbiome and clinical outcomes in patients with bronchiectasis

Dicker, Alison J.; Loneragan, Mike; Keir, Holly R.; Smith, Alexandria H.; Pollock, Jennifer; Finch, Simon

Published in:
The Lancet Respiratory Medicine

DOI:
[10.1016/S2213-2600\(20\)30557-9](https://doi.org/10.1016/S2213-2600(20)30557-9)

Publication date:
2021

Licence:
CC BY-NC-ND

Document Version
Peer reviewed version

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

Dicker, A. J., Loneragan, M., Keir, H. R., Smith, A. H., Pollock, J., Finch, S., Cassidy, A. J., Huang, J. T. J., & Chalmers, J. D. (2021). The sputum microbiome and clinical outcomes in patients with bronchiectasis: a prospective observational study. *The Lancet Respiratory Medicine*, 9(8), 885-896.
[https://doi.org/10.1016/S2213-2600\(20\)30557-9](https://doi.org/10.1016/S2213-2600(20)30557-9)

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

The sputum microbiome is associated with exacerbation frequency and mortality in patients with bronchiectasis: A prospective observational study

Alison J. Dicker PhD¹, Mike Lonergan PhD¹, Holly R Keir BSc¹, Alexandria H. Smith², Jennifer Pollock¹, Simon Finch MBChB¹, Andrew J. Cassidy PhD¹, Jeffrey T. J. Huang PhD³, Prof James D. Chalmers PhD¹

¹Scottish Centre for Respiratory Disease Research, University of Dundee, Dundee, UK

²University Cambridge, Cambridge, UK

³Division of Systems Medicine, School of Medicine, University of Dundee, Ninewells Hospital and Medical School, Dundee, UK

Corresponding author: James D. Chalmers, Scottish Centre for Respiratory Research, University of Dundee, Ninewells Hospital and Medical School, Dundee, DD1 9SY. jchalmers@dundee.ac.uk

Word count: 3869

Keywords: Microbiome, bronchiectasis, *Pseudomonas*, mortality

Funding: Funded by the GSK/British Lung Foundation Chair of Respiratory Research and European Respiratory Society through the EMBARC2 consortium. EMBARC2 is supported by project partners AstraZeneca, Chiesi, Grifols, Insmmed, Janssen, Novartis and Zambon.

Research in Context Panel

Evidence before this study: We searched PubMed from inception until Oct 9th 2019 for articles containing the terms “Bronchiectasis” and “microbiome” or “microbiota”, identifying 71 publications from 2011-2019. Searches were not limited by language and were filtered for relevance based on abstract alone. The majority of microbiome studies in bronchiectasis to date have included small numbers of patients, have been cross-sectional and a high proportion derive from a single randomized controlled trial cohort called BLESS. Existing evidence is inconsistent. Relationships between the microbiome and disease severity have been limited by the small sample size of most studies but some microbiome parameters have been shown to be associated with forced expiratory volume in 1 second and frequency of exacerbations. Some studies show a reduction in microbial diversity around exacerbation and others do not. One study showed that chronic macrolide treatment can increase the relative abundance of *Pseudomonas*. While it is known that microbiology and, in particular, chronic infection with *Pseudomonas aeruginosa* is associated with increased mortality, there have been no studies examining the association of long-term outcomes with the microbiome in bronchiectasis.

Added value of this study:

We report evidence from a prospective longitudinal cohort of 281 bronchiectasis patients with a median follow up of four years. At baseline we identified an inverse relationship between alpha diversity, measured using the Shannon-Wiener Diversity Index, and severity of disease measured using the multidimensional bronchiectasis severity index or forced expiratory volume in 1 second. Alpha diversity was also negatively associated with symptoms measured using the quality of life bronchiectasis respiratory symptom scale. Using Random Forest analysis, we identified in baseline samples that the genera *Stenotrophomonas*, *Pseudomonas* and Enterobacteriaceae were associated with a bronchiectasis severity index ≥ 9 whilst the genera *Pseudomonas* and Enterobacteriaceae were also associated with frequent exacerbations (≥ 3 exacerbations per year). Longitudinal stable and exacerbation microbiome profiles were heterogenous, with greater similarity within individual patients than between samples from different patients. A distinct microbiome profile associated with exacerbation was not identified. Over long-term follow-up, *Pseudomonas* dominance and lower alpha diversity were independently associated with all-cause mortality and hospitalization for severe exacerbations.

Implications of all available evidence:

Disease severity and long-term mortality is associated with loss of microbial diversity and dominance of Proteobacteria such as *Pseudomonas*, Enterobacteriaceae and *Stenotrophomonas* in patients with bronchiectasis. In contrast to the established paradigm that bronchiectasis exacerbations are primarily caused by bacterial infection we found little evidence of microbial changes at exacerbation. Taken

52 together this data suggests that the microbiome may identify subgroups of patients at increased risk of
53 poor outcomes who may benefit from precision treatment strategies in future.

Abstract

Introduction: Infection is a key component of the pathophysiology of bronchiectasis. Characterisation of the microbiome offers a higher degree of sensitivity and resolution than traditional culture to evaluate the importance of infection in determining the risk of exacerbation and long-term outcomes, including all-cause mortality, in bronchiectasis.

Methods: Sputum samples from 281 patients with bronchiectasis included in a longitudinal cohort of patients from two hospitals in the East of Scotland underwent 16S rRNA gene sequencing. Patients were enrolled from 2012-2015 and followed up until January 2019 for long term outcomes. Repeat sputum samples were obtained at exacerbation and stable visits during follow-up. Dominant genus was assigned based on a previously published method. Microbiome characteristics were analysed for their association with measures of clinical disease severity and long-term outcomes using PERMANOVA, Random Forest and survival analyses.

Results: In both stable bronchiectasis and during exacerbations, a microbiome dominated by Proteobacteria and Firmicutes was observed. Analysis of weighted UniFrac values indicated that patients' microbiome profiles were relatively stable over time, both during exacerbations and at disease stability, and showed less variation than that observed between patients. Random Forest analysis on baseline samples identified the genera *Pseudomonas*, Enterobacteriaceae and *Stenotrophomonas* as being associated with a BSI score ≥ 9 , whilst *Pseudomonas* and Enterobacteriaceae were also associated with ≥ 3 exacerbations. Patients in whom the dominant genus was *Pseudomonas* were at increased risk of all-cause mortality (Hazard ratio 3.12, 95% CI 1.33-7.36, $P=0.0091$) and had more frequent exacerbations (Incident rate ratio 1.69 95% CI 1.07-2.67, $P=0.024$) during follow-up.

Conclusion: A reduction in microbiome diversity, particularly associated with dominance of *Pseudomonas* and other Proteobacteria, is associated with greater disease severity, increased frequency and severity of exacerbations, and a higher risk of mortality.

Funding: Funded by the GSK/British Lung Foundation Chair of Respiratory Research and European Respiratory Society through the EMBARC2 consortium.

Introduction

Bronchiectasis is characterised pathophysiologically by a complex interaction between impaired mucociliary clearance, bacterial infection and lung inflammation which manifests clinically as a condition with variable symptoms of cough, sputum production and exacerbations. The prevalence of bronchiectasis worldwide is increasing and is associated with high medical healthcare costs and increased mortality; in Europe and North America, the prevalence is as high as 566 cases per 100,000 and estimates are even higher in China (1-4).

Despite an increase in our understanding of bronchiectasis, there are no licensed therapies and bronchiectasis remains a major clinical challenge. The disease heterogeneity in terms of its aetiology, clinical phenotype, inflammatory endotype and outcomes are the main reasons cited for unsuccessful clinical trials and the ongoing high burden of disease. Addressing this heterogeneity through translational research is currently the leading research priority in bronchiectasis.

Given the central role for infection in the pathogenesis of bronchiectasis, the microbiome represents a potential area for improved disease understanding and stratification. Several studies have described the bronchiectasis lung microbiome demonstrating some associations between diversity and lung function, and specific genera and exacerbation history. These studies have generally included <100 patients, had limited longitudinal data or have been part of one clinical trial, the BLESS study (5-7). The BLESS study included a subgroup of the bronchiectasis population using several inclusion and exclusion criteria, including a history of at least 2 exacerbations per year.

Research questions that remain in bronchiectasis include whether microbiome diversity or composition is associated with different clinical phenotypes or characteristics such as lung function or symptoms. It is not fully understood how the microbiome changes at exacerbation. Most importantly, while it is well established that chronic infection with *Pseudomonas aeruginosa* and other pathogens identified by culture are associated with significantly increased exacerbations and mortality, there are no microbiome studies which have attempted to stratify patients into clinically meaningful subtypes which could be

105 associated with long-term outcomes (8, 9). We hypothesised that patient outcomes could be associated
106 with particular microbiome profiles when clinically stable.

107 To address this question, we conducted a longitudinal cohort study of a broad representative population
108 of patients with bronchiectasis.

Methods

Study Design

Patients included in this observational study were ≥ 18 years, have a high-resolution CT confirmed diagnosis of bronchiectasis and have clinical symptoms (cough, sputum production, dyspnoea or respiratory infections) consistent with bronchiectasis. Patients were excluded if they were unable to give informed consent or had active tuberculosis or lung cancer. Patients with cystic fibrosis or pulmonary fibrosis with secondary traction bronchiectasis were also excluded. Patients were enrolled between 2012 and 2015. Patients gave written informed consent and the study was approved by the East of Scotland Research Ethics committee (12/ES/0059). Aetiology was determined after a standardised set of investigations as recommended by British Thoracic Society guidelines (10).

Clinical Assessment and Sampling

Patients provided spontaneous sputum samples at enrolment, when clinically stable (defined as no antibiotics, apart from prophylactic antibiotics, in the previous 4 weeks). Clinical (including lung function testing) and quality of life assessments were carried out as described previously (11). Patients were asked to provide further sputum samples at the onset of exacerbations (prior to antibiotic therapy where possible, but otherwise within a maximum of 48 hours after starting antibiotic therapy). Additional sputum samples were obtained from a non-consecutive convenience subset of patients at yearly follow-up visits in order to evaluate changes over time in the stable sputum microbiome (Figure 1). Severity of disease was evaluated using the Bronchiectasis Severity Index (BSI) (12). Symptoms were evaluated using the quality of life bronchiectasis respiratory symptom scale (QOL-B-RSS), a validated instrument (13). Sputum inflammatory marker concentration (IL-1 β , IL-8, and TNF α) and sputum neutrophil elastase activity was measured as described in the online supplement.

Long term follow-up

Patients were followed up from enrolment to 8th January 2019. Patients in Scotland were identified using a unique identifier which allows recording of all hospital admissions and deaths within the country. Outcomes were therefore evaluated for all patients who attended final follow-up outpatients

visits in 2019 or using linkage to medical records for patients lost to follow-up prior to 2019, ensuring complete follow-up data for all participants. Self-reported exacerbations were verified using regional prescription data which records all antibiotic use.

Exacerbations

Exacerbations were defined according to the British Thoracic Society definition of exacerbation and were treated with 14 days of antibiotic therapy (10). Severe exacerbations were those requiring hospital admission or intravenous antibiotic therapy. Our analysis of changes in the microbiome from stable state to exacerbation therefore consisted of matched samples where patients had provided both stable and exacerbation samples, with only 1 stable and exacerbation sample included per patient (Figure 1).

Sputum Microbiome

DNA was extracted from sputum and the V3 and V4 region of the bacterial 16S rRNA gene, sequenced, and the resulting data quality checked for sequencing and contamination errors as described in the online supplement. Based on the recommendation from McMurdie & Holmes (14), non-rarefied alpha diversity was measured by determining the Shannon-Wiener Diversity Index (SWDI) in PAST3, whilst beta diversity (between sample diversity) was calculated by Weighted UniFrac. PERMANOVA (Permutational multivariate analysis of variance; an approach which calculated the total of the distances between individuals within each group, and compares this to the distribution of totals obtained by repeated random re-labellings of the data) was used to test for significant differences in R v3.5.1 utilising the vegan package (15-18). To look at the relative importance of each bacterial taxa to the observed differences in clinical characteristics without bias towards predefined taxa, random forests were constructed, using the randomForest R package (19). Each forest consisted of 500 different trees with equal weighting, replacement, majority voting and all taxa included as candidates at each split. Subsequently, stratification of the microbiome samples based on the dominant taxa, regardless of abundance, was performed using a previously published stratification system (20). All sequences generated are available in the NCBI Sequence Read Archive under the Bioproject accession numbers PRJNA539959, PRJNA548310 and PRJNA316126.

Statistical Analysis

Statistical analysis was carried out using GraphPad Prism v6, R version 3.5.1 (16) and SPSS v22. Categorical data was analysed by determining the frequency and percentage in each category, whilst continuous data was displayed using median and interquartile range. Associations between linear variables and clinical parameters was performed using Spearman's correlations, and between group differences were tested using ANOVA, the Kruskal-Wallis test or the Wilcoxon matched pairs signed rank test. For long term clinical outcome analyses, exacerbations were recorded across the entire follow-up period and analysed using a negative binomial generalised linear model with follow-up time an offset. Time to first severe exacerbation and all-cause mortality were evaluated using Cox proportional hazards models with the proportional hazard assumption verified. Lung function decline was assessed using a linear regression model. For the negative binomial model, mixed models, Cox models and adjusted PERMANOVA, variables of interest were adjusted for clinically relevant confounders (age, smoking status, macrolide use, inhaled antibiotic use, prior exacerbations and FEV₁) selected *a priori* as recommended by a recent consensus document (21). Other statistical analyses were unadjusted for confounders. Sensitivity analyses were performed adjusting for BSI alone and the PERMANOVA for sequencing batch effects. Significance for statistical testing was set at $P < 0.05$. Clinically important endpoints assessed in the study were all cause mortality, time to severe exacerbation requiring hospitalisation, exacerbation frequency, FEV₁% predicted and change in FEV₁ over time, BSI score and QOL-B-RSS. Although the study related microbiome diversity and dominance with these multiple endpoints, adjustment for multiple comparisons was not performed as these endpoints all reflect bronchiectasis disease severity and burden and are not considered independent.

Role of the Funding Source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the article. All data is accessible to all the authors in this study. The corresponding author had full access to all the data and the final responsibility to submit for publication. AJD, ML and JDC had access to the raw data.

Results

Stable Microbiome and relationships with disease severity.

Three hundred patients provided at least 1 stable sample during the study and were included in the stable baseline cohort (Table S1 online), with 16S rRNA gene sequencing data being obtained from 281 patients after quality controls were applied to the sequence data (Figure 1, further details of number of sequences obtained and the negative controls are given in Figure S1 online). 52% of the cohort was female, with a median age of 69 and the median and interquartile range FEV₁% predicted was 76.0 (54.8-95.7) (Table 1).

Table 1: Patient characteristics of the stable cohort providing microbiome data (n=281). Data is presented as median and interquartile range or number and percentage of the cohort. Abbreviations: ICS: inhaled corticosteroids, BMI: Body mass index, FEV: forced expiratory volume, FVC: forced vital capacity, BSI: bronchiectasis severity index, QOL-B-RSS: quality of life- bronchiectasis, respiratory symptom scale, COPD: chronic obstructive pulmonary disease, ABPA: allergic bronchopulmonary aspergillosis

Characteristics	Baseline Microbiome Cohort Median (IQR) or N(%)
N	281
Age	69 (64-76)
Female Gender	147 (52.3%)
ICS use	169 (60.1%)
Long term macrolide use	98 (34.9%)
Inhaled antibiotics	14 (5.0%)
BMI	27 (23.7-31.1)
Exacerbation frequency (year prior to the study)	
0	45 (16.0%)
1	58 (20.6%)
2	53 (18.9%)
3 or more	125 (44.5%)
Severe exacerbation requiring hospitalisation (year prior to the study)	52 (18.5%)
Daily Sputum volume (ml)	15 (5-30)
Spirometry	
FEV ₁ (L)	1.78 (1.24-2.36)
FEV ₁ (% predicted)	76.0 (54.8-95.7)
FVC (L)	2.75 (2.16-3.48)
FEV ₁ /FVC	64.6 (53.4-74.7)
BSI Severity	
Mild	72 (25.6%)
Moderate	109 (38.8%)
Severe	100 (35.6%)
QOL-B-RSS score	66.7 (48.1-77.8)

Aetiology	
Idiopathic	144 (51.2%)
COPD	33 (11.7%)
ABPA	13 (4.6%)
Post- infective	30 (10.7%)
Rheumatoid arthritis	10 (3.6%)
Inflammatory bowel disease	9 (3.2%)
Other	42 (14.9%)

The baseline stable microbiome profiles are shown in figure 2A. The two most prevalent phyla were Proteobacteria and Firmicutes, whilst the most prevalent genera were *Haemophilus*, *Pseudomonas* and *Streptococcus*. Details about which OTUs were present in at least 75% of the baseline stable cohort (the core microbiome) are shown in Figure S2 online. Based on the % of the Operational Taxonomic Units (OTUs- a group of sequences with 97% sequence similarity and identified as likely to belong to a particular genus) for each bacterial taxa, 86 patients were identified as having a microbiome where the most abundant taxa was *Haemophilus*, 74 were *Streptococcus* dominated, 35 were *Pseudomonas*, 23 *Veillonella*, 16 *Neisseria*, 11 *Moraxella*, and 10 *Stenotrophomonas* with the median and IQR % sequences for each taxa shown in Table S2. While the data, shown on a PCoA plot using the Weighted UniFrac distances (Figure 2B), from groups with different BSI score overlap, the centres (mean values) of these clusters were distinct ($P=0.00010$, by PERMANOVA). This separation remained statistically significant after adjusting for confounders, ($P=0.023$, after adjustment for age, smoking status, macrolide use, inhaled antibiotic use, prior exacerbations and FEV₁). Utilising the random forest method and only showing those taxa with a mean decrease in Gini of >1.5 , we identified the taxa *Stenotrophomonas*, *Pseudomonas* and Enterobacteriaceae as being associated with severe bronchiectasis, defined by a BSI score ≥ 9 (Figure 2C). We correlated SWDI, as a marker of reduced microbiome alpha diversity, with measures of disease severity and quality of life such as FEV₁% predicted, BSI score, QOL-B-RSS, with no adjustments for confounders. Lower FEV₁% predicted showed a weak ($r=0.23$) but significant association ($P<0.0001$) with SWDI. Reductions in microbiome diversity, indicated by reduced SWDI, in the stable microbiome were also associated with both a more severe BSI score ($r=0.29$, $P<0.0001$) and a lower quality of life, as measured by QOL-B-RSS ($r=0.23$,

P=0.00050). A sensitivity analysis was performed after rarefying each sample to a read depth of 2725 with similar results (Table S3).

Next, we investigated whether the patient's exacerbation history was associated with the profile of their microbiome. Categorising the baseline weighted UniFrac plot according to whether the patient had 3 or more exacerbations in the previous year, no significant association between microbiome beta diversity and exacerbation frequency based on PERMANOVA was shown (Figure 3A) whilst the association between reduced SWDI and increased exacerbation frequency was not statistically significant (P=0.051, Kruskal-Wallis test, Figure 3B). Again, utilising the random forest method, only two taxa (Enterobacteriaceae and *Pseudomonas*) with a mean decrease in Gini of greater than 1.5 were associated with more frequent exacerbations (Figure 3C).

Longitudinal Changes between Stable and Exacerbation Microbiomes

49 patients provided more than 1 stable sample, for these patients, the median number of stable samples was 2, and the range in samples provided was 2 to 4. The microbiome profiles of these longitudinal stable samples are shown in Figure S4 online and the Weighted UniFrac values, both for within patient sample pairs and between stable samples from different patients indicated that a patient's stable microbiome is most similar to itself.

Exacerbation samples from 64 patients were analysed and compared to their corresponding stable sample; for 11 of the patients the exacerbation sample was collected prior to the stable sample. The median time between stable and exacerbation samples was 149 (IQR 52.25- 337.75) days. Differences in the microbiomes was determined by calculating the Weighted UniFrac, both for within patient sample pairs and between stable and exacerbation samples from different patients (Figure S5). A higher degree of similarity between stable and exacerbation samples from the same patient compared to those stable and exacerbation samples pairs from different patients was observed (Figure 4A, P=0.0054 Kolmogorov-Smirnov test). Whilst it appeared there was a trend to a reduction in SWDI at exacerbation (Figure 4B), this was statistically non-significant by Wilcoxon matched-pairs signed rank test. Using random forest analysis, multiple oral taxa were associated with stability including *Fusobacterium*, *Rothia*, *Prevotella* and *Streptococcus*. Only two bacterial taxa (*Eikenella* and *Granulicatella*) were shown to be associated

with exacerbation. (Figure 4C). Intriguingly, the two classical pathogenic taxa *Pseudomonas* and *Haemophilus* were associated with stable state rather than exacerbation (Figure 4C).

Decline in FEV₁ is not associated with reduced microbiome alpha diversity

261 patients had at least 1 longitudinal FEV₁ measurement at least 1 year after the first study visit. The mean was 3 spirometry values/patient (range 1-16). The median FEV₁ decline across the whole population was 31mL (IQR 3-73mL/year). Modelling decline in FEV₁ over time, we observed that baseline FEV₁ was a strong predictor of future FEV₁ decline and so all models included baseline FEV₁ and severity of disease using the BSI. SWDI was associated with a 12mL faster decline (95% CI -2-30mL, P=0.089) but this was not statistically significant. We also observed a slightly more rapid decline in patients with a SWDI <1 as reported by Woo *et al* (22), but the effect was not statistically significant (16mL decline, 95% CI -9-41mL, P=0.23).

Associations between sputum inflammatory markers and the microbiome

As higher inflammatory markers are associated with increased exacerbations and mortality, we first investigated whether baseline microbiome profiles were associated with inflammatory markers in sputum. Soluble sputum inflammatory marker concentration (IL-1 β (N=179), IL-8 (N=177), and TNF α (N=168)) and neutrophil elastase activity (N=218) was measured. The loadings plot for the PCoA in figure 2B is shown in Figure S6A, alongside a heatmap of Spearman correlations (Figure S6B) indicating that elastase activity in sputum was positively associated with increased levels of *Pseudomonas* and *Stenotrophomonas*. When stable samples were classified based on the taxa with the highest OTU%, between group differences were observed for TNF- α (P=0.0068), and neutrophil elastase (P<0.0001) with comparisons by Kruskal-Wallis test.

Differences in the Microbiome are Associated with Long-term Outcomes

For long term outcomes, the median duration of follow-up per subject was 1152 days (IQR 909-1634 days). Based on the observation that the Weighted UniFrac values of the stable cohort clustered on the PCoA plot by which bacterial taxa had the highest OTU% (Figure 5A), and the indication in the Forest plot that certain taxa were associated with mortality (Figure S7A), mortality rates and time to hospitalisation due to exacerbation were assessed with the patients' microbiome classified as being

dominated by *Haemophilus*, *Streptococcus*, *Veillonella*, *Pseudomonas*, or *Stenotrophomonas*. All other less common taxa were combined into an ‘Other’ group. Fifty-seven patients (20.3%) died during follow-up. The unadjusted Kaplan-Meier survival curve (Figure 5B) shows a clear association between dominant microbiome profiles and survival ($P < 0.0001$). In the adjusted Cox model, only *Pseudomonas* was independently associated with an increased risk of death- (Hazard ratio (HR) 3.12 (1.33-7.36), $P = 0.0091$). Full results including unadjusted and adjusted models for each genus are shown in Table S4.

We recorded 1436 exacerbation events during 1090 years of cumulative patient follow-up. In the adjusted model, *Pseudomonas* dominated microbiome profiles had a higher frequency of exacerbations during follow-up (incidence rate ratio (IRR) 1.69 (95% CI 1.07-2.67), $P = 0.024$). No other profiles were significantly associated with future exacerbation frequency (Table S5). 101 (35.9%) patients had at least one severe exacerbation during follow-up. The Kaplan-Meier survival analysis found a clear difference between genera in terms of hospitalisation risk ($P < 0.0001$ by log rank test), Figure 5C. In the adjusted Cox model, with the *Streptococcus* dominated microbiome group as the reference, only *Pseudomonas* dominated microbiomes were independently associated with increased hospitalisations- (HR 3.41 (95% CI 1.68-6.92), $P = 0.0011$, Table S6).

Finally, models were repeated with alpha diversity measures as predictors of future outcomes. SWDI was not significantly associated with survival when entered as a continuous variable (HR 0.79 (95% CI 0.56-1.12), $P = 0.18$), but was associated when examining the difference between groups above and below the median value of 1.25 ($P = 0.022$, Figure 5D). SWDI was associated with future risk of exacerbation (IRR (unadjusted), 0.79 (95% CI 0.66-0.93), $P = 0.0070$, IRR (adjusted) 0.75 (95% CI 0.63-0.90), $P = 0.0024$) when entered as a continuous variable. A separate model, in which prior exacerbations were included as a covariate, still found an independent effect of SWDI on exacerbation risk IRR 0.77 (95% CI 0.54-0.92, $P = 0.0041$). SWDI was associated with hospitalisation risk (HR 0.59 (95% CI 0.45-0.77), $P < 0.0001$) (Figure 5E) and the relationship with hospitalisation persisted after multivariate adjustment (HR 0.69 (95% CI 0.52-0.91), $P = 0.010$). When analysed by PERMANOVA, the microbiota of those that died was significantly different to the survivors ($P = 0.015$).

Discussion

We report, to the authors knowledge, the largest study of the bronchiectasis sputum microbiome to date, including 281 patients during disease stability with a median of over 3 years follow-up. We aimed to investigate how the sputum microbiome relates to severity of disease and long-term outcomes in bronchiectasis. We identified that severe bronchiectasis was associated with enrichment of several taxa belonging to the Proteobacteria phylum, particularly *Pseudomonas*. This is unsurprising since chronic infection with *Pseudomonas aeruginosa*, based on culture, is known to be linked to worse clinical outcomes, however 16S rRNA gene sequencing does not provide resolution to the species level (12, 23). Nevertheless, we identified that other Proteobacteria, such as Enterobacteriaceae and *Stenotrophomonas*, also appeared to be associated with worse symptoms. The two most common dominant genera identified were *Haemophilus* and *Streptococcus* and these were not strongly associated with more severe disease. Examining the data by dominant genus ignores a large amount of the information gathered by 16S rRNA gene sequencing, and previous work has suggested that alpha diversity metrics such as SWDI, which measures the richness and evenness of the taxa within a sample, was associated with lung function in one study but not in another (7, 22). We found a weak correlation between SWDI and lung function, symptoms and severity of disease suggesting that loss of diversity is associated with more severe disease. The mechanisms of reduced microbiome diversity in bronchiectasis requires further study, but antibiotic therapy is associated with lower diversity and therefore repeated exposure to antibiotics is a plausible mechanism, as is interspecies competition particularly in the context of organisms such as *Pseudomonas* (24). Consistent with the finding that *Pseudomonas* and other Proteobacteria were associated with worse outcomes (23), we confirmed the key role of *Pseudomonas* as a predictor of poor outcomes in bronchiectasis by showing that *Pseudomonas* dominant microbiome profiles were independently associated with exacerbation frequency, hospitalisation risk and all-cause mortality. This is the first time that long-term outcomes in bronchiectasis have been associated with microbiome characteristics. Furthermore, lower microbiome diversity was associated with exacerbations and hospital admissions during follow-up suggesting that this may be a “biomarker” of future outcome. The mechanisms by which alterations in the lung microbiome leads to exacerbation requires further study, which we briefly explored by assessing sputum inflammatory markers in a subset of the stable samples, but we acknowledge that because past

exacerbations are the strongest predictor of future exacerbations there is a possibility of reverse causation whereby prior exacerbations lead to lower diversity through antibiotic treatment, and therefore predict future events (25). We addressed this question by conducting a multivariable analysis adjusting for prior exacerbation events and found a persistent independent effect of SWDI on future exacerbation risk. Our results are somewhat different to the only other study to examine the ability of the microbiome to predict exacerbations, the BLESS cohort (20). In that cohort *Pseudomonas* (N=12 patients) and *Veillonella* (N=8 patients) dominated profiles were associated with a higher frequency of exacerbations. We confirmed that *Pseudomonas* is associated with greater exacerbation frequency but did not find any relationship between *Veillonella* and exacerbations. It should be noted that our sample size was greater and the BLESS study only recruited patients with at least 2 exacerbations in the prior year, and therefore may have selected for a subset of the bronchiectasis population with a *Veillonella* dominated microbiome with a higher risk of exacerbation.

While we found highly diverse microbiome profiles during exacerbations of bronchiectasis, consistent with previous findings by Cox et al (7), we found no single pattern of changes during exacerbation and analysis of the Weighted UniFrac measure of dissimilarity suggests a high degree of stability in the microbiome within patients who provided additional stable and exacerbation samples. We found an association between oral taxa and stability compared to exacerbation, but intriguingly we also found an association between the classical pathogenic taxa *Pseudomonas* and *Haemophilus* and stability. This is counter intuitive, as isolation of these taxa is classically used to define a “bacterial” exacerbation. Our results suggest that 16S rRNA gene sequencing alone is insufficient to explain the changes that occur that lead to exacerbation symptoms. Integration of additional data in future studies such as bacterial load, change in strain and viral data will be needed. Multiple pieces of evidence now point towards a more complex pattern to bronchiectasis exacerbations with the majority not being bacterial. Our previous work found only around 50% of exacerbation events were associated with a positive bacterial culture (26), whilst other studies suggest up to 50% of events are associated with isolation of a virus (27). The failure of inhaled antibiotic therapies to give consistent benefits in terms of reducing exacerbations despite substantial reductions in bacterial load also suggests that our existing paradigm

that regards Gram-negative bacteria as the primarily cause of exacerbations may be wrong (28). We demonstrate that many exacerbation events were not associated with changes in the microbiome profile. Our study was not designed to answer what caused the exacerbations in these cases and future studies are required to endotype these events.

The collection of sputum is non-invasive and is therefore used routinely in the care of patients with bronchiectasis to evaluate the presence of airway infection. In the context of microbiome research, it allows large cohorts, such as this, to be investigated which would not be possible if all subjects had to undergo multiple bronchoscopies. It nevertheless has limitations including potential contamination from the upper airway. Sputum is regarded as an intermediate between the upper and lower airway; we make no representations about the ecology of the lower respiratory tract in this study (29). Prior bronchiectasis studies have nevertheless shown a high degree of concordance between sputum and BAL (30). The other limitation of sputum is being unable to obtain samples in a subset of patients that are not productive. Studies using spontaneous sputum are inevitably biased towards patients who are able to produce high volume spontaneous sputum samples. Future studies may wish to use methods such as sputum induction or more invasive sampling to remove this selection bias. We did not evaluate bacterial load, which has recently emerged as an important predictor of treatment response in bronchiectasis, and this may provide complementary information in future studies (31). Although we have shown associations between the microbiome and clinical parameters, we cant prove causation, and further studies are required to determine if modifying the microbiome would improve clinical outcomes. Although to the authors knowledge this is the largest study of the sputum microbiome in bronchiectasis the sample size for individual subgroups was still relatively small and confidence intervals wide. Finally, we have recently reported that the fungal mycobiome was associated with increased frequency of exacerbations in the presence of high conidial loads of *Aspergillus* species (32). Future studies should seek to incorporate multiple omics methods including the characterisation of fungi and viruses to achieve a comprehensive assessment of disease endotypes.

In summary, we report to the best of the authors knowledge, the largest microbiome study in bronchiectasis to date and the first to examine its relationship with long term clinical outcomes. Reduced

microbiome diversity, particularly in the presence of *Pseudomonas* dominance, is associated with greater disease severity, increased exacerbations, increased severe exacerbations and a higher risk of all-cause mortality. Our analysis suggests the need for further studies to characterise different endotypes of bronchiectasis exacerbations and research to understand the mechanisms of reduced microbiome diversity and how this can be modulated.

Contributors

This study was conceived by AJD, ML, HRK and JDC. Data collection was performed by AJD, ML, HRK, AHS and JDC. Laboratory analysis was performed by AJD, HRK, AHS, JP, SF, AJC, JTJH and JDC. Data analysis and interpretation was carried out by AJD, ML, HRK, AHS, JTJH and JDC. The manuscript was drafted by AJD, ML and JDC. All authors reviewed, contributed to, and approved the final version of the article

Declaration of Interest

JDC reports research grants from AstraZeneca, Boehringer-Ingelheim, Chiesi, Gilead Sciences, Glaxosmithkline, Insmmed, Novartis and Zambon. All other authors declared no conflicts of interest

Data sharing

Raw sequencing files are accessible on the NCBI SRA upon publication of this paper under the Bioproject accession numbers PRJNA539959, PRJNA548310 and PRJNA316126.

Acknowledgements

We wish to acknowledge Brandon Tan and Christopher J Fong for their help in analysing the sputum inflammatory markers. We acknowledge funding from the British Lung Foundation and the European Respiratory Society through the EMBARC2 consortium.

References

1. Weycker D, Hansen GL, Seifer FD. Prevalence and incidence of noncystic fibrosis bronchiectasis among US adults in 2013. *Chronic Respir Dis*. 2017;14(4):377-84.
2. Monteagudo M, Rodriguez-Blanco T, Barrecheguren M, Simonet P, Miravittles M. Prevalence and incidence of bronchiectasis in Catalonia, Spain: A population-based study. *Respir Med*. 2016;121:26-31.
3. Quint JK, Millett ERC, Joshi M, Navaratnam V, Thomas SL, Hurst JR, et al. Changes in the incidence, prevalence and mortality of bronchiectasis in the UK from 2004 to 2013: a population-based cohort study. *Eur Respir J*. 2016;47(1):186-93.
4. Ringshausen FC, de Roux A, Diel R, Hohmann D, Welte T, Rademacher J. Bronchiectasis in Germany: a population-based estimation of disease prevalence. *Eur Respir J*. 2015;46(6):1805-7.
5. Rogers GB, Bruce KD, Martin ML, Burr LD, Serisier DJ. The effect of long-term macrolide treatment on respiratory microbiota composition in non-cystic fibrosis bronchiectasis: an analysis from the randomised, double-blind, placebo-controlled BLESS trial. *Lancet Respir Med*. 2014;2(12):988-96.
6. Taylor SL, Rogers GB, Chen AC, Burr LD, McGuckin MA, Serisier DJ. Matrix metalloproteinases vary with airway microbiota composition and lung function in non-cystic fibrosis bronchiectasis. *Ann Am Thorac Soc*. 2015;12(5):701-7.
7. Cox MJ, Turek EM, Hennessy C, Mirza GK, James PL, Coleman M, et al. Longitudinal assessment of sputum microbiome by sequencing of the 16S rRNA gene in non-cystic fibrosis bronchiectasis patients. *PLoS One*. 2017;12(2):e0170622.
8. Loebinge MR, Wells AU, Hansell DM, Chinyanganya N, Devaraj A, Meister M, et al. Mortality in bronchiectasis: a long-term study assessing the factors influencing survival. *Eur Respir J*. 2009;34(4):843-9.
9. Finch S, McDonnell MJ, Abo-Leyah H, Aliberti S, Chalmers JD. A comprehensive analysis of the impact of *Pseudomonas aeruginosa* colonization on prognosis in adult bronchiectasis. *Ann Am Thorac Soc*. 2015;12(11):1602-11.
10. Pasteur MC, Bilton D, Hill AT. British Thoracic Society guideline for non-CF bronchiectasis. *Thorax*. 2010;65(7):577.
11. Chalmers JD, Moffitt KL, Suarez-Cuartin G, Sibila O, Finch S, Furrie E, et al. Neutrophil elastase activity is associated with exacerbations and lung function decline in bronchiectasis. *Am J Respir Crit Care Med*. 2017;195(10):1384-93.
12. Chalmers JD, Goeminne P, Aliberti S, McDonnell MJ, Lonni S, Davidson J, et al. The bronchiectasis severity index. An international derivation and validation study. *Am J Respir Crit Care Med*. 2014;189(5):576-85.
13. Quittner AL, O'Donnell AE, Salathe MA, Lewis SA, Li X, Montgomery AB, et al. Quality of Life Questionnaire-Bronchiectasis: final psychometric analyses and determination of minimal important difference scores. *Thorax*. 2015;70(1):12-20.
14. McMurdie PJ, Holmes S. Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput Biol*. 2014;10(4):e1003531.
15. Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral Ecol*. 2001;26(1):32-46.
16. Team TRC. R: A language and environment for statistical computing Vienna Austria: R foundation for statistical computing; 2018 Available from: <http://www.R-project.org>.
17. Oksanen JB, F.G.; Friendly,M.; Kindt,R.; Legendre,P.; McGlinn,D.; Minchin,P.R.; O'Hara,R.B.; Simpson,G.L.; Solymos,P.; Stevens,M.H.H.; Szoecs,E.; Wagner,H.;. *vegan: Community Ecology Package* 2019 Available from: <https://CRAN.R-project.org/package=vegan>.
18. McArdle BH, Anderson MJ. Fitting multivariate models to community data: A comment on distance-based redundancy analysis. *Ecology*. 2001;82(1):290-7.
19. Liaw AW, M. Classification and regression by randomForest. *R News*. 2002;2/3:18-22.

20. Rogers GB, Zain NM, Bruce KD, Burr LD, Chen AC, Rivett DW, et al. A novel microbiota stratification system predicts future exacerbations in bronchiectasis. *Ann Am Thorac Soc*. 2014;11(4):496-503.
21. Lederer DJ, Bell SC, Branson RD, Chalmers JD, Marshall R, Maslove DM, et al. Control of confounding and reporting of results in causal inference studies. Guidance for authors from editors of respiratory, sleep, and critical care journals. *Ann Am Thorac Soc*. 2019;16(1):22-8.
22. Woo TE, Lim R, Heirali AA, Acosta N, Rabin HR, Mody CH, et al. A longitudinal characterization of the Non-Cystic Fibrosis Bronchiectasis airway microbiome. *Sci Rep*. 2019;9(1):6871.
23. Araujo D, Shteinberg M, Aliberti S, Goeminne PC, Hill AT, Fardon TC, et al. The independent contribution of *Pseudomonas aeruginosa* infection to long-term clinical outcomes in bronchiectasis. *Eur Respir J*. 2018;51(2).
24. Rogers GB, van der Gast CJ, Serisier DJ. Predominant pathogen competition and core microbiota divergence in chronic airway infection. *ISME J*. 2015;9(1):217-25.
25. Chalmers JD, Aliberti S, Filonenko A, Shteinberg M, Goeminne PC, Hill AT, et al. Characterization of the "frequent exacerbator phenotype" in bronchiectasis. *Am J Respir Crit Care Med*. 2018;197(11):1410-20.
26. Shoemark A, Cant E, Carreto L, Smith A, Oriano M, Keir HR, et al. A point of care neutrophil elastase activity assay identifies bronchiectasis severity, airway infection and risk of exacerbation. *Eur Respir J*. 2019;1900303.
27. Gao YH, Guan WJ, Xu G, Lin ZY, Tang Y, Lin ZM, et al. The role of viral infection in pulmonary exacerbations of bronchiectasis in adults: a prospective study. *Chest*. 2015;147(6):1635-43.
28. Haworth CS, Bilton D, Chalmers JD, Davis AM, Froehlich J, Gonda I, et al. Inhaled liposomal ciprofloxacin in patients with non-cystic fibrosis bronchiectasis and chronic lung infection with *Pseudomonas aeruginosa* (ORBIT-3 and ORBIT-4): two phase 3, randomised controlled trials. *Lancet Respir Med*. 2019;7(3):213-26.
29. Sulaiman I, Wu BG, Li Y, Scott AS, Malecha P, Scaglione B, et al. Evaluation of the airway microbiome in nontuberculous mycobacteria disease. *Eur Respir J*. 2018;52(4).
30. Rogers GB, van der Gast CJ, Cuthbertson L, Thomson SK, Bruce KD, Martin ML, et al. Clinical measures of disease in adult non-CF bronchiectasis correlate with airway microbiota composition. *Thorax*. 2013;68(8):731-7.
31. Sibila O, Laserna E, Shoemark A, Keir HR, Finch S, Rodrigo-Troyano A, et al. Airway bacterial load and inhaled antibiotic response in bronchiectasis. *Am J Respir Crit Care Med*. 2019;200(1):33-41.
32. Mac Aogain M, Chandrasekaran R, Lim AYH, Low TB, Tan GL, Hassan T, et al. Immunological corollary of the pulmonary mycobome in bronchiectasis: the CAMEB study. *Eur Respir J*. 2018;52(1).

Figure Legends

Figure 1: Flow chart of samples included in the study.

Figure 2: A: Overview of the stable bronchiectasis microbiome grouped according to Bronchiectasis Severity Index (BSI) then sorted according to percentage of Proteobacteria in each sample. Each patient is represented once by a vertical stacked bar, samples identified to Phyla (top) and Genera (bottom) level. Versions of these graphs but with patients sorted according to BSI and then percentage *Haemophilus* are shown in Figure S3. **B:** Weighted UniFrac PCoA plot showing clustering of stable patient samples according to mild (blue circle), moderate (red square) and severe (black triangle) BSI. Coloured circles represent two standard errors around the means for each group. Differences in groups was tested using PERMANOVA **C:** Random forest plot indicating the bacterial taxa associated with a more severe BSI score (black bars) and a less severe BSI score (blue bars). Only those taxa with a mean decrease Gini >1.5 are shown. N=281 for all analyses.

Figure 3: A: Weighted UniFrac PCoA plot comparing non-frequent exacerbating stable patient samples (blue circles) with frequent exacerbating stable patient samples (red triangles). Coloured circles represent two standard errors around the means for each group. Differences in groups was tested using PERMANOVA. **B:** Reduced alpha diversity (lower SWDI) is not significantly associated with increased prior exacerbations. Data are presented as mean with standard deviation **C:** Random Forest plot showing the bacterial taxa associated with more than 3 exacerbations in the previous year (red bars), and those associated with fewer exacerbations (blue bars). Only those taxa with a mean decrease in Gini greater than 1.5 are shown. N=281 for all analyses.

Figure 4: A: Between sample differences in the microbiome, as measured by Weighted UniFrac, comparing 1 stable and 1 exacerbation sample per patient (N=64 patients, 128 samples). Grey plot represents differences between stable and exacerbation samples from the same patient, black plot represents differences between stable and exacerbation samples from different patients. **B:** Change in alpha diversity as measured by Shannon-Wiener Diversity Index of n=64 pairs of 1 stable and 1 exacerbation sample per patient. **C:** Random Forest plot showing the bacterial taxa associated with exacerbation samples (red bars), and those associated with stable samples (blue bars) from n=64 pairs of 1 stable and 1 exacerbation sample per patient. Only those taxa with a mean decrease in Gini greater than 1 are shown.

Figure 5: Based on the clustering observed in the weighted UniFrac PCoA plot in (A) which showed patients clustered according to which taxa had the highest OTU%, mortality (B) and hospitalization (C) rates were compared by Kaplan Meier curves based on the dominant organism identified at baseline sampling. Baseline stable samples were classified as above or below the median Shannon-Wiener Diversity Index and compared to mortality (D) and hospitalization (E) rates. N=281 for all analyses.