University of Dundee

The sputum microbiome and clinical outcomes in patients with bronchiectasis

Dicker, Alison J.; Lonergan, 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

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., Lonergan, 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. Advance online publication. 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.

Download date: 03. Oct. 2023

Dicker, Alison J. et al. "The sputum microbiome and clinical outcomes in patients with bronchiectasis: a prospective observational study". The Lancet Respiratory Medicine. 2021, 9(8). 885-896. https://doi.org/10.1016/S2213-2600(20)30557-9
Released under the terms of CCBYNCND license

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

3 Alison J. Dicker PhD[1], Mike Lonergan PhD[1], Holly R Keir BSc[1], Alexandria H. Smith[2], Jennifer
4 Pollock[1], Simon Finch MBChB[1], Andrew J. Cassidy PhD[1], Jeffrey T. J. Huang PhD[3], Prof James D.
5 Chalmers PhD[1]

6 [1]Scottish Centre for Respiratory Disease Research, University of Dundee, Dundee, UK

7 [2]University Cambridge, Cambridge, UK

8 [3]Division of Systems Medicine, School of Medicine, University of Dundee, Ninewells Hospital and
9 Medical School, Dundee, UK

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

12 Word count: 3869

13 Keywords: Microbiome, bronchiectasis, *Pseudomonas*, mortality

14

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

18 **Research in Context Panel**

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

32 **Added value of this study:**

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

47 **Implications of all available evidence:**

48 Disease severity and long-term mortality is associated with loss of microbial diversity and dominance
49 of Proteobacteria such as *Pseudomonas*, Enterobacteriaceae and *Stenotrophomonas* in patients with
50 bronchiectasis. In contrast to the established paradigm that bronchiectasis exacerbations are primarily
51 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.

80 **Introduction**

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

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

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

99 Research questions that remain in bronchiectasis include whether microbiome diversity or composition
100 is associated with different clinical phenotypes or characteristics such as lung function or symptoms. It
101 is not fully understood how the microbiome changes at exacerbation. Most importantly, while it is well
102 established that chronic infection with *Pseudomonas aeruginosa* and other pathogens identified by
103 culture are associated with significantly increased exacerbations and mortality, there are no microbiome
104 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.

109 **Methods**

110 **Study Design**

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

119 **Clinical Assessment and Sampling**

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

131 **Long term follow-up**

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

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

138 **Exacerbations**

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

144 **Sputum Microbiome**

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

161 **Statistical Analysis**

162 Statistical analysis was carried out using GraphPad Prism v6, R version 3.5.1 (16) and SPSS v22.
163 Categorical data was analysed by determining the frequency and percentage in each category, whilst
164 continuous data was displayed using median and interquartile range. Associations between linear
165 variables and clinical parameters was performed using Spearman's correlations, and between group
166 differences were tested using ANOVA, the Kruskal-Wallis test or the Wilcoxon matched pairs signed
167 rank test. For long term clinical outcome analyses, exacerbations were recorded across the entire follow-
168 up period and analysed using a negative binomial generalised linear model with follow-up time an
169 offset. Time to first severe exacerbation and all-cause mortality were evaluated using Cox proportional
170 hazards models with the proportional hazard assumption verified. Lung function decline was assessed
171 using a linear regression model. For the negative binomial model, mixed models, Cox models and
172 adjusted PERMANOVA, variables of interest were adjusted for clinically relevant confounders (age,
173 smoking status, macrolide use, inhaled antibiotic use, prior exacerbations and FEV_1) selected *a priori*
174 as recommended by a recent consensus document (21). Other statistical analyses were unadjusted for
175 confounders. Sensitivity analyses were performed adjusting for BSI alone and the PERMANOVA for
176 sequencing batch effects. Significance for statistical testing was set at $P<0{\cdot}05$. Clinically important
177 endpoints assessed in the study were all cause mortality, time to severe exacerbation requiring
178 hospitalisation, exacerbation frequency, FEV_1% predicted and change in FEV_1 over time, BSI score
179 and QOL-B-RSS. Although the study related microbiome diversity and dominance with these multiple
180 endpoints, adjustment for multiple comparisons was not performed as these endpoints all reflect
181 bronchiectasis disease severity and burden and are not considered independent.

182 Role of the Funding Source

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

187 **Results**

188 **Stable Microbiome and relationships with disease severity.**

189 Three hundred patients provided at least 1 stable sample during the study and were included in the stable
190 baseline cohort (Table S1 online), with 16S rRNA gene sequencing data being obtained from 281
191 patients after quality controls were applied to the sequence data (Figure 1, further details of number of
192 sequences obtained and the negative controls are given in Figure S1 online). 52% of the cohort was
193 female, with a median age of 69 and the median and interquartile range FEV_1% predicted was 76·0
194 (54·8-95·7) (Table 1).

195 **Table 1**: Patient characteristics of the stable cohort providing microbiome data (n=281). Data is
196 presented as median and interquartile range or number and percentage of the cohort. Abbreviations:
197 ICS: inhaled corticosteroids, BMI: Body mass index, FEV: forced expiratory volume, FVC: forced vital
198 capacity, BSI: bronchiectasis severity index, QOL-B-RSS: quality of life- bronchiectasis, respiratory
199 symptom scale, COPD: chronic obstructive pulmonary disease, ABPA: allergic bronchopulmonary
200 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 1 2 3 or more	 45 (16·0%) 58 (20·6%) 53 (18·9%) 125 (44·5%)
Severe exacerbation requiring hospitalisation (year prior to the study)	52 (18·5%)
Daily Sputum volume (ml)	15 (5-30)
Spirometry FEV1 (L) FEV1 (% predicted) FVC (L) FEV1/FVC	 1·78 (1·24-2·36) 76·0 (54·8-95·7) 2·75 (2·16-3·48) 64·6 (53·4-74·7)
BSI Severity Mild Moderate Severe	 72 (25·6%) 109 (38·8%) 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%)

201

202 The baseline stable microbiome profiles are shown in figure 2A. The two most prevalent phyla were 203 Proteobacteria and Firmicutes, whilst the most prevalent genera were *Haemophilus*, *Pseudomonas* and 204 *Streptococcus*. Details about which OTUs were present in at least 75% of the baseline stable cohort (the 205 core microbiome) are shown in Figure S2 online. Based on the % of the Operational Taxonomic Units 206 (OTUs- a group of sequences with 97% sequence similarity and identified as likely to belong to a 207 particular genus) for each bacterial taxa, 86 patients were identified as having a microbiome where the 208 most abundant taxa was *Haemophilus*, 74 were *Streptococcus* dominated, 35 were *Pseudomonas*, 23 209 *Veillonella*, 16 *Neisseria*, 11 *Moraxell*a, and 10 *Stenotrophomonas* with the median and IQR % 210 sequences for each taxa shown in Table S2. While the data, shown on a PCoA plot using the Weighted 211 UniFrac distances (Figure 2B), from groups with different BSI score overlap, the centres (mean values) 212 of these clusters were distinct (P=0·00010, by PERMANOVA). This separation remained statistically 213 significant after adjusting for confounders, (P=0·023, after adjustment for age, smoking status, 214 macrolide use, inhaled antibiotic use, prior exacerbations and FEV_1. Utilising the random forest method 215 and only showing those taxa with a mean decrease in Gini of >1·5, we identified the taxa 216 *Stenotrophomonas, Pseudomonas* and Enterobacteriaceae as being associated with severe 217 bronchiectasis, defined by a BSI score ≥9 (Figure 2C). We correlated SWDI, as a marker of reduced 218 microbiome alpha diversity, with measures of disease severity and quality of life such as FEV_1% 219 predicted, BSI score, QOL-B-RSS, with no adjustments for confounders. Lower FEV_1% predicted 220 showed a weak (r=0·23) but significant association (P<0·0001) with SWDI. Reductions in microbiome 221 diversity, indicated by reduced SWDI, in the stable microbiome were also associated with both a more 222 severe BSI score (r=0·29, P<0·0001) and a lower quality of life, as measured by QOL-B-RSS (r=0·23,

223 P=0·00050). A sensitivity analysis was performed after rarefying each sample to a read depth of 2725
224 with similar results (Table S3).

225 Next, we investigated whether the patient's exacerbation history was associated with the profile of their
226 microbiome. Categorising the baseline weighted UniFrac plot according to whether the patient had 3 or
227 more exacerbations in the previous year, no significant association between microbiome beta diversity
228 and exacerbation frequency based on PERMANOVA was shown (Figure 3A) whilst the association
229 between reduced SWDI and increased exacerbation frequency was not statistically significant
230 (P=0·051, Kruskal-Wallis test, Figure 3B). Again, utilising the random forest method, only two taxa
231 (Enterobacteriaceae and *Pseudomonas*) with a mean decrease in Gini of greater than 1·5 were associated
232 with more frequent exacerbations (Figure 3C).

233 **Longitudinal Changes between Stable and Exacerbation Microbiomes**

234 49 patients provided more than 1 stable sample, for these patients, the median number of stable samples
235 was 2, and the range in samples provided was 2 to 4. The microbiome profiles of these longitudinal
236 stable samples are shown in Figure S4 online and the Weighted UniFrac values, both for within patient
237 sample pairs and between stable samples from different patients indicated that a patients stable
238 microbiome is most similar to itself.

239 Exacerbation samples from 64 patients were analysed and compared to their corresponding stable
240 sample; for 11 of the patients the exacerbation sample was collected prior to the stable sample. The
241 median time between stable and exacerbation samples was 149 (IQR 52·25- 337·75) days. Differences
242 in the microbiomes was determined by calculating the Weighted UniFrac, both for within patient sample
243 pairs and between stable and exacerbation samples from different patients (Figure S5). A higher degree
244 of similarity between stable and exacerbation samples from the same patient compared to those stable
245 and exacerbation samples pairs from different pairs was observed (Figure 4A, P=0·0054 Kolmogorov-
246 Smirnov test). Whilst it appeared there was a trend to a reduction in SWDI at exacerbation (Figure 4B),
247 this was statistically non-significant by Wilcoxon matched-pairs signed rank test. Using random forest
248 analysis, multiple oral taxa were associated with stability including *Fusobacterium*, *Rothia*, *Prevotella*
249 and *Streptococcus*. Only two bacterial taxa (*Eikenella* and *Granulicatella*) were shown to be associated

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

252 **Decline in FEV_1 is not associated with reduced microbiome alpha diversity**

253 261 patients had at least 1 longitudinal FEV_1 measurement at least 1 year after the first study visit. The
254 mean was 3 spirometry values/patient (range 1-16). The median FEV_1 decline across the whole
255 population was 31mL (IQR 3-73mL/year). Modelling decline in FEV_1 over time, we observed that
256 baseline FEV_1 was a strong predictor of future FEV_1 decline and so all models included baseline FEV_1
257 and severity of disease using the BSI. SWDI was associated with a 12mL faster decline (95% CI -2-
258 30mL, P=0·089) but this was not statistically significant. We also observed a slightly more rapid decline
259 in patients with a SWDI <1 as reported by Woo *et al* (22), but the effect was not statistically significant
260 (16mL decline, 95% CI -9-41mL, P=0·23).

261 **Associations between sputum inflammatory markers and the microbiome**

262 As higher inflammatory markers are associated with increased exacerbations and mortality, we first
263 investigated whether baseline microbiome profiles were associated with inflammatory markers in
264 sputum. Soluble sputum inflammatory marker concentration (IL-1β (N=179), IL-8 (N=177), and TNFα
265 (N=168)) and neutrophil elastase activity (N=218) was measured. The loadings plot for the PCoA in
266 figure 2B is shown in Figure S6A, alongside a heatmap of Spearman correlations (Figure S6B)
267 indicating that elastase activity in sputum was positively associated with increased levels of
268 *Pseudomonas* and *Stenotrophomonas*. When stable samples were classified based on the taxa with the
269 highest OTU%, between group differences were observed for TNF-α (P=0·0068), and neutrophil
270 elastase (P<0·0001) with comparisons by Kruskal-Wallis test.

271 **Differences in the Microbiome are Associated with Long-term Outcomes**

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

277 dominated by *Haemophilus*, *Streptococcus*, *Veillonella*, *Pseudomonas*, or *Stenotrophomonas*. All other
278 less common taxa were combined into an 'Other' group. Fifty-seven patients (20·3%) died during
279 follow-up. The unadjusted Kaplan-Meier survival curve (Figure 5B) shows a clear association between
280 dominant microbiome profiles and survival (P<0·0001). In the adjusted Cox model, only *Pseudomonas*
281 was independently associated with an increased risk of death- (Hazard ratio (HR) 3·12 (1·33-7·36),
282 P=0·0091). Full results including unadjusted and adjusted models for each genus are shown in Table
283 S4.
284 We recorded 1436 exacerbation events during 1090 years of cumulative patient follow-up. In the
285 adjusted model, *Pseudomonas* dominated microbiome profiles had a higher frequency of exacerbations
286 during follow-up (incidence rate ratio (IRR) 1·69 (95% CI 1·07-2·67), P=0·024). No other profiles were
287 significantly associated with future exacerbation frequency (Table S5). 101 (35·9%) patients had at
288 least one severe exacerbation during follow-up. The Kaplan-Meier survival analysis found a clear
289 difference between genera in terms of hospitalisation risk (P<0·0001 by log rank test), Figure 5C. In
290 the adjusted Cox model, with the *Streptococcus* dominated microbiome group as the reference, only
291 *Pseudomonas* dominated microbiomes were independently associated with increased hospitalisations-
292 (HR 3·41 (95% CI 1·68-6·92), P=0·0011, Table S6).
293 Finally, models were repeated with alpha diversity measures as predictors of future outcomes. SWDI
294 was not significantly associated with survival when entered as a continuous variable (HR 0·79 (95% CI
295 0·56-1·12), P=0·18), but was associated when examining the difference between groups above and
296 below the median value of 1·25 (P=0·022, Figure 5D). SWDI was associated with future risk of
297 exacerbation (IRR (unadjusted), 0·79 (95% CI 0·66-0·93), P=0·0070, IRR (adjusted) 0·75 (95% CI
298 0·63-0·90), P=0·0024) when entered as a continuous variable. A separate model, in which prior
299 exacerbations were included as a covariate, still found an independent effect of SWDI on exacerbation
300 risk IRR 0·77 (95% CI 0·54-0·92, P=0·0041). SWDI was associated with hospitalisation risk (HR 0·59
301 95% (95% CI 0·45-0·77), P<0·0001) (Figure 5E) and the relationship with hospitalisation persisted
302 after multivariate adjustment (HR 0·69 (95% CI 0·52-0·91), P=0·010). When analysed by
303 PERMANOVA, the microbiota of those that died was significantly different to the survivors (P=0·015)

304 **Discussion**

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

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

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

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

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

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

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

392 **Contributors**

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

398 **Declaration of Interest**

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

401 **Data sharing**

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

404 **Acknowledgements**

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

408 **References**

409 1. Weycker D, Hansen GL, Seifer FD. Prevalence and incidence of noncystic fibrosis
410 bronchiectasis among US adults in 2013. Chronic Respir Dis. 2017;14(4):377-84.
411 2. Monteagudo M, Rodriguez-Blanco T, Barrecheguren M, Simonet P, Miravitlles M. Prevalence
412 and incidence of bronchiectasis in Catalonia, Spain: A population-based study. Respir Med.
413 2016;121:26-31.
414 3. Quint JK, Millett ERC, Joshi M, Navaratnam V, Thomas SL, Hurst JR, et al. Changes in the
415 incidence, prevalence and mortality of bronchiectasis in the UK from 2004 to 2013: a population-based
416 cohort study. Eur Respir J. 2016;47(1):186-93.
417 4. Ringshausen FC, de Roux A, Diel R, Hohmann D, Welte T, Rademacher J. Bronchiectasis in
418 Germany: a population-based estimation of disease prevalence. Eur Respir J. 2015;46(6):1805-7.
419 5. Rogers GB, Bruce KD, Martin ML, Burr LD, Serisier DJ. The effect of long-term macrolide
420 treatment on respiratory microbiota composition in non-cystic fibrosis bronchiectasis: an analysis
421 from the randomised, double-blind, placebo-controlled BLESS trial. Lancet Respir Med.
422 2014;2(12):988-96.
423 6. Taylor SL, Rogers GB, Chen AC, Burr LD, McGuckin MA, Serisier DJ. Matrix metalloproteinases
424 vary with airway microbiota composition and lung function in non-cystic fibrosis bronchiectasis. Ann
425 Am Thorac Soc. 2015;12(5):701-7.
426 7. Cox MJ, Turek EM, Hennessy C, Mirza GK, James PL, Coleman M, et al. Longitudinal assessment
427 of sputum microbiome by sequencing of the 16S rRNA gene in non-cystic fibrosis bronchiectasis
428 patients. PLoS One. 2017;12(2):e0170622.
429 8. Loebinger MR, Wells AU, Hansell DM, Chinyanganya N, Devaraj A, Meister M, et al. Mortality
430 in bronchiectasis: a long-term study assessing the factors influencing survival. Eur Respir J.
431 2009;34(4):843-9.
432 9. Finch S, McDonnell MJ, Abo-Leyah H, Aliberti S, Chalmers JD. A comprehensive analysis of the
433 impact of Pseudomonas aeruginosa colonization on prognosis in adult bronchiectasis. Ann Am Thorac
434 Soc. 2015;12(11):1602-11.
435 10. Pasteur MC, Bilton D, Hill AT. British Thoracic Society guideline for non-CF bronchiectasis.
436 Thorax. 2010;65(7):577.
437 11. Chalmers JD, Moffitt KL, Suarez-Cuartin G, Sibila O, Finch S, Furrie E, et al. Neutrophil elastase
438 activity is associated with exacerbations and lung function decline in bronchiectasis. Am J Respir Crit
439 Care Med. 2017;195(10):1384-93.
440 12. Chalmers JD, Goeminne P, Aliberti S, McDonnell MJ, Lonni S, Davidson J, et al. The
441 bronchiectasis severity index. An international derivation and validation study. Am J Respir Crit Care
442 Med. 2014;189(5):576-85.
443 13. Quittner AL, O'Donnell AE, Salathe MA, Lewis SA, Li X, Montgomery AB, et al. Quality of Life
444 Questionnaire-Bronchiectasis: final psychometric analyses and determination of minimal important
445 difference scores. Thorax. 2015;70(1):12-20.
446 14. McMurdie PJ, Holmes S. Waste not, want not: why rarefying microbiome data is inadmissible.
447 PLoS Comput Biol. 2014;10(4):e1003531.
448 15. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral
449 Ecol. 2001;26(1):32-46.
450 16. Team TRC. R: A language and environment for statistical computing Vienna Austria: R
451 foundation for statistical computing; 2018 Available from: http://www.R-project.org.
452 17. Oksanen JB, F.G.; Friendly,M.; Kindt,R.; Legendre,P.; McGlinn,D.; Minchin,P.R.; O'Hara,R.B.;
453 Simpson,G.L.; Solymos,P.; Stevens,M.H.H.; Szoecs,E.; Wagner,H.;. vegan: Community Ecology Package
454 2019 Available from: https://CRAN.R-project.org/package=vegan.
455 18. McArdle BH, Anderson MJ. Fitting multivariate models to community data: A comment on
456 distance-based redundancy analysis. Ecology. 2001;82(1):290-7.
457 19. Liaw AW, M. Classification and regression by randomForest. R News. 2002;2/3:18-22.

458 20. Rogers GB, Zain NM, Bruce KD, Burr LD, Chen AC, Rivett DW, et al. A novel microbiota
459 stratification system predicts future exacerbations in bronchiectasis. Ann Am Thorac Soc.
460 2014;11(4):496-503.
461 21. Lederer DJ, Bell SC, Branson RD, Chalmers JD, Marshall R, Maslove DM, et al. Control of
462 confounding and reporting of results in causal inference studies. Guidance for authors from editors of
463 respiratory, sleep, and critical care journals. Ann Am Thorac Soc. 2019;16(1):22-8.
464 22. Woo TE, Lim R, Heirali AA, Acosta N, Rabin HR, Mody CH, et al. A longitudinal characterization
465 of the Non-Cystic Fibrosis Bronchiectasis airway microbiome. Sci Rep. 2019;9(1):6871.
466 23. Araujo D, Shteinberg M, Aliberti S, Goeminne PC, Hill AT, Fardon TC, et al. The independent
467 contribution of Pseudomonas aeruginosa infection to long-term clinical outcomes in bronchiectasis.
468 Eur Respir J. 2018;51(2).
469 24. Rogers GB, van der Gast CJ, Serisier DJ. Predominant pathogen competition and core
470 microbiota divergence in chronic airway infection. ISME J. 2015;9(1):217-25.
471 25. Chalmers JD, Aliberti S, Filonenko A, Shteinberg M, Goeminne PC, Hill AT, et al.
472 Characterization of the "frequent exacerbator phenotype" in bronchiectasis. Am J Respir Crit Care
473 Med. 2018;197(11):1410-20.
474 26. Shoemark A, Cant E, Carreto L, Smith A, Oriano M, Keir HR, et al. A point of care neutrophil
475 elastase activity assay identifies bronchiectasis severity, airway infection and risk of exacerbation. Eur
476 Respir J. 2019:1900303.
477 27. Gao YH, Guan WJ, Xu G, Lin ZY, Tang Y, Lin ZM, et al. The role of viral infection in pulmonary
478 exacerbations of bronchiectasis in adults: a prospective study. Chest. 2015;147(6):1635-43.
479 28. Haworth CS, Bilton D, Chalmers JD, Davis AM, Froehlich J, Gonda I, et al. Inhaled liposomal
480 ciprofloxacin in patients with non-cystic fibrosis bronchiectasis and chronic lung infection with
481 Pseudomonas aeruginosa (ORBIT-3 and ORBIT-4): two phase 3, randomised controlled trials. Lancet
482 Respir Med. 2019;7(3):213-26.
483 29. Sulaiman I, Wu BG, Li Y, Scott AS, Malecha P, Scaglione B, et al. Evaluation of the airway
484 microbiome in nontuberculous mycobacteria disease. Eur Respir J. 2018;52(4).
485 30. Rogers GB, van der Gast CJ, Cuthbertson L, Thomson SK, Bruce KD, Martin ML, et al. Clinical
486 measures of disease in adult non-CF bronchiectasis correlate with airway microbiota composition.
487 Thorax. 2013;68(8):731-7.
488 31. Sibila O, Laserna E, Shoemark A, Keir HR, Finch S, Rodrigo-Troyano A, et al. Airway bacterial
489 load and inhaled antibiotic response in bronchiectasis. Am J Respir Crit Care Med. 2019;200(1):33-41.
490 32. Mac Aogain M, Chandrasekaran R, Lim AYH, Low TB, Tan GL, Hassan T, et al. Immunological
491 corollary of the pulmonary mycobiome in bronchiectasis: the CAMEB study. Eur Respir J. 2018;52(1).

492

493 **Figure Legends**

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

495 **Figure 2: A**: Overview of the stable bronchiectasis microbiome grouped according to Bronchiectasis
496 Severity Index (BSI) then sorted according to percentage of Proteobacteria in each sample. Each patient
497 is represented once by a vertical stacked bar, samples identified to Phyla (top) and Genera (bottom)
498 level. Versions of these graphs but with patients sorted according to BSI and then percentage
499 *Haemophilus* are shown in Figure S3. **B:** Weighted UniFrac PCoA plot showing clustering of stable
500 patient samples according to mild (blue circle), moderate (red square) and severe (black triangle) BSI.
501 Coloured circles represent two standard errors around the means for each group. Differences in groups
502 was tested using PERMANOVA **C:** Random forest plot indicating the bacterial taxa associated with a
503 more severe BSI score (black bars) and a less severe BSI score (blue bars). Only those taxa with a mean
504 decrease Gini >1·5 are shown. N=281 for all analyses.

505 **Figure 3: A:** Weighted UniFrac PCoA plot comparing non-frequent exacerbating stable patient samples
506 (blue circles) with frequent exacerbating stable patient samples (red triangles). Coloured circles
507 represent two standard errors around the means for each group. Differences in groups was tested using
508 PERMANOVA. **B:** Reduced alpha diversity (lower SWDI) is not significantly associated with
509 increased prior exacerbations. Data are presented as mean with standard deviation **C:** Random Forest
510 plot showing the bacterial taxa associated with more than 3 exacerbations in the previous year (red
511 bars), and those associated with fewer exacerbations (blue bars). Only those taxa with a mean decrease
512 in Gini greater than 1·5 are shown. N=281 for all analyses.

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

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