

RESEARCH

Open Access



# Prospective association between the gut microbiota and incident pneumonia: a cohort study of 6419 individuals

Irina Wikki<sup>1,2\*</sup>, Joonatan Palmu<sup>1</sup>, Anni Kauko<sup>1,3</sup>, Aki Havulinna<sup>4</sup>, Pekka Jousilahti<sup>5,6</sup>, Leo Lahti<sup>4</sup>, Rob Knight<sup>7,8,9,10,11</sup>, Veikko Salomaa<sup>1,5</sup> and Teemu Niiranen<sup>1,5,12</sup>

## Abstract

**Background** Previous animal studies have identified the protective capacity of the gut microbiota against respiratory infections. Nevertheless, the prospective association between human gut microbiota and pneumonia risk remains unknown.

**Objectives** To evaluate the links between gut microbiota and incident pneumonia in a representative population sample.

**Methods** We performed shotgun metagenome sequencing on stool samples from 6419 FINRISK 2002 participants. Participants were followed up for incident pneumonia using nationwide health register data. We employed multivariable-adjusted Cox regression models and permutational multivariate analysis of variance (PERMANOVA) to assess the association of gut microbiome alpha diversity, compositional variation (beta diversity), and taxonomic composition with pneumonia risk.

**Results** Altogether, 685 patients (10.7%) developed pneumonia during a mean follow-up of 17.8 years. Alpha diversity was not associated with incident pneumonia (hazard ratio [HR] 1.00; 95% confidence interval [CI] 0.93 – 1.08), whereas community composition was (PERMANOVA  $R^2 = 0.03\%$ ;  $P = 0.02$ ). We observed an inverse association between the relative abundance of butyrate-producing bacteria and incident pneumonia (HR per 1-SD increase 0.91; 95% CI 0.85–0.98). The relative abundance of *Bacteroides\_F pectinophilus*, *Eubacterium\_G ventriosum*, *Agathobaculum butyriciproducens*, *Butyribacter intestini*, *Eubacterium\_I ramulus*, *CAG-1427 sp000435675*, and *CAG-603 sp900066105* were inversely associated with pneumonia risk. The relative abundance of *Clostridium\_AQ innocuum* was positively correlated with pneumonia risk.

**Conclusions** The gut microbiota composition, and especially the relative abundance of butyrate-producing bacteria, was associated with lower pneumonia risk in the population. These findings warrant further studies to investigate whether microbiome modulation to increase short chain fatty acid production through diet, prebiotics, or probiotics could reduce pneumonia risk.

**Keywords** Pneumonia, Butyrate, Incidence, Prospective studies, Microbiota-host interactions

\*Correspondence:

Irina Wikki  
irina.wikki@varha.fi

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Introduction

Growing evidence highlights the role of the gut microbiota in maintaining health [1, 2]. Past research has confirmed profound connections between the gut microbiota and distant organs, including connections to the brain, the lungs, the bones, and the liver [3]. The pathophysiology of many diseases is influenced by a complex interplay of the gut microbiota, its metabolites, and the host immune system [4]. Research indicates that the gut-lung axis — a communication pathway between the gut and the lung — is essential for maintaining respiratory health [5]. Disturbances in this axis, influenced by changes in the intestinal microbiota, may facilitate the initiation and progression of pneumonia through immune dysregulation, inflammation, and altered respiratory microbiome [6, 7].

The limited research on the links between the gut microbiota and pneumonia has focused on gut microbiota changes during infection, the association between the gut microbiota and immune responses, and the effect of antibiotics on host defense and disease severity [7–9]. Evidence from these studies suggests that commensal gut microbiota might contribute to the pathogenesis and prognosis of pneumonia. Our recent large, multinational, prospective cohort study demonstrated that the gut microbiota composition, specifically colonization with butyrate-producing bacteria, is related to protection against severe infections [10]. Apart from this single study, prior research has been limited to cross-sectional studies rather than prospective studies that would provide a higher level of evidence. A prospective setting would grant clarity of temporal sequence (did the exposure precede the outcome?) as the findings would not be biased by the disease or its therapy. Furthermore, many of the suggested hypotheses on links between the gut microbiota and pneumonia are based on animal models, and the generalizability of these results to humans remains open due to the considerable differences in animal and human microbiota [11–13].

The main goal of our study was to assess the long-term prospective associations between gut microbiota composition and incident pneumonia, with a special focus on butyrate-producing bacteria. To achieve this goal, we used shotgun metagenomic sequencing to characterize the composition of the gut microbiota in 6419 participants of the FINRISK population survey in 2002. Using nationwide health register data, we then followed these individuals for up to 18 years for hospital-diagnosed incident pneumonia to elucidate the links between the gut microbiota and pneumonia risk.

## Methods

### Study sample

The Finnish Institute for Health and Welfare has conducted FINRISK population surveys at 5-year intervals since 1972 to monitor risk factors for common diseases in the Finnish adult population [14]. For the FINRISK 2002 survey, a stratified random population sample of 13,437 individuals, aged 25–74 years from 6 geographic regions, was drawn from the National Population Information System [14]. The sampling was stratified by sex, region, and 10-year age group, according to the WHO MONICA protocol [15]. Willing participants were asked to provide a stool sample. In total, 8799 individuals participated in the baseline examination and metagenomic sequencing was successfully performed on 7231 stool samples. Participants with incomplete data, with fewer than 50,000 sample reads, or who were pregnant at baseline were excluded from the study. The final study sample consisted of 6419 individuals.

### Stool sampling and metagenomic sequencing

The baseline examination protocol and stool sampling procedure have been detailed in previous publications and summarized in the Online Data Supplement [14, 16]. Metagenomic sequencing was performed at the University of California, San Diego, California. The metagenomic sequencing and stool processing procedures have been described previously [16, 17]. The original sequence data were classified and assigned into microbial taxa using the GreenGenes2 reference database [18].

### Outcome and covariate definitions

Participant data were linked to nationwide health registers for prevalent diseases and incident pneumonia (Table E1). The main outcome was a diagnosis for pneumonia during the follow-up period. This incident event was defined as ICD-10 codes J10.0, J11.0, J12–J16, J17.0, J17.1–8, J18, B01.2, B06.8, and B25.0 recorded in the Hospital Discharge or Causes of Death registers. Diagnoses from both inpatient and outpatient hospital visits are included in the Hospital Discharge Register.

Documented covariates included comorbidities, age, sex, body mass index (BMI), smoking status, alcohol consumption, physical activity, and prior antibiotic use (4 months prior to sample collection). The definitions of these covariates are reported in detail in the Online Data Supplement.

As a sensitivity analysis, we included a Healthy Food Choices score as a covariate in the model to account for the confounding effects of diet [19]. These dietary data were available for 5906 individuals with 591 pneumonia cases.

### Statistical analysis

We used unadjusted and multivariable-adjusted Cox regression models to assess associations between outcomes and continuous microbiota features (alpha diversity, principal components, relative abundance of butyrate-producing and other bacteria, and microbial risk score). We also performed a permutational multivariate analysis (Bray–Curtis dissimilarity at species level, vegan R package, adonis2 function, 999 permutations) to assess the association between community composition and incident pneumonia. We adjusted all multivariable models for age, sex, BMI, alcohol use, smoking, physical activity, prior antibiotic exposure, and comorbidities (diabetes, cardiovascular diseases, cancer, hypertension, pulmonary diseases, and gastrointestinal diseases). We used Kaplan–Meier curves to visualize the temporal associations between microbiota features and incident pneumonia. We defined statistical significance as  $P < 0.05$ . We applied false discovery rate (FDR) corrected  $P$  values (Benjamini–Hochberg method) for genus- and species-level analysis. All statistical analyses were performed with R version 4.3.1. The statistical analyses are described in more detail in the Online Data Supplement.

**Table 1** Characteristics of the study sample

Characteristic	All (n = 6419)	Incident pneumonia	
		Yes (n = 685)	No (n = 5734)
Age, median (IQR), y	50.1 (39.0, 59.0)	60.3 (51.1, 67.2)	48.8 (38.1, 57.8)
Follow-up time, median (IQR), y	17.8 (17.7, 17.9)	11.5 (7.7, 14.9)	17.8 (17.8, 17.9)
Men, n (%)	2947 (45.9)	396 (57.8)	2551 (44.5)
BMI, median (IQR), kg/m <sup>2</sup>	26.4 (23.7, 29.4)	27.5 (24.4, 30.5)	26.3 (23.6, 29.2)
Smoking status, n (%)			
Non-smoker	3434 (53.5)	286 (41.8)	3148 (54.9)
Current smoker	1525 (23.8)	183 (26.7)	1342 (23.4)
Ex-smoker	1460 (22.7)	216 (31.5)	1244 (21.7)
Median alcohol usage (IQR), g	36.0 (9.0, 105.9)	30.6 (2.7, 97.5)	38.7 (9.0, 107.6)
Recent antibiotic exposure, n (%)	826 (12.9)	108 (15.8)	718 (12.5)
Physical activity intensity, n (%)			
Light	1353 (21.1)	167 (24.4)	1186 (20.7)
Moderate	3595 (56.0)	410 (59.9)	3185 (55.5)
High	1471 (22.9)	108 (15.8)	1363 (23.8)
Comorbidities, n (%)			
Hypertension	3073 (47.9)	442 (64.5)	2631 (45.9)
Diabetes	352 (5.5)	83 (12.1)	269 (4.7)
CVD	235 (3.7)	65 (9.5)	170 (3.0)
Pulmonary disease	539 (8.4)	114 (16.6)	425 (7.4)
Gastrointestinal disease	102 (1.6)	18 (2.6)	84 (1.5)
Cancer	238 (3.6)	59 (8.4)	179 (3.1)

Abbreviations: IQR Interquartile range, SD standard deviation, BMI body mass index, CVD cardiovascular disease

Analysis source codes are openly available at DOI:<https://doi.org/10.5281/zenodo.13880009>.

### Results

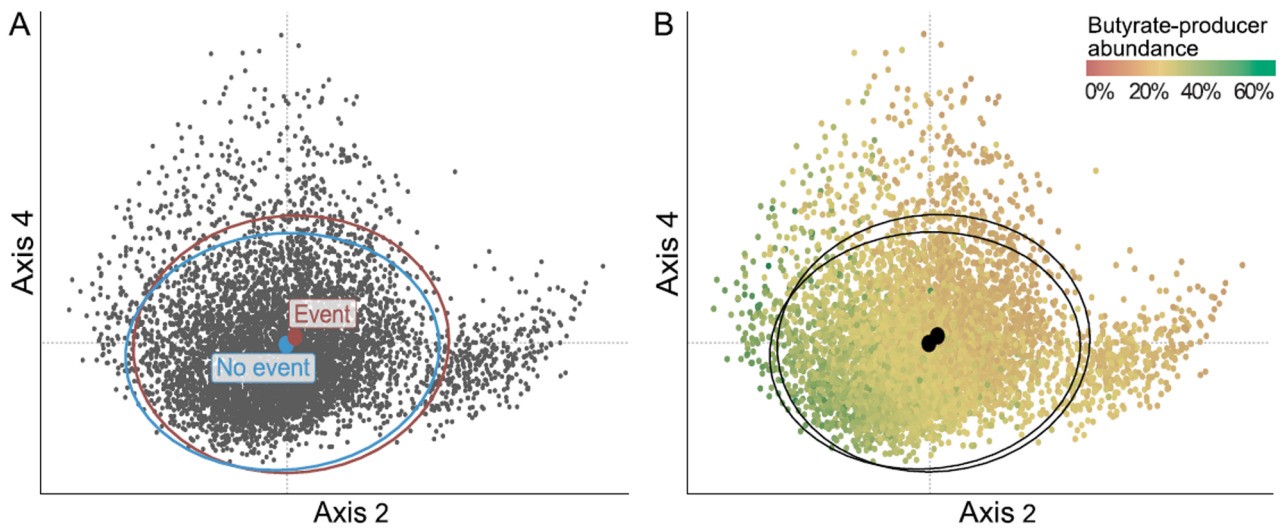
The mean follow-up time was 17.8 years, and 685 participants were diagnosed with incident pneumonia. The detailed characteristics of the study sample are reported in Table 1.

The analysis of gut microbiota composition identified 5964 species within 3352 genera. The common taxa that were included in the taxa-level analyses consisted of 384 species and 253 genera.

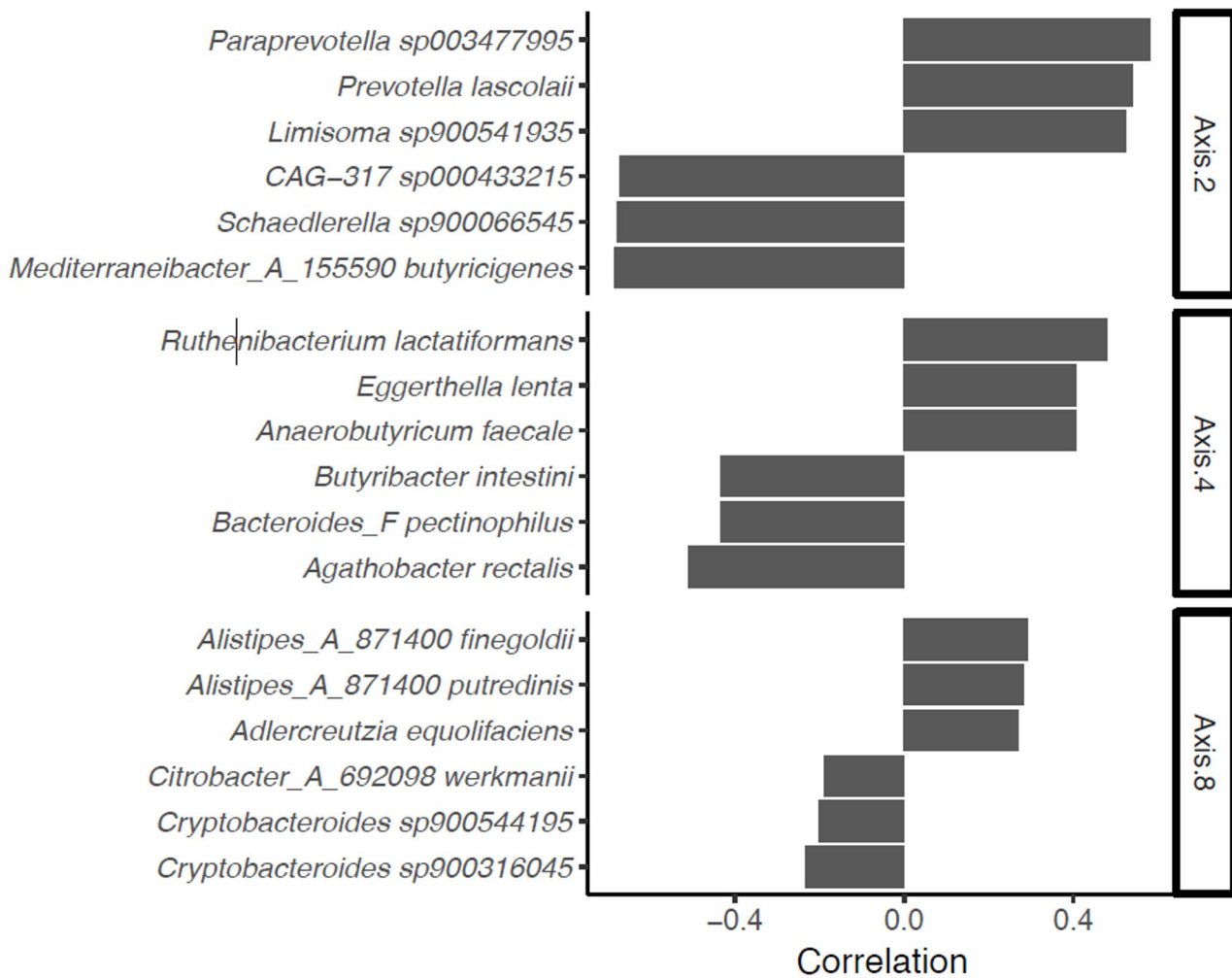
We first studied the association between alpha diversity (continuous Shannon index) and incident pneumonia using Cox proportional hazards models. Alpha diversity was not significantly associated with incident pneumonia in the unadjusted (hazard ratio [HR] 0.98; 95% confidence interval [CI] 0.91–1.06;  $P = 0.68$ ) or multivariable-adjusted (HR 1.00; 95% CI 0.93–1.08;  $P = 0.95$ ) models.

We then examined the association between gut microbiota community composition and incident pneumonia using the Bray–Curtis dissimilarity indexes calculated at species level, and the first 10 principal coordinates analysis (PCoA) axes. In unadjusted analyses, PCoA axes 8, 4, and 2 were associated with incident pneumonia ( $P \leq 0.01$  for all; Fig. 1). Figure 2 shows the bacterial species that were the strongest correlates of these axes. After multivariable-adjustment, axis 4 remained significant (HR 1.14 per 1-SD increment; 95% CI 1.06–1.22;  $P = 0.005$ ). We also performed permutational multivariate analysis of variance (PERMANOVA) to analyze the association of community composition and incident pneumonia. Both unadjusted ( $R^2 = 0.05\%$ ;  $P = 0.001$ ) and multivariable-adjusted ( $R^2 = 0.03\%$ ;  $P = 0.02$ ) analyses demonstrated a correlation between community composition and pneumonia risk.

We employed Cox models to further examine bacterial community composition and identify signature taxa linked to pneumonia risk. We observed 63 common genera significantly correlated with incident pneumonia in the unadjusted genus-level analyses (FDR  $< 0.05$ , Table E3). In the multivariable-adjusted models, *Bacteroides F*, *Eubacterium G*, *CAG-603*, *Clostridium AQ* and *Slackia A* remained significantly associated with pneumonia (Table 2). In unadjusted species-level analyses, we observed 100 common species with a significant association with incident pneumonia (FDR  $< 0.05$ , Table E4). Eight species were significantly associated in multivariable analysis: *Bacteroides F pectinophilus*, *Eubacterium G ventriosum*, *Agathobaculum butyriciproducens*, *Butyribacter intestini*, *Eubacterium I ramulus*, *CAG-1427 sp000435675*, *CAG-603 sp900066105*, and *Clostridium AQ innocuum* (Table 2). When the Healthy



**Fig. 1** Beta diversity by pneumonia incidence (A) and by the abundance of butyrate-producing bacteria (B). A PCoA 2 (horizontal axis) and PCoA 4 (vertical axis) are presented in the graph as the most significant axes in multivariable-adjusted beta diversity analysis. B Each dot represents one participant, and the color gradient visualizes the total relative butyrate producer abundance of each sample. Circles showing the centroids of our two main groups (event/no event). PCoA = Principal coordinate axis



**Fig. 2** Bacterial species most strongly correlated with PCoA axes 2, 4, and 8. Correlations are reported as Spearman correlations

**Table 2** Gut microbial taxa significantly associated with incident pneumonia using Cox proportional hazards model

Taxa	HR	95% CI	P	FDR
<b>Genus</b>				
<i>Eubacterium_G</i>	0.86	0.80–0.93	$7.7 \times 10^{-5}$	0.02
<i>Bacteroides_F</i>	0.87	0.81–0.94	$7.9 \times 10^{-4}$	0.02
<i>CAG_603</i>	0.89	0.83–0.95	$8.6 \times 10^{-4}$	0.04
<i>Clostridium_AQ</i>	1.13	1.06–1.22	$3.8 \times 10^{-4}$	0.03
<i>Slackia_A</i>	1.14	1.06–1.23	$7.9 \times 10^{-4}$	0.04
<b>Species</b>				
<i>Eubacterium_G ventriosum</i>	0.84	0.78–0.91	$7.1 \times 10^{-6}$	0.003
<i>Agathobaculum butyriciproducens</i>	0.87	0.80–0.93	$1.7 \times 10^{-4}$	0.03
<i>CAG-1427 sp000435675</i>	0.87	0.81–0.94	$7.6 \times 10^{-4}$	0.04
<i>Bacteroides_F pectinophilus</i>	0.87	0.81–0.94	$3.7 \times 10^{-4}$	0.04
<i>Butyribacter intestini</i>	0.88	0.81–0.94	$5.2 \times 10^{-4}$	0.04
<i>Eubacterium_I ramulus</i>	0.88	0.81–0.95	$9.9 \times 10^{-4}$	0.05
<i>CAG-603 sp900066105</i>	0.89	0.82–0.95	$7.4 \times 10^{-4}$	0.04
<i>Clostridium_AQ innocuum</i>	1.13	1.06–1.21	$5.1 \times 10^{-4}$	0.04

Hazard ratios are reported per 1 SD increment

Abbreviations: HR Hazard ratio, CI Confident interval, FDR False discovery rate corrected P

Food Choices score was included among the covariates, only *Eubacterium\_G ventriosum* (HR 0.85; 95% CI 0.79–0.92;  $P=0.03$ ) remained significantly associated with pneumonia in the species-level analysis, while the genus-level analysis did not yield any significant results (Table E5). As the point estimates remained relatively similar compared to the main results after adjusting for diet, this decrease in the number of significant taxa was mainly a result of decreased statistical power arising from lower number of participants with dietary data available.

**Table 3** The association of butyrate-producing bacteria abundance with pneumonia incidence

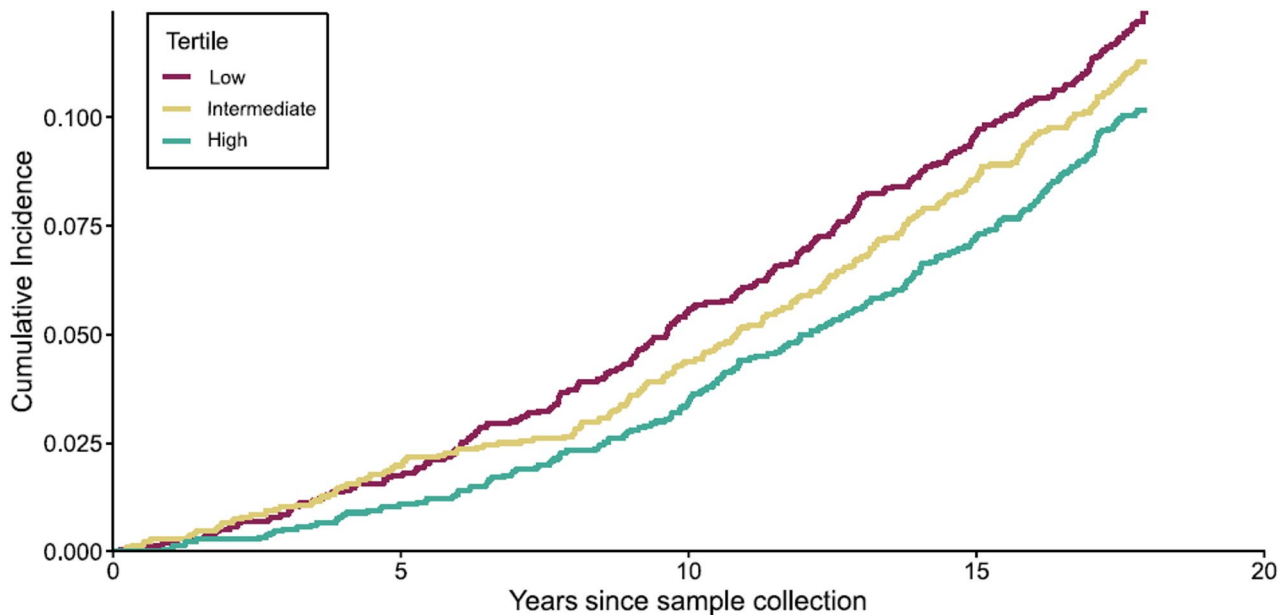
Butyrate producer abundance	Unadjusted		Adjusted	
	HR (95% CI)	P	HR (95% CI)	P
Continuous	0.90 (0.84–0.97)	0.005	0.91 (0.85–0.98)	0.01
<b>Tertiles</b>				
Low	1.00 (ref)		1.00 (ref)	
Intermediate	0.91 (0.76–1.09)	0.30	0.93 (0.77–1.11)	0.41
High	0.81 (0.67–0.97)	0.02	0.79 (0.66–0.95)	0.01

Hazard ratios are reported per 1 SD increment

Abbreviations: HR Hazard ratio, CI Confident interval, FDR False discovery rate corrected P

Our next objective was to determine whether the relative abundance of butyrate-producing bacteria is associated with pneumonia risk. We observed an inverse association between butyrate-producing bacteria relative abundance and pneumonia incidence in unadjusted analyses (HR per 1-SD increase 0.90; 95% CI 0.84–0.97;  $P=0.005$ ; Fig. 3; Table 3). This significant association was robust to adjustment for covariates (HR per 1-SD increase 0.91; 95% CI 0.85–0.98;  $P=0.01$ ).

Finally, we created a bacterial risk score for incident pneumonia using elastic nets for Cox models (see Online Data Supplement). A detailed portrayal of individual species-level taxa contribution to the risk score is reported in Figure E1 and Table E6. The risk score was calculated for each individual, and the cohort was divided



**Fig. 3** The association of butyrate-producing bacteria abundance with pneumonia incidence. Participants were divided into subgroups based on the abundance of butyrate producers. The plot shows the cumulative pneumonia incidence in subgroups of low, intermediate, and high butyrate abundance

into tertiles based on the risk score. As seen in Figure E2, the multivariable-adjusted incidence of pneumonia was higher in the high-risk score group (HR 1.31; 95% CI 1.08–1.59;  $P=0.007$ ) when compared to the low-risk group. A 1-SD increment in the risk score resulted in a 1.07-fold increase (95% CI 1.02–1.13;  $P=0.008$ ) in pneumonia risk. As a sensitivity measure, we successfully validated the risk score by dividing the cohort population into two equal-sized subpopulations (see Online Data Supplement).

## Discussion

In this prospective cohort study, we investigated for the first time the association between gut microbiota composition and long-term pneumonia incidence, using secondary health care diagnoses from a representative population sample. We demonstrated that specific axes of microbiota community composition were associated with pneumonia risk. In addition, a higher relative abundance of butyrate-producing bacteria in the gut microbiota was associated with a lower risk of incident pneumonia.

We found no correlation between alpha diversity of gut microbiota and pneumonia risk. This is in contrast to prior cross-sectional studies, which have mainly reported a loss of alpha diversity in infected patients when compared to healthy controls [20]. One of these studies suggested that the extent of alpha diversity loss may also be related to the severity of the disease [21]. However, evidence from cross-sectional and prospective studies are not directly comparable because the infection and its treatment may alter the composition of the gut microbiota. Our findings are corroborated by some of the earlier literature. Orioux et al. demonstrated that the difference between critically ill patients who later developed pneumonia and those who did not develop pneumonia was linked to differences in the relative abundances of gut microbial taxa rather than alpha diversity [22]. Li et al. also presented similar findings in a study on stroke-associated pneumonia [23]. Interestingly, Maji et al. reported an increased alpha diversity and elevated abundance of bacteria that produce butyrate and other short-chain fatty acids (SCFAs) in patients with active pulmonary tuberculosis infection [24]. These prior findings, along with our results, may suggest that the functional capacity of the gut microbiota may be more important than bacterial diversity in maintaining intestinal balance.

Data on the causal relationship between gut microbiota and lung infections are scarce and mainly based on experimental animal studies. In a small prospective study of critically ill patients, dysbiotic gut microbiota were found to be associated with increased risk of infection, accompanied by impaired immune system function [25]. Most evidence focuses on the behavior of gut microbiota during infection. Mouse studies have shown

notable shifts in intestinal microbial communities during bacterial and viral lung infections, including loss of alpha diversity and altered community composition [26, 27]. Additional data from mice suggest that the gut microbial community boosts the immune system and supports recovery from bacterial pneumonia [9, 26]. Some studies indicate that the observed changes in gut microbiota composition during infection result from the infection rather than influence susceptibility to it. Nevertheless, a growing body of research from various areas, such as lung diseases, suggests that changes in the gut microbiota may precede disease manifestation [28, 29].

Differential taxa analysis showed a positive association of two common genera and one common species with incident pneumonia. *Clostridium\_AQ\_innocuum* has been identified as a common inhabitant of the human gut and has been linked to gastrointestinal diseases [30]. It has also been identified as a rare cause of bacteremia and endocarditis. A recent study of this bacterial species suggested it may be less benign in the gut environment than previously considered [31]. The other observed genus with positive association with pneumonia, *Slackia\_A*, has previously been positively linked to various disease states such as pulmonary artery hypertension, type 2 diabetes, and colorectal cancer [32, 33]. These previous findings provide some support for our results, although the mechanism of how colonization with these bacteria may increase the risk of pneumonia remains unclear.

We observed seven species that were inversely associated with incident pneumonia. *Agathobaculum\_butyriciproducens* has recently been identified as a strictly anaerobic bacteria that produces butyric acid [34]. *Eubacterium\_I\_ramulus* has been identified as a gut commensal metabolizing dietary flavonoids that may mediate antioxidative, anti-inflammatory and antimicrobial effects [35–37]. *CAG-1427\_sp000435675* has not, to our knowledge, been studied, but belongs to the phylum *Actinobacteriota* that is referred to as “a relevant minority for the maintenance of gut homeostasis” [38]. *Bacteroides\_F\_pectinophilus*, *Butyribacter\_intestini*, *CAG-603\_sp900066105*, and *Eubacterium\_G\_ventriosum* belong to the family *Lachnospiraceae*, which are commensal anaerobic bacteria and are among the main SCFA producers in the gut. The *Lachnospiraceae* family of bacteria have been identified as enhancing the immunological function in the gut, although their impact on the host physiology has been inconsistent across different studies [39]. Many observational studies have reported a depletion of *Lachnospiraceae* family of bacteria during respiratory infections [20]. *Eubacterium\_G\_ventriosum* is recognized as one of the key butyrate-producing bacteria in the gut (Table E2). A small observational study has reported a significant depletion of *Eubacterium\_ventriosum* in both COVID-19 patients and COVID-negative

pneumonia patients, when compared to healthy controls [40]. In COVID-19 pneumonia patients, a significant inverse correlation between SCFA-producing bacteria and inflammatory markers was observed [41]. Studies on kidney transplant recipients and patients post-allogeneic HSCT have reported that reduced SCFA-producing bacteria are correlated with increased risk and incidence of viral respiratory tract infections [42, 43]. The results of our species-level analyses and the results from previous studies therefore highlight the importance of butyrate and other SCFA-producing bacteria as a potential modifier of infection risk.

Our results on butyrate-producing bacteria are in agreement with outcomes from previous cross-sectional studies. Abundant evidence from earlier studies has demonstrated the immunomodulatory and health-promoting potential of gut-derived SCFAs, especially butyrate, acetate and propionate. The proposed mechanisms underlying these findings include direct and indirect bi-directional signaling pathways between the gut microbiota and distant organs. Intraluminal production of butyrate and other SCFAs is known to contribute to the stability of the intestinal microenvironment by supporting the growth of commensal bacteria and the integrity of the intestinal barrier [44, 45]. Butyrate and other SCFAs are also shown to possess antimicrobial properties that inhibit the growth of various Gram-positive and Gram-negative bacteria [45, 46], including many opportunistic pathogens. Evidence from in vitro and in vivo mouse studies suggests that butyrate enhances the antibacterial clearance of macrophages, providing protection both locally in the intestine and distant organs such as the lungs [47]. Acetate has also been shown to increase the bactericidal activity of macrophages during pneumonia [48]. Furthermore, butyrate and other SCFAs stimulate and support the development of helper T cells, which play a role in immune regulation and neutrophil recruitment within the lungs [49, 50]. Another study demonstrated that SCFAs increase the number of activated macrophages and reduce neutrophil-attracting chemokine CXCL1 production in the airways, modulating neutrophil recruitment and thus reducing tissue damage [51]. Moreover, butyrate has been demonstrated to modulate dendritic cell production of inflammatory cytokines, which is significant particularly in the later stages of infection [47, 52]. The depletion of SCFA-producing bacteria has been linked to impaired pulmonary defense against bacterial pathogens [53], and especially acetate supplementation has resulted in reduced bacterial loads in the lungs [48]. Several other studies have also confirmed a positive association between gut SCFA production and its immunomodulatory capacity [54, 55]. These previous findings deepen our understanding of the

mechanisms underlying the potential protective capacity of SCFA-producing bacteria against pneumonia.

The taxa used for the calculation of the abundance of butyrate-producing bacteria (Table E2) do not only produce butyrate, but many are also producers of other SCFAs, including propionate and acetate. Commensal gut bacteria present cross-feeding properties, where the diversity and quantity of SCFA production varies according to available substrates and present species [55]. As both acetate and propionate may enter the systemic circulation in greater amounts than butyrate, it may be argued that they have a more potent systemic effect than butyrate. Thus, all major SCFAs, including butyrate, may have a role in pneumonia defense. It should also be noted that we did not perform direct measurement of butyrate or other bacterial metabolites. Therefore, our conclusions are somewhat speculative and based on previous literature and on the assumption that the abundance of butyrate-producing bacteria reflects the intraluminal abundance of butyrate. Further research with direct measurement of intestinal butyrate levels is necessary to validate our findings.

We identified a selection of bacterial taxa associated with pneumonia incidence, and developed a cumulative risk score for an overall risk assessment. Modulation of the gut microbiota has been shown to be effective in treatment recurrent *Clostridioides difficile* infections and is proposed as a potential treatment for many other dysbiosis-associated diseases [56]. By analyzing the composition of gut microbiota, it could be possible to find patients at high risk of future infections. The abundance of SCFA-producing bacteria could potentially be considered as a marker for infection risk given the aforementioned and other established benefits of butyrate and other SCFAs [52, 57]. Our risk score presents one useful way of examining the composition of the gut microbiota and identifying specific bacterial taxa that are positively or negatively associated with the risk of infection. Such a risk score might also be used to identify individuals who would benefit from prophylactic gut microbiota modulation. However, further experimental research is needed to confirm the benefit and safety of such a prophylactic approach. Furthermore, it remains unclear to what extent our risk score is applicable at a general level, as it has not been externally validated.

The strengths of our study include a representative population sample with comprehensive nationwide register-based follow-up with virtually no loss to follow-up. The primary strength is that this is the first report on the long-term prospective association between gut microbiota and incident pneumonia, thus avoiding many biases inherent in cross-sectional studies [58]. However, our study focused on more severe forms of pneumonia that required hospital evaluation or hospitalization. Neither

was there detailed information on the type of pneumonia (i.e., viral vs. bacterial infection, or community- vs. hospital-acquired). This limitation may confound the observed associations, as biological mechanisms for and susceptibility to different pneumonia types vary. Although the Finnish health registers are generally considered highly reliable for research use [59], errors or omissions in the coding process may still occur. Additionally, despite the relative stability of the intestinal core microbiota over time [60], a number of variables could alter it between the time of sampling and the endpoint event. These variables include exposure to long- or short-term medications, such as antibiotic treatments, dietary changes, alterations in environmental factors, and prevalent disease status. In our study, gut microbiota was assessed only at baseline, which may lead to underestimation of the association between the microbiota and pneumonia risk. Repeated stool sampling could provide a more dynamic assessment of an individual's gut microbiota. As with all observational studies, the associations do not prove causality even in a prospective setting. The presence of unknown or unmeasured confounding factors cannot be excluded and therefore residual confounding may persist.

This research represents the first steps toward a more profound understanding of the links between gut microbiota and pneumonia risk. In short, we demonstrated that the abundance of butyrate-producing bacteria is associated with a lower long-term risk of pneumonia. We identified seven species associated with reduced pneumonia incidence, many of them linked to SCFA production in the gut, and detected a signature of gut microbiota predicting the risk of incident pneumonia. Future implications include the potential for assessing individual infection risk and developing novel treatments for patients in the high-risk group. Randomized clinical trials in patients at high risk of pneumonia that focus on increasing gut SCFA levels are now needed to confirm causality.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12931-025-03453-w>.

Supplementary Material 1.

### Acknowledgements

We thank the participants of the FINRISK 2002 study for their contributions to this study.

### Authors' contributions

IW, JP, and TN conceived and designed the study. VS, AH, JP, PJ, and RK were involved in data collection and microbiota characterization. IW, JP, AK, LL, AH, PJ, and TN analyzed and interpreted the data. IW and TN wrote the manuscript. JP, LL, AK, AH, PJ, VS, and RK reviewed the manuscript critically for important intellectual content.

### Funding

IW, AK, and PJ have nothing to disclose. JP was supported by the Paavo Nurmi Foundation and the Finnish Medical Foundation. AH was supported by the Research Council of Finland (grant 369585). VS was supported by the Juho Vainio Foundation. LL was supported by the Research Council of Finland (grant 330887). TN was supported by the Research Council of Finland (grants 321351 and 354447), the Sigrid Jusélius Foundation, the Finnish Foundation for Cardiovascular Research, and the wellbeing services county of Southwest Finland.

### Data availability

The metagenomic data are available from the European Genome-Phenome Archive, (<https://ega-archive.org/studies/EGAS00001005020>) (accession number EGAD00001007035). The phenotype data contain sensitive information from healthcare registers and they are available through the THL biobank upon submission of a research plan and signing a data transfer agreement (<https://thl.fi/en/research-and-development/thl-biobank/for-researchers/application-process>).

### Declarations

#### Ethics approval and consent to participate

The FINRISK 2002 study protocol was approved by the Coordinating Ethics Committee of the Helsinki and Uusimaa University Hospital District (Ref. 558/E3/2001). All participants signed informed consent. The study was conducted according to the guidelines laid down in the Declaration of Helsinki [61].

#### Consent for publication

Not applicable.

#### Competing interests

RK is an advisor to DayTwo, Ltd.; an advisor to and equity holder of GenCirq, Inc., BiomeSense, Inc., and Cybele Microbiome, Inc.; co-founder and equity holder of Biota Technology, Inc.; and a co-founder, advisor, and equity holder of Micronoma, Inc. RK is also the director of UCSD Center for Microbiome Innovation. TN has received consulting and authoring fees from AstraZeneca, Sevier Finland and Orion Corporation.

#### Author details

<sup>1</sup>Department of Internal Medicine, University of Turku, P.O. Box 52, Turku FI-20521, Finland

<sup>2</sup>Division of Gastroenterology, Turku University Hospital, Turku, Finland

<sup>3</sup>InFLAMES Research Flagship Center, University of Turku, Turku, Finland

<sup>4</sup>Department of Computing, University of Turku, Turku, Finland

<sup>5</sup>Department of Public Health, Finnish Institute for Health and Welfare, Helsinki, Finland

<sup>6</sup>Department of Public Health, Faculty of Medicine, University of Helsinki, Helsinki, Finland

<sup>7</sup>Department of Pediatrics, University of California San Diego, La Jolla, San Diego, CA, USA

<sup>8</sup>Department of Computer Science and Engineering, University of California, La Jolla, San Diego, CA, USA

<sup>9</sup>Department of Bioengineering, University of California, La Jolla, San Diego, CA, USA

<sup>10</sup>Center for Microbiome Innovation, University of California, La Jolla, San Diego, CA, USA

<sup>11</sup>Halicioğlu Data Science Institute, University of California San Diego, La Jolla, San Diego, CA, USA

<sup>12</sup>Division of Internal Medicine, Turku University Hospital, Turku, Finland

Received: 9 August 2025 / Accepted: 10 December 2025

Published online: 22 December 2025

### References

1. Levy M, Blacher E, Elinav E. Microbiome, metabolites and host immunity. *Curr Opin Microbiol.* 2017;35:8–15.
2. Joos R, Boucher K, Lavelle A, Arumugam M, Blaser MJ, Claesson MJ, et al. Examining the healthy human Microbiome concept. *Nat Rev Microbiol.* 2025;23(3):192–205.

3. Ahlwat S, Yadav A, Sharma K. Gut–organ axis: A microbial outreach and networking. *Lett Appl Microbiol*. 2021;72(6):636–68.
4. Carding S, Verbeke K, Vipond DT, Corfe BM, Owen LJ. Dysbiosis of the gut microbiota in disease. *Microb Ecol Health Dis*. 2015;26:26191.
5. Budden KF, Gellatly SL, Wood DLA, Cooper MA, Morrison M, Hugenholtz P, et al. Emerging pathogenic links between microbiota and the gut-lung axis. *Nat Rev Microbiol*. 2017;15:55–63.
6. Wypych TP, Wickramasinghe LC, Marsland BJ. The influence of the microbiome on respiratory health. *Nat Immunol*. 2019;20:1279–90.
7. Marrella V, Nicchiotti F, Cassani B. Microbiota and immunity during respiratory infections: lung and gut affair. *Int J Mol Sci*. 2024;25:4051.
8. Ichinohe T, Pang IK, Kumamoto Y, Peaper DR, Ho JH, Murray TS, et al. Microbiota regulates immune defense against respiratory tract influenza A virus infection. *Proc Natl Acad Sci USA*. 2011;108:5354–9.
9. Schuijt TJ, Lankelma JM, Scicluna BP, De Sousa E, Melo F, Roelofs JJTH, et al. The gut microbiota plays a protective role in the host defence against pneumococcal pneumonia. *Gut*. 2016;65:575–83.
10. Kullberg RJ, Wikki I, Haak BW, Kauko A, Galenkamp H, Peters-Sengers H, et al. Association between butyrate-producing gut bacteria and the risk of infectious disease hospitalisation: results from two observational, population-based microbiome studies. *Lancet Microbe*. 2024;5:100864.
11. Cahana I, Iraqi FA. Impact of host genetics on gut microbiome: take-home lessons from human and mouse studies. *Anim Models Exp Med*. 2020;3:229–36.
12. Hugenholtz F, de Vos WM. Mouse models for human intestinal microbiota research: a critical evaluation. *Cell Mol Life Sci*. 2018;75:149–60.
13. Nguyen TLA, Vieira-Silva S, Liston A, Raes J. How informative is the mouse for human gut microbiota research? *Dis Model Mech*. 2015;8:1–16.
14. Borodulin K, Tolonen H, Jousilahti P, Jula A, Juolevi A, Koskinen S, et al. Cohort profile: the National FINRISK Study. *Int J Epidemiol*. 2018;47:696–696i.
15. The World Health Organization MONICA Project. (monitoring trends and determinants in cardiovascular disease): a major international collaboration. WHO MONICA project principal investigators. *J Clin Epidemiol*. 1988;41:105–14.
16. Salosensaari A, Laitinen V, Havulinna AS, Meric G, Cheng S, Perola M, et al. Taxonomic signatures of cause-specific mortality risk in human gut microbiome. *Nat Commun*. 2021;12:2671.
17. Qin Y, Havulinna AS, Liu Y, Jousilahti P, Ritchie SC, Tokolyi A, et al. Combined effects of host genetics and diet on human gut microbiota and incident disease in a single population cohort. *Nat Genet*. 2022;54:134–42.
18. McDonald D, Jiang Y, Balaban M, Cantrell K, Zhu Q, Gonzalez A, et al. GreenGenes2 unifies microbial data in a single reference tree. *Nat Biotechnol*. 2024;42:715–8.
19. Koponen KK, Salosensaari A, Ruuskanen MO, Havulinna AS, Männistö S, Jousilahti P, et al. Associations of healthy food choices with gut microbiota profiles. *Am J Clin Nutr*. 2021;114:605.
20. Woodall CA, McGeoch LJ, Hay AD, Hammond A. Respiratory tract infections and gut microbiome modifications: a systematic review. *PLoS ONE*. 2022;17(1):e0262057.
21. Yeoh YK, Zuo T, Lui GCY, Zhang F, Liu Q, Li AYL, et al. Gut microbiota composition reflects disease severity and dysfunctional immune responses in patients with COVID-19. *Gut*. 2021;70:698–706.
22. Orioux A, Enaud R, Imbert S, Boyer P, Begot E, Camino A, et al. The gut microbiota composition is linked to subsequent occurrence of ventilator-associated pneumonia in critically ill patients. *Microbiol Spectr*. 2023;11:e00641–23.
23. Li Z, Gu M, Sun H, Chen X, Zhou J, Zhang Y. The potential of gut microbiota in prediction of stroke-associated pneumonia. *Brain Sci*. 2023;13:1217.
24. Maji A, Misra R, Dhakan DB, Gupta V, Mahato NK, Saxena R, et al. Gut microbiome contributes to impairment of immunity in pulmonary tuberculosis patients by alteration of butyrate and propionate producers. *Environ Microbiol*. 2018;20:402–19.
25. Schlechte J, Zucoloto AZ, Yu I, Doig CJ, Dunbar MJ, McCoy KD, et al. Dysbiosis of a microbiota-immune metasytem in critical illness is associated with nosocomial infections. *Nat Med*. 2023;29:1017–27.
26. Lankelma JM, Birnie E, Weehuizen TAF, Scicluna BP, Belzer C, Houtkooper RH, et al. The gut microbiota as a modulator of innate immunity during melioidosis. *PLoS Negl Trop Dis*. 2017;11(4):e0005548.
27. Groves HT, Cuthbertson L, James P, Moffatt MF, Cox MJ, Tregoning JS. Respiratory disease following viral lung infection alters the murine gut microbiota. *Front Immunol*. 2018;9:182.
28. Hrnčir T. Gut microbiota dysbiosis: triggers, consequences, diagnostic and therapeutic options. *Microorganisms*. 2022;10:578.
29. Liu Y, Teo SM, Meric G, Tang HHH, Zhu Q, Sanders JG, et al. The gut Microbiome is a significant risk factor for future chronic lung disease. *J Allergy Clin Immunol*. 2022;151:943–52.
30. Cherny KE, Muscat EB, Reyna ME, Kociolek LK. *Clostridium innocuum*: microbiological and clinical characteristics of a potential emerging pathogen. *Anaerobe*. 2021;71:102418.
31. Bhattacharjee D, Flores C, Woelfel-Monsivais C, Seekatz AM. Diversity and prevalence of *Clostridium innocuum* in the human gut microbiota. *mSphere*. 2022;8:e00569-22.
32. Neri-Rosario D, Martinez-Lopez YE, Esquivel-Hernandez DA, Sanchez-Castaneda JP, Padron-Manrique C, Vazquez-Jimenez A, et al. Dysbiosis signatures of gut microbiota and the progression of type 2 diabetes: a machine learning approach in a Mexican cohort. *Front Endocrinol*. 2023;14:1170459.
33. Zhang C, Xi Y, Zhang Y, He P, Su X, Li Y, et al. Causal effects between gut microbiota and pulmonary arterial hypertension: a bidirectional Mendelian randomization study. *Heart Lung*. 2024;64:189–97.
34. Ahn S, Jin T-E, Chang D-H, Rhee M-S, Kim HJ, Lee SJ, et al. *Agathobaculum butyriciproducens* gen. Nov Sp Nov., a strict anaerobic, butyrate-producing gut bacterium isolated from human faeces and reclassification of *Eubacterium desmolans* as *Agathobaculum desmolans* comb. Nov. *Int J Syst Evol Microbiol*. 2016;66:3656–61.
35. Yang G, Hong S, Yang P, Sun Y, Wang Y, Zhang P, et al. Discovery of an ene-reductase for initiating flavone and flavonol catabolism in gut bacteria. *Nat Commun*. 2021;12:790.
36. Shamsudin NF, Ahmed QU, Mahmood S, Shah SAA, Khatib A, Mukhtar S, et al. Antibacterial effects of flavonoids and their structure-activity relationship study: a comparative interpretation. *Molecules*. 2022;27:1149.
37. Fan X, Fan Z, Yang Z, Huang T, Tong Y, Yang D, et al. Flavonoids-natural gifts to promote health and longevity. *Int J Mol Sci*. 2022;23:2176.
38. Binda C, Lopetuso LR, Rizzatti G, Gibiino G, Cennamo V, Gasbarrini A. Actinobacteria: a relevant minority for the maintenance of gut homeostasis. *Dig Liver Dis*. 2018;50:421–8.
39. Vacca M, Celano G, Calabrese FM, Portincasa P, Gobbetti M, De Angelis M. The controversial role of human gut lachnospiraceae. *Microorganisms*. 2020;8:573.
40. Zuo T, Zhang F, Lui GCY, Yeoh YK, Li AYL, Zhan H, et al. Alterations in gut microbiota of patients with COVID-19 during time of hospitalization. *Gastroenterology*. 2020;159:944–55.
41. Tang L, Gu S, Gong Y, Li B, Lu H, Li Q, et al. Clinical significance of the correlation between changes in the major intestinal bacteria species and COVID-19 severity. *Engineering*. 2020;6:1178–84.
42. Haak BW, Littmann ER, Chaubard J-L, Pickard AJ, Fontana E, Adhi F, et al. Impact of gut colonization with butyrate-producing microbiota on respiratory viral infection following allo-HCT. *Blood*. 2018;131:2978–86.
43. Lee JR, Huang J, Magruder M, Zhang LT, Gong C, Sholi AN, et al. Butyrate-producing gut bacteria and viral infections in kidney transplant recipients: a pilot study. *Transpl Infect Dis*. 2019;21:e13180.
44. Zhu L-B, Zhang Y-C, Huang H-H, Lin J. Prospects for clinical applications of butyrate-producing bacteria. *World J Clin Pediatr*. 2021;10:84–92.
45. Du K, Bereswill S, Heimesaat MM. A literature survey on antimicrobial and immune-modulatory effects of butyrate revealing non-antibiotic approaches to tackle bacterial infections. *Eur J Microbiol Immunol*. 2021. <https://doi.org/10.1556/1886.2021.00001>.
46. Kadry AA, El-Antrawy MA, El-Ganiny AM. Impact of short chain fatty acids (SCFAs) on antimicrobial activity of new  $\beta$ -lactam/ $\beta$ -lactamase inhibitor combinations and on virulence of *Escherichia coli* isolates. *J Antibiot (Tokyo)*. 2023;76:225–35.
47. Schulthess J, Pandey S, Capitani M, Rue-Albrecht KC, Arnold I, Franchini F, et al. The short chain fatty acid butyrate imprints an antimicrobial program in macrophages. *Immunity*. 2019;50:432–45.
48. Sencio V, Barthelemy A, Tavares LP, Machado MG, Soulard D, Cuinat C, et al. Gut dysbiosis during influenza contributes to pulmonary pneumococcal superinfection through altered short-chain fatty acid production. *Cell Rep*. 2020;30:2934–2947.e6.
49. Remcho TP, Kolls JK. Unfolding the role of Th17 cells in neutrophilic lung inflammation. *Am J Respir Cell Mol Biol*. 2024;71:137–8.
50. Saadh MJ, Allela OQB, Ballal S, Mahdi MS, Chahar M, Verma R, et al. The effects of microbiota-derived short-chain fatty acids on T lymphocytes: from autoimmune diseases to cancer. *Semin Oncol*. 2025;52:152398.
51. Trompette A, Gollwitzer ES, Pattaroni C, Lopez-Mejia IC, Riva E, Pernot J, et al. Dietary fiber confers protection against flu by shaping Ly6c<sup>+</sup> patrolling

- monocyte hematopoiesis and CD8+T cell metabolism. *Immunity*. 2018;48:992-1005.e8.
52. Siddiqui MT, Cresci GAM. The immunomodulatory functions of butyrate. *J Inflamm Res*. 2021;14:6025–41.
  53. Dörner PJ, Anandakumar H, Röwekamp I, Fiocca Vernengo F, Millet Pascual-Leone B, Krzanowski M, et al. Clinically used broad-spectrum antibiotics compromise inflammatory monocyte-dependent antibacterial defense in the lung. *Nat Commun*. 2024;15:2788.
  54. Maslowski KM, Vieira AT, Ng A, Kranich J, Sierro F, Yu D, et al. Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature*. 2009;461:1282–6.
  55. Verma A, Bhagchandani T, Rai A, Nikita, Sardarni UK, Bhavesh NS, et al. Short-Chain fatty acid (SCFA) as a connecting link between microbiota and Gut-Lung axis: potential therapeutic intervention to improve lung health. *ACS Omega*. 2024;9:14648–71.
  56. Ortigao R, Pimentel-Nunes P, Dinis-Ribeiro M, Libanio D. Gastrointestinal microbiome - what we need to know in clinical practice. *GE Port J Gastroenterol*. 2020;27:336–51.
  57. Feng Q, Chen W-D, Wang Y-D. Gut microbiota: an integral moderator in health and disease. *Front Microbiol*. 2018;9:151.
  58. Wang X, Cheng Z. Cross-sectional studies: strengths, weaknesses, and recommendations. *Chest*. 2020;158:S65-71.
  59. Sund R. Quality of the Finnish hospital discharge register: a systematic review. *Scand J Public Health*. 2012;40:505–15.
  60. Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, et al. The long-term stability of the human gut microbiota. *Science*. 2013;341:1237439.
  61. World Medical Association. World medical association declaration of helsinki: ethical principles for medical research involving human subjects. *JAMA*. 2013;310:2191–4.

### Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.