

1 **Multiple indicators of gut dysbiosis predict all-cause and cause-specific**  
2 **mortality in solid organ transplant recipients**

3 J. Casper Swarte<sup>1,2,3</sup>, Shuyan Zhang<sup>1</sup>, Lianne M. Nieuwenhuis<sup>3,4</sup>, Ranko Gacesa<sup>1,5</sup>, Tim J.  
4 Knobbe<sup>2,3</sup>, **TransplantLines Investigators**, Vincent E. de Meijer<sup>3,4</sup>, Kevin Damman<sup>3,6</sup>, Erik  
5 A.M. Verschuuren<sup>3,6</sup>, C. Tji Gan<sup>3,6</sup>, Jingyuan Fu<sup>5,7</sup>, Alexandra Zhernakova<sup>5</sup>, Hermie J.M.  
6 Harmsen<sup>9</sup>, Hans Blokzijl<sup>1,3</sup>, Stephan J.L. Bakker<sup>2,3</sup>, Johannes R. Björk<sup>1,5#</sup> and Rinse K.  
7 Weersma<sup>1#</sup>  
8 # contributed equally to this study.

9 <sup>1</sup> Department of Gastroenterology and Hepatology, University of Groningen, University Medical  
10 Center Groningen, Groningen, the Netherlands

11 <sup>2</sup> Division of Nephrology, Department of Internal Medicine, University of Groningen,  
12 University Medical Center Groningen, Groningen, the Netherlands

13 <sup>3</sup> UMCG Comprehensive Transplant Center, Groningen, the Netherlands

14 <sup>4</sup> Department of Surgery, section of Hepatobiliary Surgery and Liver Transplantation, University of  
15 Groningen, University Medical Center Groningen, Groningen, the Netherlands

16 <sup>5</sup> Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen,  
17 the Netherlands

18 <sup>6</sup> University of Groningen, Department of Cardiology, University Medical Center Groningen,  
19 Groningen, The Netherlands.

20 <sup>7</sup> Department of Pediatrics, University of Groningen, University Medical Center Groningen, Groningen,  
21 the Netherlands.

22 <sup>8</sup> Department of Pulmonary Diseases, Tuberculosis and Lung transplantation, University of  
23 Groningen, University Medical Center Groningen, Groningen, The Netherlands

24 <sup>9</sup> Department of Medical Microbiology and Infection prevention, University of Groningen, University  
25 Medical Center Groningen, Groningen, The Netherlands

26 **Corresponding authors: Dr. J.R. Björk ([bjork.johannes@gmail.com](mailto:bjork.johannes@gmail.com)) and Prof R.K. Weersma**  
27 **([r.k.weersma@umcg.nl](mailto:r.k.weersma@umcg.nl))**

28 **Abstract**

29 **Objective**

30 Gut microbiome composition is associated with multiple diseases, but relatively little is  
31 known about its relationship with long-term outcome measures. While gut dysbiosis has  
32 been linked to mortality risk in the general population, the relation with overall survival in  
33 specific diseases has not been extensively studied. In the current study, we present in-depth  
34 analyses regarding the relationship between gut dysbiosis and all-cause and cause-specific  
35 mortality in the setting of solid organ transplant recipients (SOTR).

36 **Design** We analyzed 1,337 metagenomes derived from fecal samples of 766 kidney, 334  
37 liver, 170 lung and 67 heart transplant recipients from the TransplantLines Biobank and  
38 Cohort; a prospective cohort study including extensive phenotype data with 6.5 years of  
39 follow up. To quantify gut dysbiosis, we included additional 8,208 metagenomic samples  
40 from a general population from the same geographical location. Multivariable Cox regression  
41 and a machine learning algorithm were used to analyze the association of indicators of gut  
42 dysbiosis and species abundances, with all-cause and cause-specific mortality.

43 **Results** We identified two patterns representing overall microbiome community variation that  
44 were associated with both all-cause and cause specific mortality. Gut microbial distance to  
45 the average of the general population was associated with all-cause mortality and infection-,  
46 malignancy- and cardiovascular disease related mortality. Using multivariable Cox  
47 regression, we identified 23 species that were associated with all-cause mortality. By using a  
48 machine learning algorithm, we identified a log-ratio of 19 species predictive of all-cause  
49 mortality, all of which were also independently associated in the multivariable Cox-  
50 regression analysis.

51 **Conclusion** Gut dysbiosis is consistently associated with mortality in SOTR. Our results  
52 support the observations that gut dysbiosis is predictive of long-term survival. Since our data  
53 do not provide causative evidence, further research needs to be done to see determine  
54 whether gut-microbiome targeting therapies might improve long term outcomes

55

56 **Summary box**

57 **Significance of this study**

58 **What is already known on this subject?**

- 59     • Current literature suggests that the gut microbiome signature might be associated  
60         with mortality risk in the general population.
- 61     • Higher diversity of gut microbiota is associated with lower mortality in allogeneic  
62         hematopoietic-cell transplantation recipients.
- 63     • Liver and kidney transplant recipients suffer from gut dysbiosis and an analysis with a  
64         relatively low number of events showed that dysbiosis is associated with mortality.

65 **What are the new findings?**

- 66     • Across kidney, liver, heart and lung transplant recipients, we identified two overall  
67         microbial community variation patterns that are associated with all-cause mortality  
68         independent of the organ transplant and specifically to death from malignancy and  
69         infection.
- 70     • We find that multiple indicators of gut dysbiosis predict all-cause mortality and death  
71         by cardiovascular diseases, malignancy and infection.
- 72     • We find multiple microbial species associated with all-cause and cause-specific  
73         mortality. Using three different methods, we identify multiple bacterial species  
74         (shared between different analytical approaches) that are associated with an  
75         increased or decreased risk of mortality following solid organ transplantation.
- 76     • Using a machine learning algorithm, we identify a log-ratio of 19 bacterial species  
77         that was associated with all-cause mortality.

## 78 Introduction

79 Gut dysbiosis, while not clearly defined, is a condition typically characterized by the growth  
80 of pathogens at the expense of commensal bacteria when compared to a healthy  
81 microbiome. A dysbiotic gut microbiome has been observed in many diseases, including  
82 inflammatory bowel disease, obesity, diabetes mellitus and cancer.<sup>1-4</sup> Population-based  
83 studies report a large overlap in microbial associations to general health, suggesting a  
84 common dysbiotic signature in compromised health.<sup>5-7</sup>

85 Recent evidence suggests that such dysbiotic signatures are not only associated with  
86 a subject's health status at time of sampling but also predictive of long-term survival. The gut  
87 microbiome, characterized by metagenomic sequencing of stool samples, was associated  
88 with mortality in a well-characterized population-based study of 7,211 adults with a follow-up  
89 of 15 years in Finland.<sup>8</sup> This study found that members from the *Enterobacteriaceae* family  
90 were especially associated with an increased mortality risk.<sup>8</sup> While studies linking gut  
91 dysbiosis with patient survival in specific disease populations are scarce, the relationship  
92 has been studied more extensively in allogeneic hematopoietic-cell transplantation where a  
93 lower alpha diversity is associated with increased mortality risk.<sup>9,10</sup> We recently reported, in  
94 the setting of solid organ transplantation, that the extent of gut dysbiosis (in terms of the  
95 average distance from the general population) in both liver and kidney transplant recipients  
96 was associated with higher all-cause mortality risk.<sup>11</sup> However, the number of events in this  
97 study was relatively small due to the limited follow-up time.

98 In the current study, we present an in-depth survival analysis of both all-cause and  
99 cause-specific mortality in a large population of solid organ transplant recipients (SOTR). In  
100 the current study, we have 3.7 times more samples and 2.7 times more events from liver,  
101 kidney, heart, and lung transplantation recipients compared with our previous analyses. This  
102 population of SOTR represents an appropriate model to study the relationship between gut  
103 dysbiosis and long-term survival, because the population SOTR is characterized by  
104 polypharmacy, multimorbidity and the prevalence of dysbiosis is high compared with the  
105 general population.<sup>11,12</sup> Using this unique population of transplant recipients (n=1,337) and  
106 metagenomics samples from the general population (n=8,208), we analyzed the relationship  
107 between the gut microbiome and mortality. These findings are of interest for the  
108 transplantation community but also of interest for our general understanding of the gut  
109 microbiome and its relation to health.

## 110 **Results**

### 111 **Characteristics of solid organ transplant recipients**

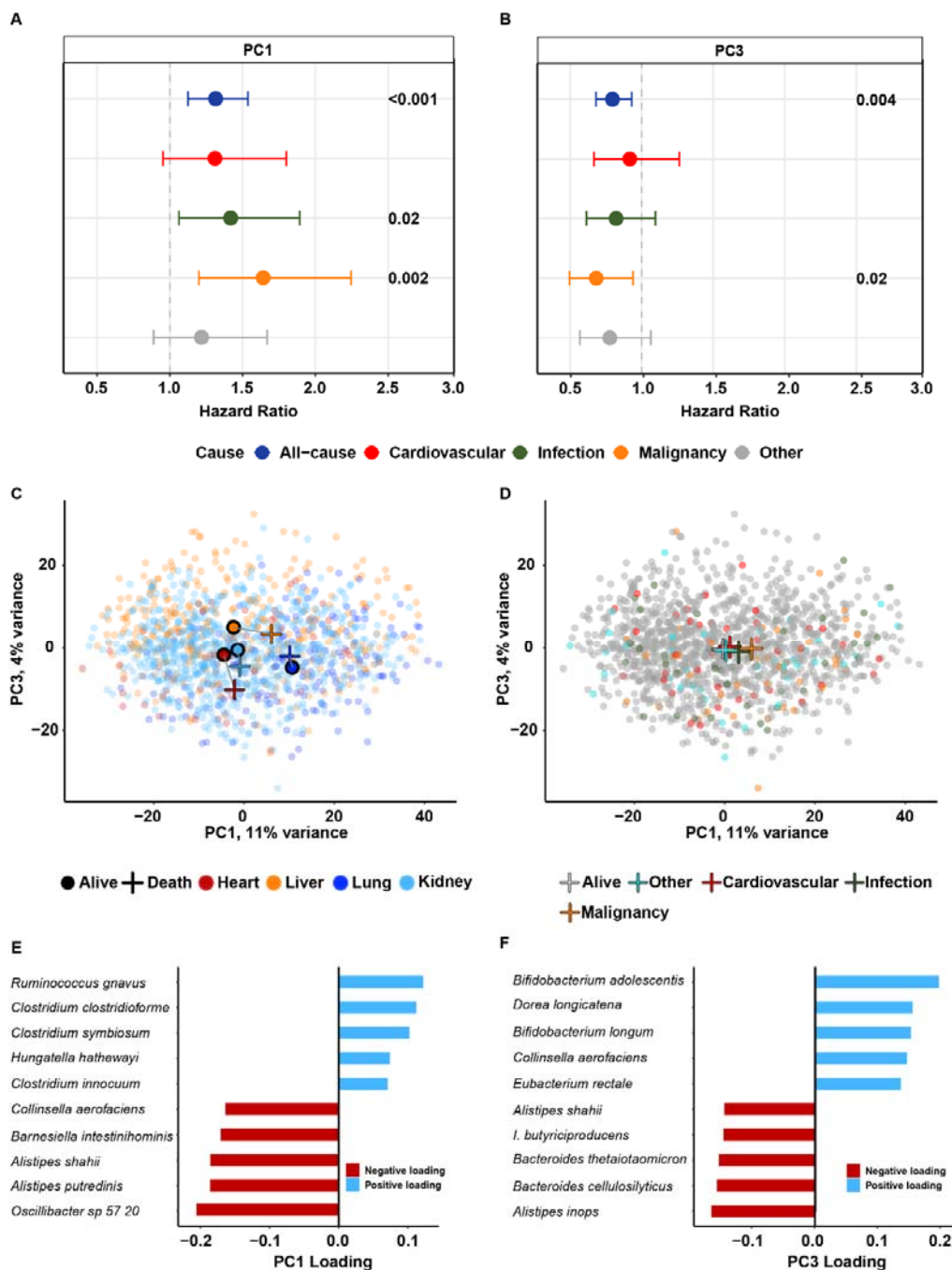
112 In total, 1,337 solid organ transplant recipients who provided a fecal sample at variable time  
113 after transplantation, including 766 kidney (KTR), 334 liver (LTR), 170 lung (LuTR) and 67  
114 heart (HTR) transplant recipients from the TransplantLines Biobank and Cohort study were  
115 included. The average age ( $\pm$  standard deviation, SD) of all recipients was 57 ( $\pm$  13.0) years,  
116 784 recipients (59%) were male and the average time since transplantation across all organ  
117 types was 7.6 ( $\pm$  8.0) years (**Supplementary Figure 1A, 1B and 1C**). During the follow-up  
118 period with a minimum of 2.8 years and a maximum of 6.5 years, a total of 162 participants  
119 (88 KTR, 33 LTR, 35 LuTR and 6 HTR) died (Supplementary Table 1). Of these, 48 (28%)  
120 died due to infection related mortality, 38 (23%) due to cardiovascular related mortality, 38  
121 (23%) due to malignancy related mortality, and 40 (25%) died from other causes  
122 (**Supplementary Table 1**).

123 In the first part of the analysis, we computed for each transplant recipient's gut  
124 microbiome, indicators of gut dysbiosis: Shannon diversity index, the distance to the average  
125 microbiome composition of the general population, richness of antibiotic resistance genes  
126 (ARGs) and virulence factors (VFs). With multivariable Cox regression including age, sex,  
127 BMI and years since transplantation, we subsequently investigated the relationship between  
128 each of these indicators and recipient all-cause and cause-specific mortality. While we  
129 analyzed all-cause mortality for each transplant organ type separately, we only performed  
130 the cause-specific analysis on all SOTR pooled, because of smaller numbers of these  
131 events. In the second part of the analysis, we aimed at identifying microbial species that  
132 individually or jointly predict mortality. We investigated the relationship between transplant  
133 recipient mortality and each species' CLR-transformed abundance, the quantile of each  
134 species in the general population. Lastly, we used a machine learning algorithm to identify  
135 which log-ratio of species best predict mortality.

### 136 **Overall community variation is associated with mortality risk**

137 We first performed a Principal Component Analysis (PCA) on CLR-transformed species  
138 abundances. In all SOTR we observed a significant relationship between principal  
139 component (PC) 1 and increased all-cause mortality and decreased all-cause mortality and  
140 PC3 (PC1: HR=1.32, 95% CI=1.13-1.54, FDR=5.8x10<sup>-4</sup>; PC3: HR=0.80, 95% CI=0.68-0.93,  
141 FDR=4.0x10<sup>-3</sup>; **Figure 1A and 1B**). PC3 was also associated with a lower mortality for KTR  
142 and HTR (KTR: HR=0.73, 95% CI=0.59-0.89, FDR=2.2x10<sup>-3</sup>; HTR: HR=0.34, 95% CI=0.12-  
143 0.95, FDR=0.03; **Figure 1C**). In the cause-specific analysis, we found that PC1 was related  
144 to death from malignancy and infection (malignancy: HR=1.64, 95% CI=1.20-2.24, FDR=2.0

145  $\times 10^{-3}$ ; infection: HR=1.42, 95% CI=1.06-1.89, FDR=0.01; **Figure 1A, 1B and 1D**), and PC3  
146 with a decreased risk of death from malignancy (HR=0.68, 95% CI=0.49-0.94, FDR= $2.0 \times$   
147  $10^{-2}$ ). The five species that exhibited the largest positive loadings onto PC1 (and thus were  
148 associated with mortality) were *Ruminococcus gnavus*, *Clostridium clostridioforme*,  
149 *Clostridium symbiosum*, *Hungathella hathewayi*, and *Clostridium innocuum* (**Figure 1E**), and  
150 the five species that exhibited the largest positive loadings onto PC3 (and thus were  
151 inversely associated with mortality) were *Bifidobacterium adolescentis*, *Dorea longitena*,  
152 *Bifidobacterium longum*, *Collinsella aerofaciens* and *Eubacterium rectale* (**Figure 1F**).

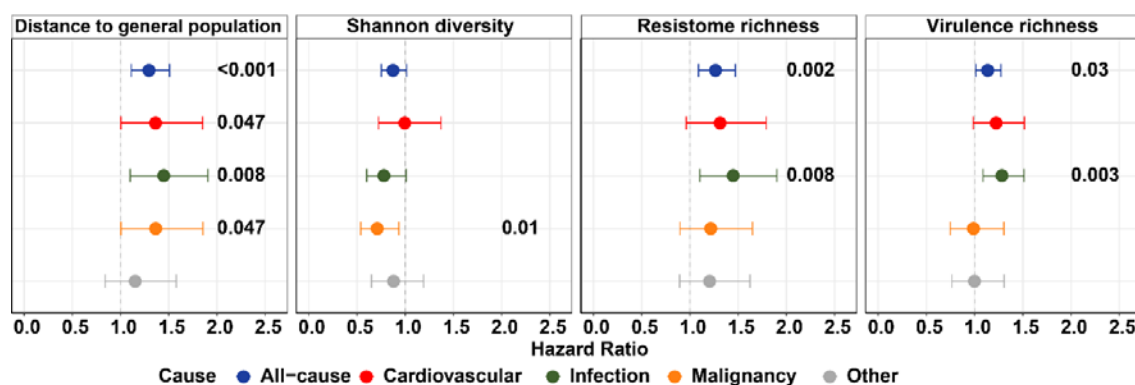


153 **Figure 1 (A and B)** Forest plot depicting results from Cox-regression analysis performed in  
 154 principal component 1 and 3 for all-cause and cause-specific mortality. **(C)** Principal  
 155 component analysis (PCA) colored by transplantation type. Circles and crosses represent  
 156 centroids of SOTR that are alive and dead at the time of the follow-up, respectively. **(D)** PCA  
 157 colored by cause of death showing the centroids (represented by crosses) of infection and  
 158 malignancy related mortality are separated from cardiovascular and other causes of death in  
 159 PC1. **(E and F)** Principal component loadings consisting of the 10 species with the largest

160 positive and negative loadings for PC1 and PC3, thus were associated with mortality. I.:  
 161 *Intestinimonas*.

162 **Multiple gut dysbiosis indicators are associated with all-cause and cause-**  
 163 **specific mortality**

164 In the cause-specific analyses performed in all SOTR, we found that the Shannon diversity  
 165 index was related to death from malignancy (HR=0.71, 95% CI=0.54-0.93, FDR=0.01). We  
 166 calculated how far each transplant recipient's gut microbiome was from the average  
 167 composition of the general population (Aitchison distance) and performed Cox-regression  
 168 analyses. The distance to the general population was significantly associated with higher  
 169 mortality risk in all SOTR (HR=1.29, 95% CI=1.11-1.51; FDR =  $9.3 \times 10^{-4}$ ; **Figure 2**,  
 170 **Supplementary Figure 2**). In the cause-specific analyses, the distance to the general  
 171 population was related to death from infection (HR=1.46, 95% CI=1.11-1.93, FDR= $7.2 \times 10^{-3}$ ),  
 172 malignancy (HR=1.39, 95% CI=1.02-1.89, FDR=0.03) and cardiovascular disease (HR=1.36,  
 173 95% CI=1.01-1.87, FDR=0.04; **Figure 2**). Finally, we found that harboring a higher richness  
 174 of antibiotic resistance genes (ARG) s and virulence factors (VFs) were associated with an  
 175 increased all-cause mortality risk (ARGs: HR=1.27, 95% CI=1.09-1.47; FDR= $2.2 \times 10^{-3}$ ,  
 176 **Figure 2**; VFs: HR=1.14, 95% CI=1.01-1.27; FDR=0.03; **Figure 2**) and with death from  
 177 infection in the cause-specific analyses (ARGs: HR=1.45, 95% CI=1.10-1.90, FDR= $8.0 \times 10^{-3}$ ;  
 178 VFs: HR=1.28, 95% CI=1.09-1.51, FDR= $3.0 \times 10^{-3}$ ; **Figure 2**). We did not observe a  
 179 significant association between the Shannon diversity index and all-cause mortality when we  
 180 analyzed recipients together and stratified by organ type (FDR>0.05, **Supplementary Table**  
 181 **2**). Overall, higher distance to the general population, increased in richness of ARG and VF,  
 182 and decreased in Shannon diversity were linked to increased mortality.



183 **Figure 2** Alpha diversity metrics for all-cause and cause-specific mortality analysis. We  
 184 calculated the Aitchison distance of each transplant recipient from the average composition  
 185 of the general population and identified a relationship with all-cause and cause specific

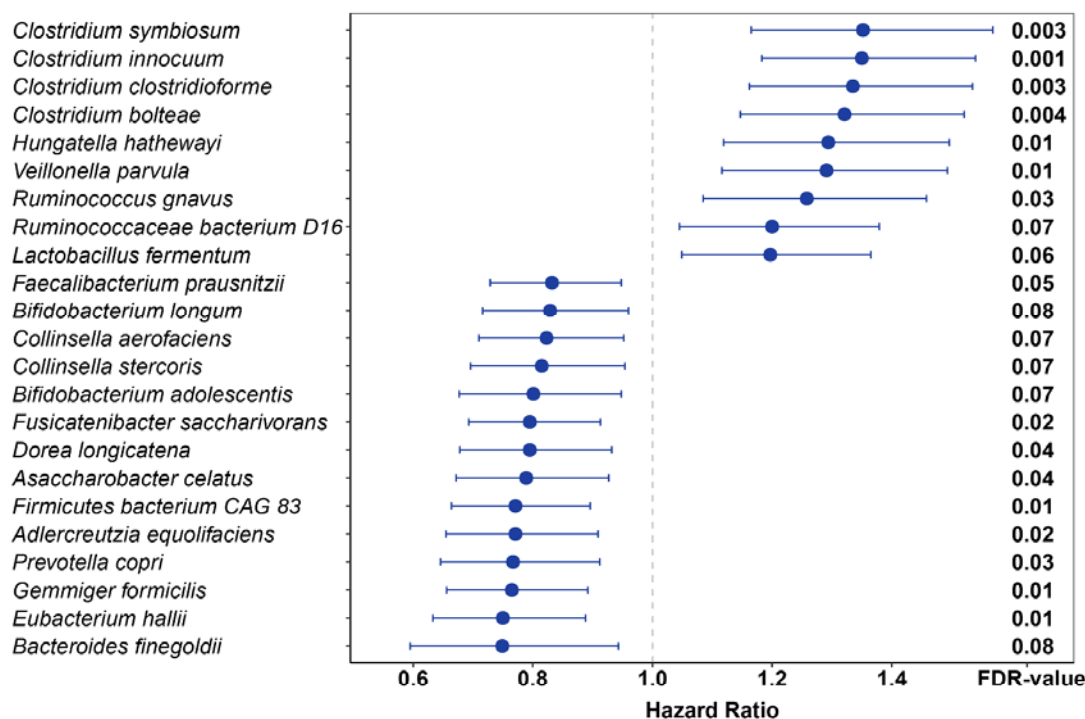


186 mortality. The Shannon diversity index was associated with malignancy related mortality.  
187 The richness of antibiotic resistance genes and virulence factors were both associated with  
188 all-cause and infection related mortality.

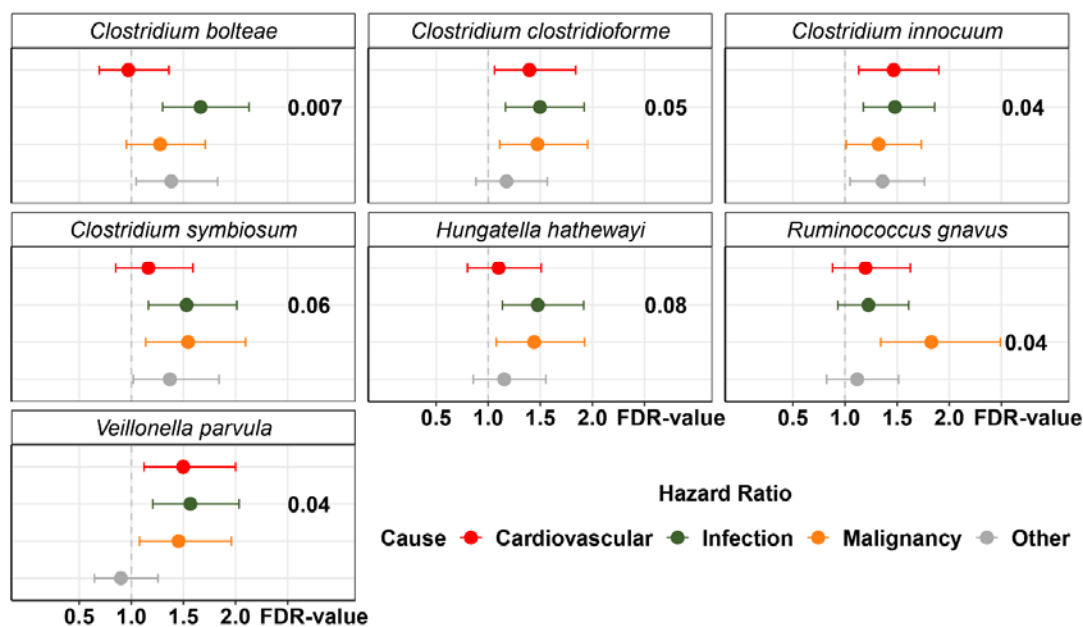
### 189 **Multiple species are associated with mortality**

190 We identified a total of 23 (16%; FDR<0.10) species whose CLR-transformed abundance  
191 were associated with all-cause mortality using multivariate Cox-regression including age,  
192 sex, BMI and years since transplantation (**Supplementary Table 5**). When we analyzed all  
193 SOTR, we found four *Clostridium* species (*C. innocuum* [HR=1.35, 95% CI=1.18-1.54,  
194 FDR=0.001], *C. clostridioforme* [HR=1.34, 95% CI=1.16-1.54, FDR=0.003], *C. symbiosum*  
195 [HR=1.35, 95% CI=1.17-1.57, FDR=0.003] and *C. boltea* [HR=1.32, 95% CI=1.157-1.52,  
196 FDR=0.004]) that were positively associated with all-cause mortality (**Figure 3A**;  
197 **Supplementary Table 5**) and death from infection in the cause-specific analyses  
198 (**Supplementary Table 6**; **Figure 3B**). Other species that were associated with an elevated  
199 mortality risk included *H. hathewayi* (HR=1.29, 95% CI=1.12-1.50, FDR=0.01), *Veillonella*  
200 *parvula* (HR=1.29, 95% CI=1.12-1.49, FDR=0.01) and *R. gnavus* (HR=1.26, 95% CI=1.09-  
201 1.46, FDR=0.03; **Figure 3A**; **Supplementary Table 5**). In the cause-specific analyses, we  
202 found that the abundance of *H. hathewayi* (HR=1.48, 95% CI=1.18-1.92, FDR=0.08) and *V.*  
203 *parvula* (HR=1.57, 95% CI=1.21-2.03, FDR=0.04) were related to death from infection  
204 (**Supplementary Table 6**; **Figure 3B**) and that the abundance of *R. gnavus* was related to  
205 death from malignancy (HR=1.83, 95% CI=1.34-2.49, FDR=0.04; **Supplementary Table 6**;  
206 **Figure 3B**). We also identified multiple species that were associated with a lower mortality  
207 risk when we analyzed all SOTR. For example, butyrate producers *Eubacterium hallii*  
208 (HR=0.75, 95% CI=0.63-0.89, FDR=0.01), *Firmicutes bacterium CAG 83* (HR=0.77, 95%  
209 CI=0.66-0.90, FDR=0.01), *Gemmiger formicilis* (HR=0.77, 95% CI=0.66-0.89, FDR=0.01)  
210 and *Faecalibacterium prausnitzii* (HR=0.83, 95% CI=0.73-0.95, FDR=0.05) were negatively  
211 associated with mortality (**Figure 3A**; **Supplementary Table 5**). Other commensals that  
212 were associated with a lower mortality risk included *Adlercreutzia equolifaciens* (HR=0.77,  
213 95% CI=0.66-0.91, FDR=0.02), *Prevotella copri* (HR=0.77, 95% CI=0.65-0.91, FDR=0.03),  
214 *Asaccharobacter celatus* (HR=0.79, 95% CI= 0.67-0.93, FDR=0.04) and *D. longicatena*  
215 (HR=0.80, 95% CI=0.68-0.93, FDR=0.04).

A

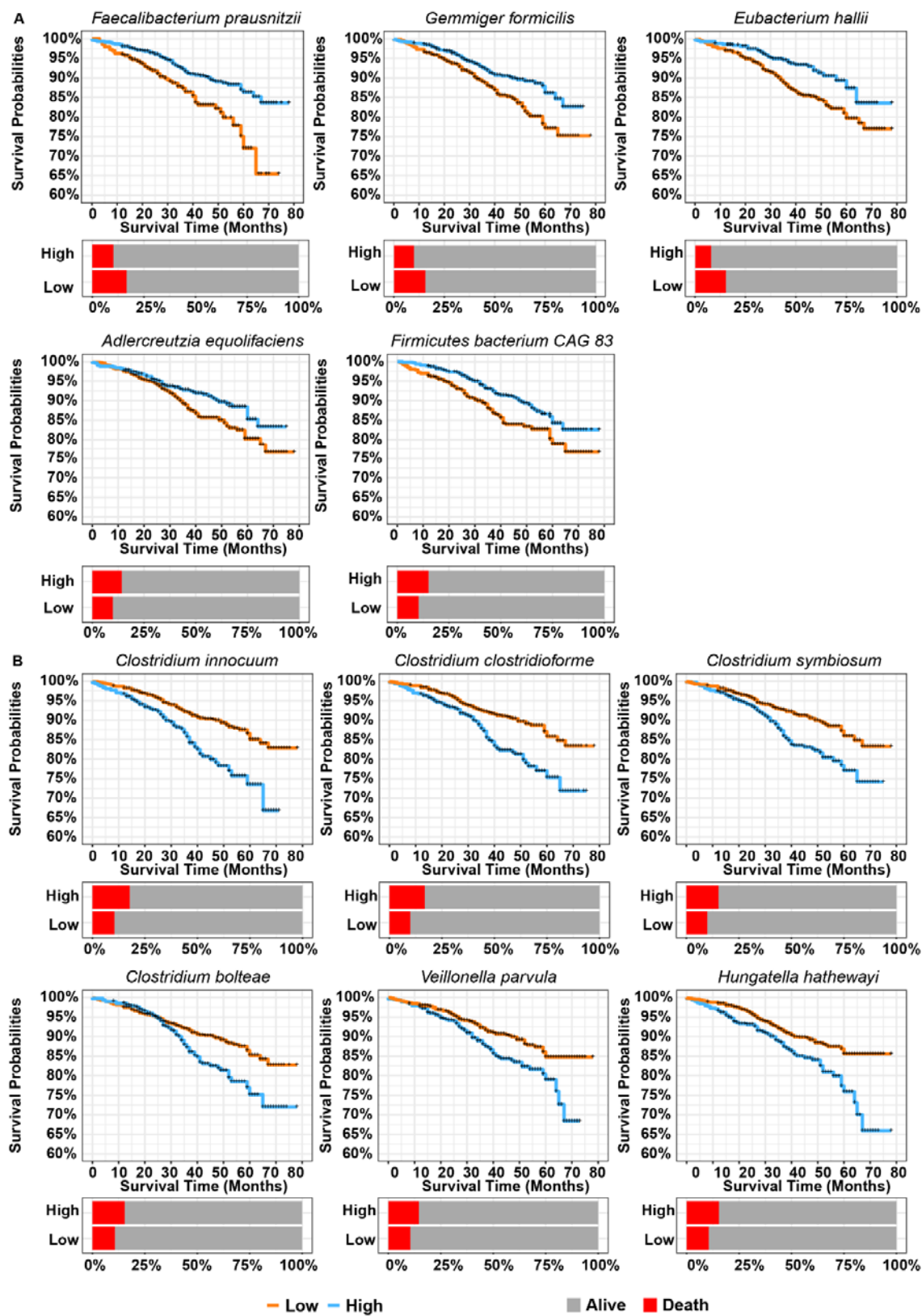


B



216 **Figure 3 (A)** Forest plot depicting all-cause mortality associated species from Cox-  
 217 regression analysis. Hazard-ratio and 95% confidence interval and the FDR-corrected value  
 218 are shown. **(B)** Forest plot depicting cause-specific mortality associated species from Cox-  
 219 regression analysis. Hazard-ratio and 95% confidence interval and the FDR-corrected value  
 220 are shown.

221           We took the analyses one step further by categorizing species based on whether  
222 their CLR-transformed abundance in each transplantation recipient were outside of its  
223 ‘normal’ range in the general population (higher [ $>75\%$  quantile] or lower [ $<25\%$  quantile])  
224 and whether this binary classification associated more strongly with mortality. This measure  
225 is relatively similar to our dysbiotic indicator ‘distance to the average of the general  
226 population’ but instead gives an indication of dysbiosis on the species as opposed to the  
227 microbial community level. When we analyzed all SOTR, we found many of the same  
228 associations (11/16 and 22/23 species at an  $FDR<0.05$  and  $FDR<0.1$ , respectively) but with  
229 stronger hazard ratios (at an  $FDR<0.05$ ; **Figure 4**). For example, the four *Clostridium*  
230 species (*C. innocuum*, *C. clostridioforme*, *C. symbiosum* and *C. bolteae*) that were positively  
231 associated with all-cause mortality in the previous analysis exhibited up to 1.8 times higher  
232 hazard ratios in this analysis (range of increase: min=1.3, median=1.35, max=1.8; **Figure 4**).

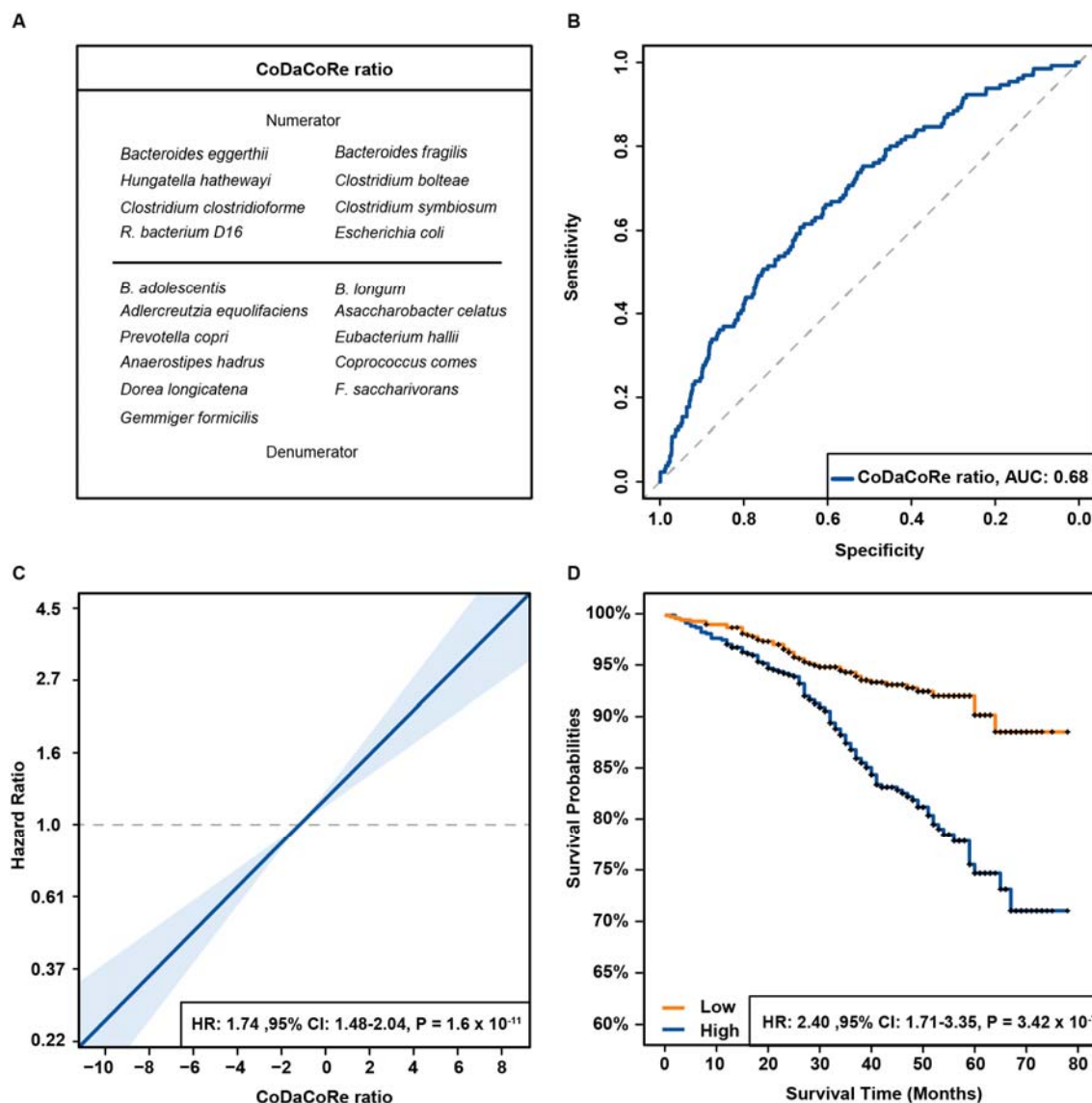


233 **Figure 4** Kaplan-Meier curves showing the mortality probabilities of recipients compared  
 234 with the general population. Bar plots depict the percent of deceased patients in the low and

235 high group, respectively. **(A)** Decreased mortality risk; i.e., if the relative abundance of the  
236 bacteria of a SOTR corresponds to the highest quantile of the general population there was  
237 a significantly lower risk for mortality. **(B)** Increased mortality risk; i.e., if the relative  
238 abundance of the bacteria of a SOTR corresponds to the highest quantile of the general  
239 population there was a significantly higher risk for mortality.

#### 240 **Gut microbiome as predictive biomarker of mortality**

241 Finally, to identify a predictive biomarker we used a machine learning algorithm  
242 (CoDaCoRe). This algorithm has been developed to identify the log-ratio most predictive of  
243 an outcome, in this case death (dead/alive at the time of censoring) in all SOTR (**see**  
244 **Methods**). We first split our data into a training and testing set (80/20) with a proportional  
245 number of events in each set. This algorithm identified a log-ratio consisting of 19 species  
246 (AUC=0.68; **Figure 5A and 5B**) that had a classification accuracy of 88% in the test set. In  
247 this log-ratio, the numerator – or species whose joint abundance is predictive of death  
248 consisted of eight species (*Bacteroides eggerthii*, *B. fragilis*, *H. hathewayi*, *C. bolteae*, *C.*  
249 *clostridioforme*, *C. symbiosum*, *Ruminococcaceae bacterium D16* and *Escherichia coli*;  
250 **Figure 5A; Supplementary Table 9**), and the denominator – or species whose joint  
251 abundance is predictive of mortality consisted of 11 species (*B. adolescentis*, *B. longum*,  
252 *Adlercreutzia equolifaciens*, *A. celatus*, *P. copri*, *E. hallii*, *Anaerostipes hadrus*, *Coprococcus*  
253 *comes*, *D. longicatena*, *Fusicatenibacter saccharivorans* and *G. formicilis*; **Figure 5A;**  
254 **Supplementary Table 9**). All of these species were individually identified to be associated  
255 with mortality in our previous analyses (FDR<0.10; **Figure 3A; Supplementary Table 5**).  
256 Finally, we tested whether the identified log-ratio also could predict mortality in a  
257 multivariable Cox regression including age, sex, BMI and years since transplantation. We  
258 found that this log-ratio (i.e., harboring more of the numerator species and less of the  
259 denominator species) is indeed associated with an elevated mortality risk (HR=1.74 ,95%  
260 CI=1.48-2.04, P=1.60 x 10<sup>-11</sup>, **Figure 5C**). When we categorized individual transplant  
261 recipients based on whether they harbored lower or higher than the median value of this log-  
262 ratio across all SOTR, we found that the hazard ratio increased by almost a factor of 1.4  
263 (HR=2.40, 95% CI=1.71-3.35, P=3.40 x 10<sup>-7</sup>, **Figure 5D**).



264 **Figure 5** Machine learning algorithm identifies a log-ratio predictive of mortality. **(A)** Species  
 265 included in the identified CoDaCoRe ratio. **(B)** AUC-ROC curve demonstrating discriminative  
 266 power of the most predictive log-ratio identified by CoDaCoRe in the training set. **(C)** Blue  
 267 line indicates the estimated hazard ratio compared to the log-ratio values with light blue area  
 268 representing the 95% confidence interval (CI) of the hazard ratio. **(D)** Kaplan-Meier curves  
 269 for recipients harboring lower (orange) and higher (blue) the median log-ratio value across  
 270 all SOTR. R.: Ruminococcaceae, B.: Bifidobacterium, F.: Fusicatenibacter.

## 271 Discussion

272 In the current study, we report an in-depth analysis of the gut microbiome in relation with  
273 both all-cause and cause-specific mortality in a population of SOTR from the  
274 TransplantLines cohort and biobank study.<sup>13</sup> We observed gut microbial signatures  
275 associated with both all-cause and cause-specific mortality, especially death from infection.  
276 The distance of the gut microbiome to general population controls, resistome and virulence  
277 factor richness were associated with a higher mortality risk. We found a consistent mortality  
278 related gut microbial signal consisting of previously disease-associated species.  
279 Interestingly, we discovered that if the abundance of a species among SOTR is outside what  
280 is considered the 'normal' or 'healthy' range in the general population, it also predicts  
281 mortality. Overall, our results show that gut dysbiosis related gut microbial signatures are  
282 associated with mortality across different SOTR.

283 Diversity analysis was partly consistent with previously reported results. Patients with  
284 a lower Shannon diversity index had 29% higher risk of malignancy related mortality.  
285 However, we did not observe a significant relationship between the Shannon diversity index  
286 and all-cause mortality in all the pooled SOTR analysis or stratified by organ-type. Thus, we  
287 are unable to confirm previous reported associations between the Shannon diversity index  
288 and mortality in HSCT-recipients and liver transplant recipients.<sup>10,11</sup> Similar to the gut  
289 microbiome mortality analysis in the general population, we observed a significant  
290 relationship between all-cause and cause-specific mortality and the PCA signature.<sup>8</sup> SOTR  
291 that were one standard deviation higher than average in PC1 had a 32% higher mortality risk  
292 while SOTR lower in PC3 had a 20% lower mortality risk. We observed a consistent gut  
293 microbial signal in all of our mortality analyses, i.e., the PCA mortality analysis, the gut  
294 dysbiosis indicator analysis, and the per-species and machine learning mortality analysis.

295 Many of the species that were associated with mortality in our study, have previously  
296 been associated with disease in the general population.<sup>5</sup> Specifically, we found that  
297 abundances of several clostridium species, including *C. clostridioforme*, *C. symbiosum*, *C.*  
298 *bolteae*, and *V. parvula* and *R. gnavus* were significantly associated with higher mortality in  
299 SOTR and with multiple diseases in the general population. In contrast, *P. copri*, *D.*  
300 *longicatena* and *F. prausnitzii* were associated with lower mortality in SOTR and with general  
301 health in the general population.<sup>5</sup> Furthermore, we observed that the extent of dissimilarity of  
302 the gut microbiome compared with the general population is associated with all-cause-,  
303 infection related-, cardiovascular related- and malignancy related mortality. This relationship  
304 was consistent with individual bacterial species in our study, but when compared to in the  
305 FINRISK study we were only able to confirm the observation of a lower mortality risk for  
306 SOTR with a higher abundance of *Faecalibacterium prausnitzii*.<sup>8</sup> However, in the FINRISK

307 study Cox-regression analysis was performed on the *genus* level and SHOGUN was used  
308 for taxonomical classification while we performed our analysis on the *species* level and used  
309 MetaPhlAn for taxonomic profiling, potentially limiting the comparability between the two  
310 studies.<sup>14</sup> Thus, further research using diverse populations and standardized methodology is  
311 needed to test whether our findings generalize to broader populations.

312 We observed a lower abundance of four butyrate producing bacteria is linked to  
313 increased mortality; *Gemmiger formicilis*, *Firmicutes bacterium* CAG 83, *Eubacterium hallii*  
314 and *Faecalibacterium prausnitzii*.<sup>15-18</sup> Butyrate is a short-chain fatty acid (SCFA) with a  
315 broad range of functionality including; microbiome modulation, anti-inflammatory activity,  
316 anti-obesity effect and antioxidant functions.<sup>19</sup> It was previously observed that KTR and LTR  
317 have a lower abundance of butyrate producing bacteria compared with controls and that a  
318 lower abundance of butyrate producing bacteria is associated with a lower health related  
319 quality of life in KTR.<sup>11,20-22</sup> We now find a higher mortality risk for SOTR with a lower  
320 abundance of butyrate producing bacteria. These results suggest that reduced butyrate  
321 levels could potentially have a direct role in mortality for SOTR. A lower abundance of  
322 butyrate producing bacteria was associated with increased occurrence of graft-vs-host  
323 disease and transplantation related mortality in HSCT recipients.<sup>23</sup> While measuring fecal  
324 SCFA was outside of the scope of the current study, future studies should evaluate fecal  
325 SCFA in relation with mortality. Our results warrant further studies into the role of butyrate  
326 producing bacteria and mortality in SOTR. The use of butyrate producing probiotics might  
327 offer a promising way to improve outcomes of solid organ transplantation.<sup>24</sup>

328 Strengths of the current study include a large sample size of the population of SOTR  
329 and the availability of a large control group from the Dutch population. With this dataset we  
330 were able to pinpoint a gut microbiome - mortality related signal in a group of SOTR with a  
331 high prevalence of dysbiosis. A limitation of the current study is that samples were not  
332 obtained at uniform time points after transplantation due to the cross-sectional nature of the  
333 cohort. Furthermore, we report results from an observational study which limits us to identify  
334 any causal relationships. It is possible that reverting dysbiosis will improve survival after  
335 transplantation, but it is also possible that the gut microbial signature that we observe is the  
336 effect of poor overall health and that the effect is not causal.

337 This study highlights a dysbiosis gut microbial - mortality signal in a population of  
338 SOTR with a high prevalence of dysbiosis. These findings are of interest for the transplant  
339 community as well as our general understanding of the relationship between the gut  
340 microbiome and health. Our results support emerging evidence showing that gut dysbiosis is  
341 predictive of long-term survival, indicating that gut-microbiome targeting therapies might  
342 improve patient outcomes although causal links should be identified first.



## 343 **Methods**

344

### 345 **Study design**

346 All SOTR cross-sectional gut microbiome data from the TransplantLines Biobank and Cohort  
347 study (Trial registration number NCT03272841) was included.<sup>13</sup> The TransplantLines study  
348 has been previously described in detail and aimed to include all potential adult solid organ  
349 transplant recipients and kidney donors at the University Medical Center Groningen  
350 (UMCG), The Netherlands, starting from June 2015.<sup>13</sup> We included 1337 fecal samples from  
351 SOTR. 8,208 subjects from the Dutch Microbiome Project were included to quantify the  
352 extent of dysbiosis and per species dysbiosis analysis.<sup>5</sup> Fecal samples from TransplantLines  
353 and DMP were collected using the same procedures and processed with the same DNA  
354 extraction protocols (see below). All participants signed an informed consent form prior to  
355 sample collection. TransplantLines (METc 2014/077) and Lifelines (METc 2017/152) were  
356 approved by the local institutional ethics review board (IRB) from the UMCG. Both studies  
357 adhere to the UMCG Biobank Regulation and are in accordance with the World Medical  
358 Association (WMA) Declaration of Helsinki and the Declaration of Istanbul.

### 359 **Clinical data**

360 In the TransplantLines study, every transplant recipient was asked to fill in questionnaires  
361 and blood, urine and fecal samples were collected. A detailed description, including details  
362 regarding the rationale of the study design, inclusion/exclusion criteria and randomization of  
363 the TransplantLines study is given by Eisenga *et al.*<sup>13</sup> In the current study, the primary  
364 outcome was overall survival. Clinical records were assessed to verify if a participant was  
365 alive or deceased, using a censoring date of January 1<sup>st</sup> 2022. If a patient was deceased, we  
366 assessed the cause of death and classified the cause of death into cardiovascular, infection,  
367 malignancy or other related mortality categories.

### 368 **Sample selection and gut microbiome data generation**

369

#### 370 **Fecal sample collection and subsequent processing**

371 Patients were asked to collect a fecal sample the day prior to the study visit. A FecesCatcher  
372 (TAG Hemi VOF, Zeijen, The Netherlands) was sent to the patients at home. Feces were  
373 collected and stored in appropriate tubes and frozen at home (at -18°C) immediately after  
374 collection. Frozen fecal samples were collected by UMCG personnel and stored at -80°C  
375 until DNA extraction.

## 376 **DNA extraction**

377 Microbial DNA was extracted using QIAamp Fast DNA Stool Mini Kit (Qiagen, Germany)  
378 according to the manufacturer's instructions. The QIAcube (Qiagen, Germany) automated  
379 sample preparation system was used for this purpose. Library preparation was performed  
380 using NEBNext® Ultra™ DNA Library Prep Kit for Illumina for samples with total DNA  
381 amount lower than 200ng, as measured using Qubit 4 Fluorometer, while samples with DNA  
382 yield higher than 200ng were prepared using NEBNext® Ultra™ II DNA Library Prep Kit for  
383 Illumina®. Libraries were prepared according to the manufacturer's instructions.  
384 Metagenomic shotgun sequencing was performed using Illumina HiSeq 2000 sequencing  
385 platform and generated approximately 8 Gb of 150 bp paired-end reads per sample (mean  
386 7.9 gb, st.dev 1.2 gb). Library preparation and sequencing were performed at Novogene and  
387 MGI.

## 388 **Metagenomic data processing**

389 Illumina adapters and low-quality reads (Phred score <30) were filtered out using KneadData  
390 (v0.5.1)<sup>25</sup>. Then Bowtie2 (v2.3.4.1)<sup>26</sup> was used to remove reads aligned to the human  
391 genome (hg19). The quality of the reads was examined using FastQC toolkit (v0.11.7).  
392 Taxonomy alignment was done by MetaPhlan3 (v3.7.2)<sup>26,27</sup> with the database of marker  
393 genes mpa\_v20\_m200. Metacyc pathways were profiled by HUMAnN2 (v0.11.1)<sup>28</sup>. Bacterial  
394 virulence factors and antibiotic resistance genes were identified using shortBRED  
395 [shortbred\_identify.py (v0.9.5) (51) and shortbred\_quantify.py tool (v0.9.5)] against virulence  
396 factors of pathogenic bacteria (VFDB) database (<http://www.mgc.ac.cn/VFs/main.htm>) and  
397 comprehensive antibiotic resistance database (CARD) (<https://card.mcmaster.ca/>)  
398 separately. Samples were further excluded in case of an eukaryotic or viral abundance  
399 >25% of total microbiome content or a total read depth <10 million. In total, we identified  
400 1132 taxa (17 phyla, 27 class, 52 order, 98 family, 231 genera and 705 species). After  
401 filtering for a prevalence of 10% and relative abundance threshold of 0.01%, 141 species  
402 were left. Hereafter, total-sum normalization was applied. Analyses were performed using  
403 locally installed tools and databases on CentOS (release 6.9) on the high-performance  
404 computing infrastructure available at UMCG and University of Groningen (RUG). An  
405 example of scripts used for microbiome processing is available at  
406 <https://github.com/GRONINGEN-MICROBIOME-CENTRE/TransplantLines>.

## 407 **Statistical analysis**

408 Centered log-ratio normalization was used due to the compositional nature of the  
409 metagenomic sequencing data<sup>29</sup>. PCA was performed using Euclidean distance between clr-  
410 transformed abundances (Aitchison distance<sup>30</sup>) of bacterial species. The Shannon diversity

411 index was calculated using the *vegan*<sup>31</sup> package in R. Cox proportional hazard models using  
412 the R packages *survival* and *rms* was used including age, sex, BMI and years since  
413 transplantation to analyze the association between diversity metrics and mortality. We used  
414 the *cox\_wrapper* function from Salosensaari *et al.* to analyze the relationship between the  
415 gut microbiome and mortality per species.<sup>8</sup> To further analyze the relationship between  
416 dysbiosis and the gut microbiome we used gut microbiome data from the general population  
417 and reclassified species abundance for SOTR according to quantiles of the general  
418 population. Cox-regression analysis was performed on this newly classified data.<sup>5</sup> To assess  
419 which ratio of bacteria best predicted mortality in SOTR we applied CoDaCore with mortality  
420 status as the dependent variable using balances as the log-ratio type, a lambda of 0 and a  
421 maximum for base learners of 1.<sup>32</sup> The most predictive ratio was then calculated per SOTR  
422 and Cox-regression analysis was performed on this ratio. False discovery rate was applied  
423 as a correction for multiple testing in all analysis.<sup>33</sup>

#### 424 **Data availability**

425 The raw microbiome sequencing data and basic phenotypes used in this study are available  
426 at the European Genome-Phenome Archive under accession numbers EGAD00001008907  
427 (<https://ega-archive.org/datasets/EGAD00001008907>), EGAS00001006257 ([https://ega-](https://ega-archive.org/studies/EGAS00001006257)  
428 [archive.org/studies/EGAS00001006257](https://ega-archive.org/studies/EGAS00001006257)) and EGAS00001006258 ([archive.org/studies/EGAS00001006258](https://ega-</a><br/>429 <a href=)). Due to patient confidentiality, the clinical datasets  
430 associated with the metagenomic datasets are available upon request to the University  
431 Medical Centre Groningen. Access to this clinical dataset requires a minimal access  
432 procedure consisting of a request per email ([datarequest.transplantlines@umcg.nl](mailto:datarequest.transplantlines@umcg.nl)) for a  
433 data access form. A response will be provided within two working weeks. This access  
434 procedure is to ensure that the data are being requested for research/scientific purposes  
435 only and thus complies with the informed consent signed by TransplantLines participants,  
436 which specifies that the collected data will not be used for commercial purposes.

437

#### 438 **Acknowledgements**

439 We would like to thank all participants from the TransplantLines- and Lifelines cohort and  
440 biobank study. We would like to thank the Center for Information Technology of the

441 University of Groningen (RUG) for support and for providing access to the Peregrine high-  
442 performance computing cluster and the Genomic Coordination Center (UMCG and RUG) for  
443 support and for providing access to Calculon and Boxy high-performance computing  
444 clusters. The TransplantLines Biobank and Cohort study received funding from Astellas BV  
445 (TransplantLines Biobank and Cohort study) and Chiesi Pharmaceuticals BV (PA-SP/PRJ-  
446 2020-9136) and was co-financed by the Dutch Ministry of Economic Affairs and Climate  
447 Policy by means of the PPP-allowance made available by the Top Sector Life Sciences &  
448 Health to stimulate public-private partnerships. Sequencing of the kidney part of the  
449 TransplantLines cohort was funded by a grant from the Dutch NWO/TTW/DSM partnership  
450 program Animal Nutrition and Health (project number 14939) to S.J.L.B. R.K.W. is supported  
451 by the Seerave Foundation, the Netherlands Organization for Scientific Research (NWO),  
452 and the EU Horizon Europe Program grant miGut-Health: personalized blueprint of intestinal  
453 health (101095470).

454

## 455 **References**

456

- 457 1. Vila, A. V., Imhann, F. & Collij, V. Gut microbiota composition and functional changes in  
458 inflammatory bowel disease and irritable bowel syndrome. *Sci. Transl. Med.* (2018).
- 459 2. Greathouse, K. L. *et al.* Gut microbiome meta-analysis reveals dysbiosis is independent  
460 of body mass index in predicting risk of obesity-associated CRC. *BMJ Open*  
461 *Gastroenterol.* **6**, e000247 (2019).
- 462 3. Zhou, Z., Sun, B., Yu, D. & Zhu, C. Gut Microbiota: An important player in type 2  
463 diabetes mellitus. *Front. Cell. Infect. Microbiol.* **12**, 834485 (2022).
- 464 4. Yang, Y. *et al.* Dysbiosis of human gut microbiome in young-onset colorectal cancer.  
465 *Nat. Commun.* **12**, 6757 (2021).
- 466 5. Gacesa, R. *et al.* Environmental factors shaping the gut microbiome in a Dutch  
467 population. *Nature* **604**, 732–739 (2022).

- 468 6. Gupta, V. K. *et al.* A predictive index for health status using species-level gut  
469 microbiome profiling. *Nat. Commun.* **11**, 4635 (2020).
- 470 7. Abbas-Egbariya, H. *et al.* Meta-analysis defines predominant shared microbial  
471 responses in various diseases and a specific inflammatory bowel disease signal.  
472 *Genome Biol.* **23**, 61 (2022).
- 473 8. Salosensaari, A. *et al.* Taxonomic signatures of cause-specific mortality risk in human  
474 gut microbiome. *Nat. Commun.* **12**, 2671 (2021).
- 475 9. Taur, Y. *et al.* The effects of intestinal tract bacterial diversity on mortality following  
476 allogeneic hematopoietic stem cell transplantation. *Blood* **124**, 1174–1182 (2014).
- 477 10. Peled, J. U. *et al.* Microbiota as Predictor of Mortality in Allogeneic Hematopoietic-Cell  
478 Transplantation. *N. Engl. J. Med.* **382**, 822–834 (2020).
- 479 11. Swarte, J. C. *et al.* Gut microbiome dysbiosis is associated with increased mortality after  
480 solid organ transplantation. *Sci. Transl. Med.* **14**, eabn7566 (2022).
- 481 12. Annavajhala, M. K. *et al.* Colonizing multidrug-resistant bacteria and the longitudinal  
482 evolution of the intestinal microbiome after liver transplantation. *Nature Communications*  
483 vol. 10 at <https://doi.org/10.1038/s41467-019-12633-4> (2019).
- 484 13. Eisenga, M. F. *et al.* Rationale and design of TransplantLines: a prospective cohort  
485 study and biobank of solid organ transplant recipients. *BMJ Open* **8**, e024502 (2018).
- 486 14. Usyk, M. *et al.* Comprehensive evaluation of shotgun metagenomics, amplicon  
487 sequencing, and harmonization of these platforms for epidemiological studies. *Cell Rep*  
488 *Methods* **3**, 100391 (2023).
- 489 15. Gophna, U., Konikoff, T. & Nielsen, H. B. Oscillospira and related bacteria - From  
490 metagenomic species to metabolic features. *Environ. Microbiol.* **19**, 835–841 (2017).
- 491 16. Engels, C., Ruscheweyh, H.-J., Beerenwinkel, N., Lacroix, C. & Schwab, C. The  
492 Common Gut Microbe *Eubacterium hallii* also Contributes to Intestinal Propionate  
493 Formation. *Front. Microbiol.* **7**, 713 (2016).
- 494 17. Kircher, B. *et al.* Predicting butyrate- and propionate-forming bacteria of gut microbiota  
495 from sequencing data. *Gut Microbes* **14**, 2149019 (2022).

- 496 18. Lopez-Siles, M., Duncan, S. H., Garcia-Gil, L. J. & Martinez-Medina, M.  
497 Faecalibacterium prausnitzii: from microbiology to diagnostics and prognostics. *ISME J.*  
498 **11**, 841–852 (2017).
- 499 19. Amiri, P. *et al.* Role of Butyrate, a Gut Microbiota Derived Metabolite, in Cardiovascular  
500 Diseases: A comprehensive narrative review. *Front. Pharmacol.* **12**, 837509 (2021).
- 501 20. Lee, J. R. *et al.* Butyrate-producing gut bacteria and viral infections in kidney transplant  
502 recipients: A pilot study. *Transpl. Infect. Dis.* **21**, e13180 (2019).
- 503 21. Wong, H. J. *et al.* Predictive and Prognostic Roles of Gut Microbial Variation in Liver  
504 Transplant. *Front. Med.* **9**, 873523 (2022).
- 505 22. Swarte, J. C. *et al.* Health-related quality of life is linked to the gut microbiome in Kidney  
506 transplant recipients. *Research Square* (2023) doi:10.21203/rs.3.rs-2578408/v1.
- 507 23. Meedt, E. *et al.* Prolonged Suppression of Butyrate-Producing Bacteria Is Associated  
508 With Acute Gastrointestinal Graft-vs-Host Disease and Transplantation-Related  
509 Mortality After Allogeneic Stem Cell Transplantation. *Clin. Infect. Dis.* **74**, 614–621  
510 (2022).
- 511 24. Zhu, L.-B., Zhang, Y.-C., Huang, H.-H. & Lin, J. Prospects for clinical applications of  
512 butyrate-producing bacteria. *World J Clin Pediatr* **10**, 84–92 (2021).
- 513 25. McIver, L. J. *et al.* bioBakery: a meta'omic analysis environment. *Bioinformatics* vol. 34  
514 1235–1237 at <https://doi.org/10.1093/bioinformatics/btx754> (2018).
- 515 26. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature*  
516 *Methods* vol. 9 357–359 at <https://doi.org/10.1038/nmeth.1923> (2012).
- 517 27. Truong, D. T. *et al.* MetaPhlan2 for enhanced metagenomic taxonomic profiling. *Nat.*  
518 *Methods* **12**, 902–903 (2015).
- 519 28. Franzosa, E. A. *et al.* Species-level functional profiling of metagenomes and  
520 metatranscriptomes. *Nat. Methods* **15**, 962–968 (2018).
- 521 29. Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V. & Egozcue, J. J. Microbiome  
522 Datasets Are Compositional: And This Is Not Optional. *Front. Microbiol.* **8**, 2224 (2017).
- 523 30. Pawlowsky-Glahn, V. & Egozcue, J. J. Compositional data and their analysis: an

- 524           introduction. *Geological Society, London, Special Publications* **264**, 1–10 (2006).
- 525   31. Dixon, P. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* **14**,
- 526           927–930 (2003).
- 527   32. Gordon-Rodriguez, E., Quinn, T. P. & Cunningham, J. P. Learning sparse log-ratios for
- 528           high-throughput sequencing data. *Bioinformatics* **38**, 157–163 (2021).
- 529   33. Korthauer, K. *et al.* A practical guide to methods controlling false discoveries in
- 530           computational biology. *Genome Biol.* **20**, 118 (2019).