



HAL
open science

Early nasopharyngeal microbial signature associated with severe influenza in children: a retrospective pilot study

Stanley Langevin, Maxime Pichon, Elise Smith, Juliet Morrison, Zachary Bent, Richard Green, Kristi Barker, Owen Solberg, Yves Gillet, Etienne Javouhey, et al.

► **To cite this version:**

Stanley Langevin, Maxime Pichon, Elise Smith, Juliet Morrison, Zachary Bent, et al.. Early nasopharyngeal microbial signature associated with severe influenza in children: a retrospective pilot study. *Journal of General Virology*, 2017, 98 (10), pp.2425-2437. 10.1099/jgv.0.000920 . hal-01909982

HAL Id: hal-01909982

<https://hal.science/hal-01909982v1>

Submitted on 20 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Early nasopharyngeal microbial signature associated with severe 2 influenza in children: a retrospective pilot study

3 Stanley Langevin^{1,2*}, Maxime Pichon^{3,4,5,*}, Elise Smith¹, Juliet Morrison¹, Zachary Bent²,
4 Richard Green¹, Kristi Baker¹, Owen Solberg², Yves Gillet⁶, Etienne Javouhey⁷, Bruno
5 Lina^{3,4,5}, Michael G. Katze^{1,8}, Laurence Josset^{3,4,5,#}

6 ¹Department of Microbiology, School of Medicine, University of Washington, Seattle, Washington, United
7 States of America

8 ²Department of Systems Biology, Sandia National Laboratories, Livermore, California, United States of
9 America

10 ³ University Lyon, Virpath, CIRI, INSERM U1111, CNRS UMR5308, ENS Lyon, Université Claude Bernard
11 Lyon 1, F-69372 Lyon, France

12 ⁴Laboratoire de Virologie, IAI, CBN, Groupement Hospitalier Nord, Lyon, France

13 ⁵Centre national de Reference Virus Influenzae, IAI, CBN, Groupement Hospitalier Nord, Lyon, France

14 ⁶Department of Pediatric Emergency, Hospices Civils de Lyon, Hôpital Femme Mère Enfant, Bron, France

15 ⁷Department of Pediatric Intensive Care, Hospices Civils de Lyon, Hôpital Femme Mère Enfant, Bron, France

16 ⁸Washington National Primate Research Center, University of Washington, Seattle, Washington, USA.

17 * These authors contributed equally to this work

18 # To whom correspondence should be addressed. Tel: +33 (0)4 72 07 10 22; Email: laurence.josset@chu-
19 lyon.fr

20 **KEYWORDS:** Microbiome; Influenza; Biomarker; Children; Emergency unit; Intensive care

21 **ABSTRACT WORD COUNT:** 240

22 **MANUSCRIPT WORD COUNT:** 5,075

23 **RUNNING TITLE:** Nasopharyngeal microbiota associated with influenza severity

24 **SEQUENCE ACCESSION NUMBER:** The Sequence Read Archive (SRA) accession number for the
25 raw sequences is SUB1938413.

Abbreviations: ACE: Abundance-based Coverage Estimator; ANOVA: ANalysis Of VAriance; CCEIRS: Centers of Excellence for Influenza Research and Surveillance; Ct: Cycle threshold; DNA: Desoxyribo Nucleic Acid; glm : generalized linear model; H: Shannon diversity index; ICU: Intensive Care Unit; IRB: Institutional Review Board; LOOCV: Leave-One-Out Cross-Validation; NMDS: Non-Metric MultiDimensional Scaling ; NP : NasoPharyngeal; OTU: Operational Taxonomic Unit; PRISM : Pediatric RISK of Mortality;RNA: RiboNucleic Acid; V1/V3 : Variable region of the 16S DNA;

26 **ABSTRACT**

27 A few studies have highlighted the importance of the respiratory microbiome in modulating
28 the frequency and outcome of viral respiratory infections. However, sufficient data is
29 lacking on the use of microbial signatures as prognostic biomarkers to predict respiratory
30 disease outcomes. In this study, we aimed to evaluate whether specific bacterial community
31 compositions in the nasopharynx of children at the time of hospitalization are associated
32 with different influenza clinical outcomes.

33 We utilized retrospective nasopharyngeal (NP) samples (n=36) collected at the time of
34 hospital arrival from children who were infected with influenza virus and had been
35 symptomatic for less than two days. Based on their clinical course, children were classified
36 into two groups: patients with mild influenza, and patients with severe respiratory or
37 neurological complications. We implemented custom 16S rRNA gene sequencing,
38 metagenomic sequencing and computational analysis workflows to classify the bacteria
39 present in NP specimens at the species level.

40 We found that increased bacterial diversity in the nasopharynx of children was strongly
41 associated with influenza severity. In addition, patients with severe influenza had decreased
42 relative abundance of *Staphylococcus aureus*, and increased abundance of *Prevotella*
43 (including *P. melaninogenica*), *Streptobacillus*, *Porphyrromonas*, *Granulicatella* (including
44 *G. elegans*), *Veillonella* (including *V. dispar*), *Fusobacterium* and *Haemophilus* in their
45 nasopharynx.

46 This pilot study provides proof-of-concept data for the use of microbial signatures as
47 prognostic biomarkers of influenza outcome. Further large prospective cohort studies are
48 needed to refine and validate the performance of such microbial signatures in clinical
49 settings.

50 INTRODUCTION

51 Seasonal influenza in children is a major burden on public health. It is estimated that
52 between 20 to 30% of children are infected with influenza each winter (WHO). While most
53 of the children develop uncomplicated influenza illness, between 1.5 to 4% of infected
54 children are hospitalized for influenza complications (1) including pneumonia, asthma
55 exacerbations, neurological complications, and death (2). Risk factors for hospital
56 admission in children include age <2 years and having underlying medical conditions (such
57 as neurological disorders, prematurity, immunosuppression, diabetes, or sickle cell disease)
58 (3). However, about half of pediatric hospitalizations from seasonal influenza occur in
59 children without a high-risk medical condition (4).

60 From a clinical perspective, failure to identify influenza-infected patients that have a higher
61 risk of developing severe disease delays the delivery of appropriate treatment, and may
62 have profound consequences for the recovery and long-term health of the patient. It has
63 been reported, in an adult population, that a delay of one day from onset of symptoms to
64 hospital admission increased the risk of death from influenza-caused disease by 5.5% (5).
65 Several attempts have been made to develop prognostic indicators that would help identify
66 patients at risk of severe influenza disease progression. Influenza viral load is not predictive
67 of clinical outcome (6), but increased serum levels for several cytokines in adults infected
68 with 2009 H1N1 influenza virus are correlated with disease progression (7), and increased
69 nasal levels of two proinflammatory cytokines are associated with severe influenza (6).
70 However, at present, there are no biomarkers readily available to the physician at the time
71 of hospital admission able to predict the likelihood of disease progression in all patients.

72 Improving severe influenza prognosis and treatment requires a better understanding of
73 influenza disease. Influenza severity is determined by a complex interplay between viral,
74 environmental and host factors, which is not yet fully understood and which cannot be

75 comprehended without looking at the system as a whole (8). One factor which has been
76 poorly studied and needs to be integrated in influenza pathogenesis models is the
77 microbiome, defined as the microbial communities colonizing the human body (microbiota)
78 together with the surrounding environment (9,10). Secondary bacterial pneumonia is a
79 well-known cause of pulmonary complications resulting from influenza infection, and was
80 reported as a major cause of mortality during the 1918 pandemic (11). However, the idea
81 that the microbiota might also be involved in influenza pathogenesis is only a recent
82 concept. Leung *et al.* recently reported that the oropharyngeal microbiota of a pool of
83 patients with H1N1 pneumonia was different from that of patients with non-H1N1
84 pneumonia (12). A second metagenomic study observed that patients infected with H1N1
85 had a higher amount of *Proteobacteria* in the upper respiratory tract than normal reference
86 nasopharyngeal (NP) microbial profiles previously published (13). This study also showed
87 a variation in NP microbiota depending on patient age. However, sequencing was
88 performed on pools of anonymized samples with no associated clinical data and could
89 therefore not link specific changes in microbiota composition with different influenza
90 outcomes.

91 In addition, several studies using mice have shown that the gut microbiota protects against
92 influenza disease. Antibiotic-treated mice exhibit impaired antiviral immune responses
93 following influenza virus infection, resulting in delayed viral clearance (14), while oral
94 administration of probiotics in mice prior to influenza infection reduces viral titers (15).
95 Finally, a recent mouse study showed that *Staphylococcus aureus* colonization in the upper
96 respiratory mucosa attenuated influenza infection compared with germ-free mice (16).
97 However, germ-free mice may not be appropriate to study the role of the microbiome in
98 influenza as they have important physiological defects such as altered immune systems,

99 which can perturb responses to influenza infection (reviewed in (17)). Additionally, there
100 are significant differences between the microbiomes of mice and humans (17).

101 Due both to the heterogeneity of the respiratory microbiome in humans and to the low
102 proportion of severe cases in seasonal influenza, large prospective studies are needed to
103 link the respiratory microbiome with influenza severity in patients. As a proof-of-concept,
104 we utilized NP specimens from a retrospective collection of influenza-positive samples
105 from children infected with influenza virus with different disease outcomes, mild or severe,
106 and tested whether specific bacterial species in the nasopharynx of patients at the time of
107 hospitalization correlated with influenza disease severity. We determined a differential
108 microbial signature discriminating patients with severe and mild influenza. Such a signature
109 could be used as a prognostic biomarker for early diagnosis of severe influenza.

110 **RESULTS**

111 **Selected patients**

112 We profiled the microbial communities of 36 NP samples collected at the time of hospital
113 arrival from patients classified retrospectively based on their clinical course into two
114 groups: patients with mild influenza (n=22) or severe influenza (n=14) outcome. Patients
115 had been symptomatic for less than 2 days before hospital admission. None of the patients
116 received antibiotics for up to three weeks before sample collection). After sampling, severe
117 patients were pre-emptively treated. Demographics and clinical parameters of the 36
118 patients are summarized in **Table 1** (baseline characteristics for **Table 1A** and follow-up
119 characteristics for **Table 2A**). Among patients who developed severe influenza, four
120 patients had severe respiratory complications requiring ventilatory support and
121 hospitalization in Intensive Care unit (ICU), and ten patients had neurological
122 complications requiring hospitalization in ICU (n=6 from 10, 60%) or in other units. The

123 most common neurological complications were coma (n=4 from 10, 40%) and epilepsy
124 (n=5 from 10, 50%). Average time of hospitalization was similar among the two severe
125 groups (7.4 and 9.8 days for the patients with neurological and respiratory complications,
126 respectively). Patients with mild influenza were either discharged or hospitalized in
127 emergency short-stay units. Their median time of hospitalization of one day corresponds to
128 a monitoring of more than 6 hours in emergency service, which is considered in France as a
129 hospitalization. Due to a too small number of patients included, no age- and sex- matched
130 analyses could be performed on this cohort with sufficient power. As respiratory
131 microbiome changes with age and sex (18), our results have been adjusted for these
132 parameters in the rest of the study using multivariate analysis (see Material and Methods
133 section). Finally, Cycle threshold (Ct) values from influenza real-time RT-PCR, as
134 described in the Material and Methods section, were used as a proxy measure of viral load
135 and were not significantly different between the groups, suggesting no relationship between
136 initial viral load and disease outcome, as previously described (6).

137 **NP microbiota composition of children infected with influenza virus**

138 Across all 36 respiratory samples, > 5.2 million high-quality 16S RNA sequences were
139 classified into 300 operational taxonomic units (OTUs), with 101 OTUs supported by >
140 1,000 reads each (**Fig S1**). All OTUs were classified up to the genus level and 49% were
141 classified up to the species level. Relative abundance for all 300 OTUs in each sample is
142 given (**Fig 1A**). Five OTUs dominated the NP bacterial profiles (defined as a relative
143 abundance > 50% in at least one sample). Among these five OTUs, the genera were
144 *Moraxella*, *Staphylococcus*, *Streptococcus*, and *Haemophilus* (**Fig 1B**). *Moraxella*
145 *catharrhalis* was found as the dominant species in 11 of 36 samples (30%). *Moraxella*
146 *nonliquefaciens* was also detected and was the dominant species in 4 samples.
147 *Staphylococcus aureus* was the dominant species in 2 samples. OTU_4 was classified only

148 to the genus level: *Streptococcus*, and was the dominant species in 5 samples. Finally,
149 OTU_5 (*Haemophilus*) was the dominant OTU in 4 samples. In the remaining 10 (28%)
150 other samples, there were more than a single dominant OTU: for 3 samples, there were two
151 major OTUs (with relative abundance between 25 and 50%), and for 7 samples, bacterial
152 profiles were more complex. Importantly the presence of a dominant OTU was not
153 significantly associated with influenza outcome after adjusting with covariates (**Table S1**).
154 To validate the bacteria taxonomy classifications inferred by the 16S ribosomal profiling,
155 we performed shotgun metagenomic sequencing on a subset of NP specimens obtained
156 from patients with mild (n=10) and severe (n=10) influenza. Comparable bacterial
157 compositions were found in the 16S targeted sequencing and shotgun metagenomic datasets
158 (**Fig S2**).

159 **NP microbiota composition differentiates influenza outcome**

160 To examine whether variations in NP bacterial communities composition were related to
161 clinical and/or demographic characteristics of the patients, a UniFrac distance matrix was
162 calculated and used in PerMANOVA analyses against five clinical variables (clinical
163 outcome, days between sampling and symptom onset, patient's age, patient's sex, sample
164 nature). This analysis revealed that differences in NP microbiota composition were strongly
165 associated with influenza outcome (p-value= 0.00025) (**Fig 2, Table S2**). This analysis,
166 confirmed by unsupervised clustering of the samples based on whole microbiota profiles,
167 was able to segregate mild from severe patients. All but one mild patient belonged in the
168 first cluster, and all but 4 severe patients belonged in the second cluster (**Fig S3**). Patient's
169 age was also associated with NP microbiota (p-value= 0.03450) but less strongly than
170 influenza outcome. There was no evidence of significant interaction between age and
171 clinical outcome on NP microbiota composition (**Table S2**). Finally, we tested whether
172 there was a difference in microbiota composition between patients with respiratory or

173 neurological complications. After adjusting for covariates, type of severe influenza
174 complication was not significantly associated with whole microbiota profiles (**Table S3**).

175 **Patients with mild influenza have a less diverse NP microbiota than patients that**
176 **develop severe influenza**

177 We further tested whether NP microbiota diversity (alpha diversity) was also associated
178 with patient's demographics and/or clinical features. Several diversity indices were
179 calculated (**Fig 3**) and were further used in statistical analysis to examine whether species
180 diversity in NP samples were related to five variables (clinical outcome, days between
181 sampling and symptom onset, patient's age, patient's sex, sample nature). Using multiway
182 ANOVA, we found that NP microbiota diversity, as estimated by the Shannon diversity
183 index (H), ACE, Chao1, Fisher, and Simpson indices, was strongly associated with
184 influenza outcome (**Table S4**). The Simpson's reciprocal index was the only one not
185 significantly associated with influenza outcome (**Table S4**). This observation could be
186 linked to the size of the cohort,
187 this index being the most sensitive to a lack of power. Patients with mild influenza had
188 significantly lower bacterial diversity (average $H=0.57$) than patients with severe influenza
189 (average $H=2.03$). The NP microbiota diversity was weakly associated with patient's age
190 (**Table S4**). There was no evidence of significant interaction between age and clinical
191 outcome, between sample nature and clinical outcome, or between age and sample nature
192 on NP microbiota diversity (**Table S4**). In addition, there was no significant difference in
193 microbiota diversity between patients with neurological and respiratory complications
194 (**Table S5**). Overall, we found that increased bacterial diversity in the upper respiratory
195 tract was associated with influenza severity.

196 **Differential microbial abundance between patients with mild or severe influenza**

197 Statistical analysis to characterize the differential microbial species abundance between
198 patients with mild influenza and patients with severe influenza was performed using a
199 generalized linear model adjusting for age, sex, time since symptom onset and sample
200 nature. Twelve OTUs were found differentially abundant between the two groups of
201 patients (**Fig 4A**). Average abundance in each group for the 12 differential OTUs is given
202 **Table 2**. Among the differential OTUs, one was among the five most abundant OTUs
203 detected in the respiratory microbiome profiles: *Staphylococcus aureus* was more abundant
204 in patients with mild influenza than in patients with severe influenza (**Table 2**). The
205 remaining 11 differential OTUs had limited abundance (<3% on average) and were all
206 more abundant in patients with severe influenza. The genera of these 11 OTUs were:
207 *Prevotella* for two of them, *Streptobacillus*, *Porphyromonas*, *Veillonella* for three of them,
208 *Fusobacterium*, *Haemophilus*, *Lachnospiraceae incertae sedis*, and *Granulicatella*. Some of
209 these OTUs were detected almost exclusively in patients with severe influenza, for example
210 *Porphyromonas* (OTU_31) was detected in 11 (76%) patients with severe influenza vs. in
211 only 2 (9%) patients with mild influenza.

212 In addition, we performed a similar statistical analysis on the severe influenza group only,
213 to determine whether there were differences between patients with neurological and
214 respiratory complications. Seven OTUs (*Moraxella catarrhalis*, *Actinomyces*,
215 *Corynebacterium*, *Dolosigranulum pigrum*, *Chryseobacterium*, *Prevotella* and *Parvimonas*
216 *micra*) were found differentially abundant (**Fig 4B, Table 3**). Except for *Moraxella*
217 *catarrhalis* and *Parvimonas micra*, these OTUs were significantly more abundant in
218 neurological severe forms than in respiratory severe forms. As already described,
219 *Moraxella catarrhalis* remain abundant in both severe form (17.39% vs. 29.89%), unlike all
220 others OTUs, with limited abundance in both groups (<1% in average).

221 **Minimal microbial signatures classify mild vs. severe cases**

222 To determine whether a smaller subset of OTUs could efficiently classify patients into mild
223 and severe groups, we used an association rule mining algorithm among differentially
224 abundant OTUs. To test the signatures' predictive performance, we used the trained rules to
225 determine influenza disease severity of patients excluded from the training set (leave-one-
226 out cross-validation (LOOCV)). In addition, only rules that were found in all cross-
227 validation instances were considered as robust. This analysis revealed 30–5 robust
228 association rules comprising two to three OTUs able to classify mild vs. severe patients
229 with cross-validated sensitivity up to 93% and specificity up to 100% (**Table 4**). For
230 instance, the quantification of OTU_22 (*Streptobacillus*) at a relative abundance higher
231 than 8.8e-06 and the quantification of OTU_85 (*Haemophilus*) at a relative abundance
232 higher than 4.8e-05 predict that the patient will develop a severe influenza with a sensitivity
233 of 92.86% and a specificity of 90.91%. The performances of these different rules need to be
234 validated in prospective cohort studies but these data indicate that microbial signatures
235 comprising of very few bacterial species have the potential to predict influenza outcome
236 with substantial accuracy.

237 Altogether, these results demonstrate that differences in NP microbiota composition early
238 after infection are associated with influenza severity and provide proof of concept data for
239 the use of microbial signatures as prognostic biomarkers of influenza outcomes in a clinical
240 setting.

241 **DISCUSSION**

242 Influenza virus transmission and pathogenesis in humans is unpredictable, at the population
243 and individual level in terms of clinical severity. While rapidly identifying patients at risk
244 of influenza complications is crucial for patient management and survival, there is currently

245 no host biomarker assay available to predict influenza severity at the time of infection. In
246 this study, we used a retrospective collection of early NP samples from children infected
247 with influenza virus to test the feasibility of using a prognostic microbial signature to
248 predict influenza disease outcome in a clinical setting. By comparing the relative
249 abundance and diversity of microbial populations in the upper respiratory tract of children
250 at the time of hospitalization, we identified bacterial profiles that differentiated children
251 that developed severe disease vs. non-progressive mild influenza infection, using five
252 different diversity indexes. All but one have shown statistical differences between these
253 two groups. This difference was not significant for the Simpson's reciprocal index, the
254 most dependent to cohort size (as reviewed in (19)). To our knowledge, this is the first
255 study associating influenza severity with NP bacterial composition in children.

256 There are a limited number of studies describing the composition of the respiratory tract
257 microbiota in healthy and diseased children. Studies profiling the NP microbiota have
258 reported that healthy children younger than 2 years have their upper respiratory tract
259 colonized by one predominant commensal bacterial genera consisting of *Moraxella*,
260 *Haemophilus*, *Streptococcus*, *Prevotella*, *Dolosigranulum*, *Staphylococcus*,
261 *Corynebacterium*, *Flavobacteria* or *Neisseria* (18,20–22). Similarly to these studies, we
262 found that most of the children infected with influenza had microbial compositions in their
263 nasopharynx that were dominated by one bacterial species, with *Moraxella*, *Streptococcus*,
264 *Haemophilus* and *Staphylococcus* representing the most abundant bacterial genera.
265 However, the nasopharynx microbiomes of children infected with influenza were not
266 colonized with gram positive commensal species (*Anoxybacillus*, *Dolosigranulum*,
267 *Corynebacterium* and *Lactococcus*). It is possible that influenza virus induced a decrease in
268 the abundance of these species early after infection. Interestingly, these gram positive
269 commensal species have been shown to control the abundant colonization of pathobionts

270 such as *H. influenza*, *S. pneumonia*, *S. aureus*, or *M. catarrhalis* and enhanced risk of
271 acquiring acute otitis media (21,23), pneumonia and bronchiolitis (24). In future studies, it
272 will be interesting to follow respiratory microbiome dynamics longitudinally prior and post
273 influenza infection to characterize early virus-microbiome interactions.

274 The most common genus found in our population was *Moraxella*, with two main species *M.*
275 *catarrhalis* and *M. nonliquefaciens*. High abundance of *M. nonliquefaciens* was previously
276 associated with viral pneumonia caused by adenovirus, rhinovirus or enterovirus,
277 respiratory syncytial virus, or human metapneumovirus infections (20). In addition,
278 *Moraxella* vs. *Alloiococcus* colonization is associated with increased respiratory syncytial
279 virus infections and severity (22). As we could not profile NP microbiota of healthy
280 children in parallel with influenza-infected children, we cannot determine whether
281 *Moraxella* colonization was also associated with influenza infection but we found that *M.*
282 *catarrhalis* and *M. nonliquefaciens* relative abundance were not predictive of influenza
283 severity, after adjusting for covariates. Further evaluation of the effect of specific virus-
284 microbiome interactions is necessary to better define the respiratory microbiota and its
285 influence on clinical disease outcome of acute respiratory infections in children.

286 In our study, severe influenza infection was predicted by increased bacterial diversity in the
287 nasopharynx of children. In addition, a microbial signature of 12 OTUs was able to
288 discriminate patients developing severe vs. mild influenza. Eleven of these 12 OTUs had
289 higher abundance in severe patients and included the following genera: *Prevotella*,
290 *Streptobacillus*, *Porphyromonas*, *Granulicatella*, *Veillonella*, *Fusobacterium*,
291 *Lachnospiracea incertae sedis* and *Haemophilus* spp. Except for *Haemophilus*, these
292 genera all belong to obligate or facultative anaerobic bacteria. Species that were found
293 more abundant in patients with severe influenza included: *Prevotella melaninogenica*, and
294 *Veillonella dispar* which are anaerobic gram-negative bacteria part of the normal oral (25)

295 and respiratory tract microbiota (26) and can cause anaerobic pulmonary infections (25) or
296 meningitis and endocarditis (27); and *Granulicatella elegans*, a member of the nutritionally
297 variant streptococci, which are common components of the oral microbiota that have been
298 associated with endocarditis and septicemia (28). Only one OTU was more abundant in
299 patients with mild symptoms: *Staphylococcus aureus*. This is in line with the study of
300 Wang *et al.* who reported that *S. aureus* priming prevents influenza-mediated lung injury in
301 a mouse model (16). The proposed mechanism was that *S. aureus* induces monocyte
302 polarization into M2 alveolar macrophages via Toll-like receptor 2 signaling, which inhibits
303 influenza-mediated lethal inflammation (16). Based on our pilot study, the higher
304 abundance of *S. aureus* in mild non-progressors could explain the lack of severe clinical
305 manifestations observed in our patient cohort, while the presence of a diverse nasopharynx
306 microbiota containing multiple pathobionts significantly increased the risk of acquiring
307 severe influenza-related disease. Furthermore, we have compared each of the two severe
308 influenza groups. Seven OTUs were significantly more diverse in one group, respiratory
309 group for two of them (*Moraxella catarrhalis* and *Parvimonas micra*), and neurological
310 group for the rest of them (*Actinomyces*, *Corynebacterium*, *Dolosigranulum pigrum*,
311 *Chryseobacterium* and *Prevotella*). Physiologically, all these bacteria were commensal of
312 the respiratory tract and oral cavity. They were mainly described in pneumonia and
313 periodontitis infections but could also be the cause of central nervous system infections,
314 bone infections and endocarditis (29–31). Interestingly, *Dolosigranulum pigrum*, associated
315 to the development of kerato-conjunctivitis, has already been characterized in respiratory
316 tract microbiome as belonging to an health-associated community, developing lately in
317 children born by caesarean section (32). Note that this analysis remains mostly
318 underpowered, due to the small number of patients in each severe group in this study.

319 It is noteworthy that microbial signature predicting influenza clinical outcome could be
320 reduced to only 2-3 OTUs. If proven to be robust in future studies, such signature creates
321 potential for the development of a real-time clinical assay using current technology. Rapid
322 PCR panels accurately quantifying both bacterial and viral species (35) could facilitate a
323 practical implementation of prognostic microbial signature-based assays simultaneously
324 with acute respiratory virus diagnostics in a clinical setting.

325 Our study has limitations inherent in its retrospective nature. First, despite the large number
326 of samples (n=372) from children infected with influenza virus available in our collection,
327 only 36 samples (9.7%) met our selection criteria. Because of the relatively low number or
328 our cohort, we could not match patients groups in sex and age. As age and sex were
329 described as important factors modifying the microbiota in the nasopharynx (18,22), we
330 had to adjust for these potential confounders in statistical models. We found that the main
331 factor influencing the NP microbiota was age. Sex was not associated with microbiota
332 composition and diversity, as expected for young prepubescent children. The microbiome
333 was also described as changing with seasons (18). However in our study, samples were
334 collected over different years but all during the influenza season (Jan-March) therefore
335 there was no seasonal effect to account for. Finally, due to the retrospective nature of our
336 pilot study we could not compare the performances of our signature with existing clinical
337 scores of severity, such as the Pediatric Risk of Mortality (PRISM III). Differences in
338 number between neurological and respiratory groups (n=10 vs. n=4) could be explained by
339 the administration of antimicrobial therapeutic in patients with respiratory complications
340 before the NP sample was collected. Indeed children with respiratory distress were most
341 likely to have pre-emptive antibiotics administrated at hospital admission before sampling,
342 and so were excluded from analyses. In contrast, when children developing neurological
343 complication had antibiotics administrated, this was mainly during hospitalization after NP

344 sample was collected, and treatment administrated prior to admission were mostly anti-
345 epileptic treatment.

346 Future studies will involve the enrollment of prospective cohorts of children admitted in the
347 emergency departments with longitudinal sample and clinical data collection to validate our
348 microbial signature performance and analyze host-virus-microbiome interactions. Studies
349 integrating microbiome analysis with a systems biology approach to influenza pathogenesis
350 will accelerate the development of novel diagnostic tools and personalized therapeutics to
351 combat influenza virus infections in children.

352 **METHODS**

353 **Ethical statement**

354 Respiratory samples (NP aspirate or swab) were collected for regular disease management
355 during hospital stay and no additional samples were taken. For the purpose of this study,
356 patient confidentiality was strictly protected and informed consent has been obtained. This
357 study was approved by the ethical committee of Hospices Civils de Lyon on Oct 14th, 2014.
358 The study was also approved by the University of Washington IRB under Expedited
359 Category 5 (Human Subjects Application #49811).

360 **Patient selection**

361 We analyzed 372 clinical records from children (≤ 15 years) who were admitted to a
362 pediatric hospital (Hôpital Femme-Mère-Enfant, Hospices Civils de Lyon, France) in 2011-
363 2014. All analyzed samples were tested positive for influenza A or B virus using
364 Respiratory Multi Well System MWS r-gene® kit (bioMérieux, Marcy-l'Étoile, France)
365 during the routine testing of the virology department of the University Hospital of Lyon.

366 Doing so, Ct could be obtained and were considered as an estimated viral load in the
367 sample.

368 In order to define a prognostic signature, only patients who were symptomatic for less than
369 2 days and had a respiratory sample collected at the time of hospital arrival were selected.
370 Based on their clinical course after sample collection, patients were classified into two
371 groups: patients with mild influenza or patients with respiratory or neurological
372 complications (the severe influenza group). Criteria used for classifying children in the
373 severe influenza group with respiratory complications were: utilization of invasive or non-
374 invasive ventilation, blood gas alteration (hypoxemia <95% in arterial sample), and
375 hospitalization in ICU for respiratory complications. Neurological complications were
376 defined as symptoms that affected the central or peripheral nervous system including
377 seizures, encephalopathy, encephalitis, or any focal neurologic symptom (33). Patients with
378 neurological complications were hospitalized either in ICU or in neurology units
379 specialized in the management of encephalitis or epilepsy. We excluded patients with
380 incomplete clinical files as well as patients receiving antibiotics prior to respiratory sample
381 collection as this treatment likely perturbs the microbiome. Other exclusion criteria
382 included known risk factors for severe influenza (defined as chronic respiratory diseases,
383 cardiac diseases, metabolic diseases, immunosuppression,...) and known chronic disease
384 associated with respiratory microbiota dysbiosis (such as severe asthma, ...)(3). Note that
385 two patients had Down's syndrome, one in the mild influenza group and one in the severe
386 influenza group. In addition, in the respiratory group, one patient presented a Prader-Willi
387 syndrome and one an uncategorized myopathy. While individuals with Down's syndrome,
388 Prader-Willi syndrome, and myopathy are known to be at risk of respiratory problems,
389 these comorbidities have not been associated with modifications in NP microbiome and we
390 therefore did not exclude these patients. In total, 72 patients were selected, including 42

391 patients with mild influenza, 11 patients with respiratory distress symptoms and 19 patients
392 with neurological complications. We extracted DNA and RNA using Nuclisens EasyMag.
393 Only 36 patients had samples with sufficient DNA quantity and quality for sequencing,
394 including 22 patients with mild influenza, 4 patients with respiratory distress symptoms and
395 10 patients with neurological complications.

396 **16S rRNA sequencing and analysis**

397 Genomic DNA was extracted from de-identified patient respiratory samples (n=36).
398 Custom primers were synthesized to amplify a 510bp fragment containing variable regions
399 V1-V3 of the 16S ribosomal gene. Briefly, 5ng of genomic DNA from each sample was
400 used as starting template to generate the V1-V3 amplicon libraries. All genomic DNA was
401 subjected to 20 cycles of PCR (Failsafe, Epicentre) and 16S amplicons were cleaned using
402 0.8X AMPure XP beads (Beckman Couter Inc) following manufacturer's instructions.
403 Nextera dual index adaptors (Nextera XT adaptors, Illumina Inc.) were incorporated by
404 performing 10 PCR cycles (Failsafe, Epicentre), cleaned using 1.1X AMPure XP beads
405 (Beckman Couter Inc), quantified using a Qubit (DNA High Sensitivity, Life
406 Technologies), and multiplexed using equal molar ratio of DNA for each sample. Final 16S
407 libraries were loaded on a MiSeq sequencer at 2 pM with 5% PhiX control and sequenced
408 using custom Illumina read primers to eliminate sequencing the V1-V3 16S primer
409 sequences (**Table S6**). Each sample had an average of 150K read depth (300bp, paired end)
410 and all 16S ribosomal sequences were classified using the UPARSE metagenomic pipeline
411 (34). Briefly, this pipeline removed low quality reads, merged the paired reads to generate
412 ~430bp fragment, removed sequence artifacts (chimeric sequences, sequence errors),
413 mapped full-length reads to a highly curated 16S ribosomal database derived from the RDP
414 16S training set (v14), identified sequence reads to genus/species level with a 97% cut-off
415 based on OTU classification, and generated summary tables for downstream statistical

416 analyses. In total, all samples passed our cut-off of at least 50,000 high-quality 16S DNA
417 sequence reads per sample and all OTU were classified to the genus level (**Fig S1**).
418 Rarefaction curves were performed to verify that sampling had exhausted the diversity at
419 the sequence read depth cut off of 50,000 reads per sample (**Fig S1C**). For classification to
420 the species level, we blasted the OTU representative sequences against the SILVA Ref
421 NR99 database (release 119) from which uncultured species were removed. We used blastn
422 with the following parameters ‘-evalue 1e-50 -perc_identity 99’. Blast results were further
423 filtered to keep only results that were identical (up to the genus level) to UPARSE
424 classification.

425 Raw sequence data have been deposited at SRA (SUB1938413).

426 **Metagenome sequencing and analysis**

427 Genomic DNA (5ng) from 10 mild and 10 severe influenza patients were fragmented and
428 tagged using the Nextera™ transposase following manufacturer’s instructions (Nextera™
429 XT DNA library kit, Illumina). Dual index adaptors were incorporated by PCR (14 cycles)
430 and final libraries were multiplexed for sequencing. The final library was loaded on a
431 NextSeq 500 using a 300 cycle high-output kit. All reads were de-multiplexed, adaptor
432 sequences were removed, and low quality reads were discarded. High quality reads were
433 aligned to a reference mammalian ribosomal RNA database and the human hg19 genome
434 using STAR aligner. Unmapped nonhuman reads were classified using Kraken metagenome
435 analysis software and the MiniKraken DB using default settings (35). Genus/species level
436 identification including total read counts for each bacterial group was determined.

437 **Statistical analysis**

438 All statistical analyses were performed using R Statistical Software. The R “phyloseq”
439 package (36) was used to calculate abundance-based richness estimators (e.g. Chao1,

440 ACE), and diversity indices (e.g. Shannon, Simpson and Fisher). Association between the
441 Shannon diversity index (H) and patient group was assessed by one way ANOVA with
442 Tukey multiple comparisons of means post hoc test. Multi-way ANOVA (type III) was
443 used to identify covariates associated with each diversity index. The R “phyloseq” package
444 was also used to calculate distance metrics (Bray-Curtis and weighted Unifrac) and
445 visualize samples into two dimensions using non-metric multidimensional scaling (NMDS).
446 Beta diversity analysis was performed on relative abundance data (*i.e* proportions: counts
447 are divided by total library size). Permutational multivariate analysis of variance
448 (PerMANOVA) was performed using adonis function in the R “vegan” package (37) to test
449 the ability of variables (patient group, sex and age, sample nature, days since symptom
450 onset) to account for observed variance in microbial profiles. Finally, statistical analyses for
451 the differential abundance of mapped OTU reads were conducted by the ALDEx2 R
452 package using a generalized linear model (glm) adjusting for covariates (age and sex) using
453 `aldex.glm` function (38). We have modified the `aldex.glm` function to enable testing the
454 effect of several variables in the glm model, instead of testing only one variable as in the
455 default `aldex.glm` function. Covariates that were taken into account were age, sex, time
456 since symptom onset and sample nature. Differential OTUs were defined using a p-value
457 cut-off of 0.05.

458 To determine whether small sets of OTUs could classify the patients, we used classification
459 based on association rules using the R “arules” package. For this analysis, the relative
460 abundance matrix of differentially abundant OTUs was transformed into a binary matrix
461 based on the median relative of each OTU, with 1 meaning that the relative abundance of
462 OTU_x in sample (i) is greater than the median value of OTU_x across all samples. Rules
463 were mined with $\text{support} > 0.3$, and $\text{confidence} > 0.8$. Redundant rules were further pruned.
464 We judged the robustness of the rules by using an LOOCV approach. For each subject,

465 association rules were predicted on all other subjects, excluding the one for which we made
466 a prediction. After we had performed this procedure for all subjects, we selected the rules
467 that were found in all instances and calculated their sensitivities and specificities from
468 predicted versus true classifications.

469 **FUNDING INFORMATION**

470 This work was supported by Federal funds from the National Institute of Allergy and
471 Infectious Diseases, National Institutes of Health, Department of Health and Human
472 Services, under Centers of Excellence for Influenza Research and Surveillance (CEIRS)
473 Contract No. HHSN272201400005C. Funding sources did not played any role in the study
474 or in the preparation of the article or decision to publish

475 **ACKNOWLEDGMENTS**

476 The authors thank Marcus Korth and Bruno Simon for valuable feedback on the
477 manuscript.

478 **CONFLICTS OF INTEREST**

479 The authors declare that there are no conflicts of interest.

480 **REFERENCES**

- 481 1. Belongia EA, Irving SA, Waring SC, Coleman LA, Meece JK, Vandermause M, et al.
482 Clinical characteristics and 30-day outcomes for influenza A 2009 (H1N1), 2008-2009
483 (H1N1), and 2007-2008 (H3N2) infections. *JAMA*. 2010;304(10):1091-8.
- 484 2. Dawood FS, Chaves SS, Pérez A, Reingold A, Meek J, Farley MM, et al.
485 Complications and associated bacterial coinfections among children hospitalized with
486 seasonal or pandemic influenza, United States, 2003-2010. *J Infect Dis*.
487 2014;209(5):686-94.
- 488 3. Gill PJ, Ashdown HF, Wang K, Heneghan C, Roberts NW, Harnden A, et al.
489 Identification of children at risk of influenza-related complications in primary and

- 490 ambulatory care: a systematic review and meta-analysis. *Lancet Respir Med.*
491 2015;3(2):139-49.
- 492 4. Committee on infectious diseases. Recommendations for Prevention and Control of
493 Influenza in Children, 2012-2013. *Pediatrics.* 2012;130(4):780-92.
- 494 5. Campbell A, Rodin R, Kropp R, Mao Y, Hong Z, Vachon J, et al. Risk of severe
495 outcomes among patients admitted to hospital with pandemic (H1N1) influenza. *Can*
496 *Med Assoc J.* 2010;182(4):349-55.
- 497 6. Oshansky CM, Gartland AJ, Wong S-S, Jeevan T, Wang D, Roddam PL, et al.
498 Mucosal Immune Responses Predict Clinical Outcomes during Influenza Infection
499 Independently of Age and Viral Load. *Am J Respir Crit Care Med.* 2014;189(4):449-
500 62.
- 501 7. Davey RT, Lynfield R, Dwyer DE, Losso MH, Cozzi-Lepri A, Wentworth D, et al.
502 The association between serum biomarkers and disease outcome in influenza
503 A(H1N1)pdm09 virus infection: results of two international observational cohort
504 studies. *PLoS One.* 2013;8(2):e57121.
- 505 8. Josset L, Tisoncik-Go J, Katze MG. Moving H5N1 studies into the era of systems
506 biology. *Virus Res.* 2013;178(1):151-67.
- 507 9. Weinstock GM. Genomic approaches to studying the human microbiota. *Nature.*
508 2012;489(7415):250-6.
- 509 10. Marchesi JR, Ravel J. The vocabulary of microbiome research: a proposal.
510 *Microbiome.* 2015;3:31.
- 511 11. Morens DM, Taubenberger JK, Fauci AS. Predominant role of bacterial pneumonia as
512 a cause of death in pandemic influenza: implications for pandemic influenza
513 preparedness. *J Infect Dis.* 2008;198(7):962-70.
- 514 12. Leung RK-K, Zhou J-W, Guan W, Li S-K, Yang Z-F, Tsui SK-W. Modulation of
515 potential respiratory pathogens by pH1N1 viral infection. Raoult D, éditeur. *Clin*
516 *Microbiol Infect.* 2013;19(10):930-5.
- 517 13. Chaban B, Albert A, Links MG, Gardy J, Tang P, Hill JE. Characterization of the
518 Upper Respiratory Tract Microbiomes of Patients with Pandemic H1N1 Influenza.
519 *PLoS ONE.* 2013;8(7).
- 520 14. Ichinohe T, Pang IK, Kumamoto Y, Peaper DR, Ho JH, Murray TS, et al. Microbiota
521 regulates immune defense against respiratory tract influenza A virus infection. *PNAS*
522 *USA.* 2011;108(13):5354-9.
- 523 15. Hori T, Kiyoshima J, Shida K, Yasui H. Augmentation of cellular immunity and
524 reduction of influenza virus titer in aged mice fed *Lactobacillus casei* strain Shirota.
525 *Clin Diagn Lab Immunol.* 2002;9(1):105-8.

- 526 16. Wang J, Li F, Sun R, Gao X, Wei H, Li L-J, et al. Bacterial colonization dampens
527 influenza-mediated acute lung injury via induction of M2 alveolar macrophages. *Nat*
528 *Commun.* 2013;4.
- 529 17. Fritz JV, Desai MS, Shah P, Schneider JG, Wilmes P. From meta-omics to causality:
530 experimental models for human microbiome research. *Microbiome.* 2013;1:14.
- 531 18. Bogaert D, Keijsers B, Huse S, Rossen J, Veenhoven R, van Gils E, et al. Variability
532 and Diversity of Nasopharyngeal Microbiota in Children: A Metagenomic Analysis.
533 Semple M, éditeur. *PLoS ONE.* 2011;6(2):e17035.
- 534 19. Chao A, Chiu C-H, Jost L. Unifying Species Diversity, Phylogenetic Diversity,
535 Functional Diversity, and Related Similarity and Differentiation Measures Through
536 Hill Numbers. *Annu Rev Ecol Evol Syst.* 23 nov 2014;45(1):297-324.
- 537 20. Sakwinska O, Bastic Schmid V, Berger B, Bruttin A, Keitel K, Lepage M, et al.
538 Nasopharyngeal Microbiota in Healthy Children and Pneumonia Patients. *J Clin*
539 *Microbiol.* 2014;52(5):1590-4.
- 540 21. Hilty M, Qi W, Brugger SD, Frei L, Agyeman P, Frey PM, et al. Nasopharyngeal
541 Microbiota in Infants With Acute Otitis Media. *J Infect Dis.* 2012;205(7):1048-55.
- 542 22. Teo SM, Mok D, Pham K, Kusel M, Serralha M, Troy N, et al. The infant
543 nasopharyngeal microbiome impacts severity of lower respiratory infection and risk of
544 asthma development. *Cell Host Microbe.* 2015;17(5):704-15.
- 545 23. Pettigrew MM, Laufer AS, Gent JF, Kong Y, Fennie KP, Metlay JP. Upper
546 Respiratory Tract Microbial Communities, Acute Otitis Media Pathogens, and
547 Antibiotic Use in Healthy and Sick Children. *Appl Environ Microbiol.*
548 2012;78(17):6262-70.
- 549 24. Vissing NH, Chawes BLK, Bisgaard H. Increased Risk of Pneumonia and
550 Bronchiolitis after Bacterial Colonization of the Airways as Neonates. *Am J Respir*
551 *Crit Care Med.* 2013;188(10):1246-52.
- 552 25. Schreiner A. Anaerobic pulmonary infections. *Scand J Infect Dis Suppl.* 1979;(19):77
553 -9.
- 554 26. de Steenhuijsen Pijpers WAA, Heinonen S, Hasrat R, Bunsow E, Smith B, Suarez-
555 Arrabal M-C, et al. Nasopharyngeal Microbiota, Host Transcriptome, and Disease
556 Severity in Children with Respiratory Syncytial Virus Infection. *Am J Respir Crit*
557 *Care Med.* 2016;194(9):1104-15.
- 558 27. Loughrey AC, Chew EW. Endocarditis caused by *Veillonella dispar*. *J Infect.*
559 1990;21(3):319-21.
- 560 28. Yacoub AT, Krishnan J, Acevedo IM, Halliday J, Greene JN. Nutritionally variant
561 streptococci bacteremia in cancer patients: a retrospective study, 1999-2014. *Mediterr*
562 *J Hematol Infect Dis.* 2015;7(1):2015030.

- 563 29. Abou Abdallah R, Cimmino T, Baron S, Cadoret F, Michelle C, Raoult D, et al.
564 Description of *Chryseobacterium timonianum* sp. nov., isolated from a patient with
565 pneumonia. *Antonie Van Leeuwenhoek*. 16 mai 2017;
- 566 30. Cobo F, Rodríguez-Granger J, Sampedro A, Aliaga-Martínez L, Navarro-Marí JM.
567 Pleural effusion due to *Parvimonas micra*. A case report and a literature review of 30
568 cases. *Rev Esp Quimioter*. août 2017;30(4):285-92.
- 569 31. Wong VK, Turmezei TD, Weston VC. Actinomycosis. *BMJ*. 11 oct 2011;343:d6099.
- 570 32. Bosch AATM, Levin E, van Houten MA, Hasrat R, Kalkman G, Biesbroek G, et al.
571 Development of Upper Respiratory Tract Microbiota in Infancy is Affected by Mode
572 of Delivery. *EBioMedicine*. juill 2016;9:336-45.
- 573 33. Khandaker G, Zurynski Y, Buttery J, Marshall H, Richmond PC, Dale RC, et al.
574 Neurologic complications of influenza A(H1N1)pdm09: Surveillance in 6 pediatric
575 hospitals. *Neurology*. 2012;79(14):1474-81.
- 576 34. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads.
577 *Nat Methods*. 2013;10(10):996-8.
- 578 35. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using
579 exact alignments. *Genome Biol*. 2014;15(3):46.
- 580 36. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive
581 Analysis and Graphics of Microbiome Census Data. Watson M, éditeur. *PLoS ONE*.
582 2013;8(4):e61217.
- 583 37. Oksanen J, Kindt R, Legendre P, O'Hara RB. *Vegan: Community Ecology Package*.
584 *Bioconductor*; 2015.
- 585 38. Fernandes AD, Reid JNS, Macklaim JM, McMurrough TA, Edgell DR, Gloor GB.
586 Unifying the analysis of high-throughput sequencing datasets: characterizing RNA-
587 seq, 16S rRNA gene sequencing and selective growth experiments by compositional
588 data analysis. *Microbiome*. 2014;2(1):15.
- 589

590 **FIGURES LEGENDS**

591 **Figure 1: NP microbiota composition associated with each respiratory sample.** A.
592 Heatmap displaying the relative abundance of the 300 OTUs detected in each NP sample.
593 Relative abundance is shown with a black to white gradient scale, with OTUs that were not
594 quantified in white, and OTUs with a relative abundance of 1 being in black. The samples
595 were ordered based on patients group and sample number. OTUs were clustered based on
596 their relative abundance values across samples using Euclidian distances and complete
597 linkage function. B. Barplot showing relative abundance of the 5 most abundant OTUs in
598 each sample, colored by OTUs. Samples were ordered as in panel A.

599 **Figure 2: The NP microbiota composition differentiates influenza severity.** Non-metric
600 Multidimensional Scaling (NMDS) plot comparing the global NP microbial profiles with
601 distances calculated using Unifrac distance. Each dot represents a sample: samples with
602 similar microbial profiles are close together in the NMDS plot, while increasing distances
603 in the plot between samples suggest divergent microbial profiles. Patients' groups are
604 depicted with different colors, and sample natures (NP Aspirate or Swabs) with different
605 shapes. Stress = 0.1102

606 **Figure 3: Patients developing severe influenza have more diverse NP bacterial**
607 **communities than patients with mild influenza.** Boxplots showing different alpha
608 diversity measures for the 3 groups of patients (Mild, Neurological, and Respiratory
609 complications).

610 **Figure 4: Bacterial species differentially abundant between patients with mild and**
611 **severe influenza (A) and between patients with respiratory and neurological**
612 **complications (B).** The heatmap shows the relative abundance of 12 OTUs that were found
613 differentially abundant between patients with mild or severe disease after adjusting for

614 covariates (p -value < 0.05). Relative abundance is shown with a black to white gradient
615 scale, with OTUs that were not quantified in white, and OTUs with a relative abundance of
616 1 being in black. The samples were ordered based on patients group and sample number.
617 OTUs were clustered based on their relative abundance values across samples using
618 Euclidian distances and complete linkage function.

619 **TABLE LEGENDS**

620 **Table 1: (A) Baseline demographic and clinical characteristics of all patients included**
621 **in the study. (B) Clinical evolution and therapeutic management of all patients**
622 **included in the study.**

623 **Table 2: Taxonomy and average relative abundance for the 12 OTUs differentially**
624 **abundant between mild and severe influenza groups after adjusting for covariates**
625 **(age, sex, sample nature, time since symptom onset).**

626 **Table 3: Taxonomy and average relative abundance for the 7 OTUs differentially**
627 **abundant between patients with respiratory and neurological complications after**
628 **adjusting for covariates (age, sex, sample nature, time since symptom onset).**

629 **Table 4: LOOCV performance characteristics of minimal microbial signatures in**
630 **discriminating patients with mild vs. severe influenza outcome.** Only rules predicted in
631 all cross-validation instances are displayed in this table. Sensitivity and specificity were
632 calculated using out-of-sample predictions. Rules are sorted based on specificity values in
633 descending order.