

1

2 **Title page**

3 Full title: Identifying Human Encounters that Shape the Transmission of *Streptococcus*

4 *Pneumoniae* and Other Respiratory Infections

5 Authors: Olivier le Polain de Waroux<sup>1</sup>, Stefan Flasche<sup>1</sup>, Adam J Kucharski<sup>1</sup>, Celine

6 Langendorf<sup>2</sup>, Donny Ndazima<sup>3</sup>, Juliet Mwanga-Amumpaire<sup>3,4</sup>, Rebecca F Grais<sup>2</sup>, Sandra

7 Cohuet<sup>5</sup>, W John Edmunds<sup>1</sup>

8 Affiliations:

9 1. Department of Infectious Disease Epidemiology, London School of Hygiene and

10 Tropical Medicine, Keppel Street, WC1E 7HT, London, United Kingdom

11 2. Department of Research, Epicentre, 8 Rue Saint-Sabin, 75011 Paris, France

12 3. Epicentre Mbarara Research Centre, PO Box 1956, Mbarara, Uganda

13 4. Mbarara University of Science and Technology, Mbarara University, PO Box 1410,

14 Mbarara, Uganda

15 5. Department of Field Epidemiology, Epicentre, 8 Rue Saint-Sabin, 75011 Paris, France

16

17 Corresponding author: Olivier Le Polain de Waroux. Email: [Olivier.lepolain@gmail.com](mailto:Olivier.lepolain@gmail.com)

18

19

20

21

22 **Abstract**

23 Although patterns of social contacts are believed to be an important determinant of  
24 infectious disease transmission, there is little empirical evidence to back this up. Indeed, no  
25 previous study has linked individuals' risk of respiratory infection with their current pattern  
26 of social contacts. We explored whether the frequency of different types of social  
27 encounters were associated with current pneumococcal carriage and self-reported acute  
28 respiratory symptoms (ARS), though a survey in Uganda in 2014. In total 566 participants  
29 were asked about their daily social encounters and about symptoms of ARS in the last two  
30 weeks. A nasopharyngeal specimen was also taken from each participant. We found that the  
31 frequency of physical (i.e. skin-to-skin), long ( $\geq 1\text{h}$ ) and household contacts – which capture  
32 some measure of *close* (i.e. relatively intimate) contact –, was higher among pneumococcal  
33 carriers than non-carriers, and among people with ARS compared to those without,  
34 irrespective of their age. With each additional physical encounter the age-adjusted risk of  
35 carriage and ARS increased by 6% (95%CI 2-9%) and 9% (1-18%) respectively. In contrast, the  
36 number of *casual* contacts (<5 minutes long) was not associated with either pneumococcal  
37 carriage or ARS. A detailed analysis by age of contacts showed that the number of close  
38 contacts with young children (<5 years) was particularly higher among older children and  
39 adult carriers than non-carriers, while the higher number of contacts among people with  
40 ARS was more homogeneous across contacts of all ages. Our findings provide key evidence  
41 that the frequency of *close* interpersonal contact is important for transmission of respiratory  
42 infections, but not that of *casual* contacts. Such results strengthen the evidence for public  
43 health measures based upon assumptions of what contacts are important for transmission,

44 and are important to improve disease prevention and control efforts, as well as inform  
45 research on infectious disease dynamics.

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62 **Author summary**

63 Although social contacts are an important determinant for the transmission of many  
64 infectious diseases it is not clear how the nature and frequency of contacts shape individual  
65 infection risk. We explored whether frequency, duration and type of social encounters were  
66 associated with someone's risk of respiratory infection, using nasopharyngeal carriage (NP)  
67 of *Streptococcus pneumoniae* and acute respiratory symptoms as endpoints. To do so, we  
68 conducted a survey in South-West Uganda collecting information on people's social  
69 encounters, respiratory symptoms, and pneumococcal carriage status. Our results show  
70 that both pneumococcal carriage and respiratory symptoms are independently associated  
71 with a higher number of social encounters, irrespective of a person's age. More specifically,  
72 our findings strongly suggest that the frequency of *close* contacts is important for  
73 transmission of respiratory infections, particularly pneumococcal carriage. In contrast, our  
74 study showed no association with the frequency of short *casual* contacts. Those results are  
75 essential for both improving disease prevention and control efforts as well as informing  
76 research on infectious disease dynamics and transmission models.

77

78

79

80

81

82

83 **Introduction**

84 The transmission of respiratory infections is likely to depend on the frequency and age  
85 structure of human social contacts, as well as other factors including pre-existing immunity  
86 from prior infection or vaccination [1-3]. To understand the dynamics of such infections,  
87 studies to quantify social mixing patterns have been conducted in various settings, under  
88 the assumption that self-reported encounters reflect transmission probabilities of  
89 pathogens transmitted through close contact [2, 4-11]. Combined with disease transmission  
90 models, these data are increasingly being used to inform infection control policies [12, 13].

91 There is evidence from population-based models that self-reported social mixing patterns  
92 can reproduce observed aggregated seroprevalence data for chickenpox [14], mumps [4],  
93 parvovirus [15], influenza [4, 16, 17] and whooping cough [18]. Moreover, it has been  
94 suggested that age-stratified social mixing patterns can capture individual influenza risk, as  
95 measured by a four-fold rise in neutralization titres over the course of an epidemic (17, 18).  
96 However, it remains unclear precisely how risk of infection is related to the frequency and  
97 nature of an individual's social encounters around the time of infection.

98 To establish how social behaviour shapes individual-level infection, we explored whether  
99 the frequency and duration of different types of social encounters were associated with an  
100 individual's risk of respiratory infection, using nasopharyngeal (NP) carriage of *Streptococcus*  
101 *pneumoniae* (the pneumococcus) and self-reported acute respiratory symptoms (ARS) as  
102 endpoints. *S.pneumoniae* is one of the main causes of pneumonia and sepsis globally [19],  
103 disproportionately so in low-income settings [19-21]. Colonization of the nasopharynx is a  
104 precondition to disease, and the main source of human-to-human transmission. Given that

105 most episodes of carriage remain asymptomatic, social behaviour is unlikely to change as a  
106 result of carriage, making pneumococcal carriage a more suitable endpoint than  
107 symptomatic illness to explore the association between social behaviour and infection risk,  
108 given that people tend to limit their contacts during symptomatic illness [22]. In addition, as  
109 natural immunity to carriage is weak [23], pre-existing immunity is less likely to confound  
110 associations between disease and social contact patterns than in studies using immunizing  
111 infections as biological endpoint [16, 17].

112 We analysed data from a social contact study nested within a survey of NP carriage of  
113 *S.pneumoniae* conducted across a rural South-West Uganda in 2014, before the introduction  
114 of the pneumococcal conjugate vaccine (PCV). Participants across age groups were  
115 randomly selected from 56 villages and one small town. Participants were asked to name all  
116 individuals with whom they had a conversational encounter lasting  $\geq$  5 minutes between  
117 wake-up on the day prior to the survey day to wake-up on the survey day ( $\sim$  24 hours).  
118 Those encounters are further referred to as '*ordinary*' contacts. For very short contacts (i.e.  
119 <5 minutes), which we here define as '*casual*' contacts, an estimate of the number of  
120 contacts was asked (<10, 10-19, 20-29,  $\geq$ 30) without further details. ARS were defined as  
121 any of the following symptoms in the previous two weeks: sore throat, sneezing, difficulty  
122 breathing, and runny nose. A nasopharyngeal specimen was taken from each survey  
123 participant and was processed and analysed as per WHO guidelines [24]. Using these  
124 individually-matched data, we examined whether the type and frequency of social contacts  
125 differed by carriage status and between individuals with and without ARS.

126

127

128 **Results**

129 Study population

130 Of the 687 individuals initially targeted for inclusion, 568 (83%) individuals from unique  
131 households responded to the survey, with data on both social contacts and carriage or ARS  
132 available from 566 participants. Participant's age spanned across age groups and the sex  
133 distribution was reasonably balanced (58% female). The majority (98%) of children aged 6 –  
134 15 years attended school or college. More than a third of all adults (36%) worked in  
135 agriculture and 22% were homemakers/housewife.

136 On average, people made seven '*ordinary*' contacts (defined as contacts  $\geq 5$  minutes long),  
137 ranging from 0 to 25, the majority of which were physical (i.e. 'skin-to-skin' contact or  
138 indirect physical contact through utensils passed from mouth-to-mouth). There was no  
139 evidence that the average number of contacts differed by weekday or between weekdays  
140 and weekends ( $P=0.623$ ). Children aged 5–9 years reported most contacts and children <5  
141 years the fewest (Figure 1). The majority of contacts made by children were physical.

142

143 **Figure 1:** Number of reported contacts by age group and type of contact. Legend: Boxplots  
144 showing the median (horizontal line), interquartile range (boxes) and 95% confidence  
145 interval around the mean (whiskers). The outliers are shown as bubbles above or under the  
146 whiskers

147

148 The most intense mixing tended to be between individuals of the same age group (i.e.  
149 assortative mixing), but there was also substantial mixing between age groups. Contact from  
150 and with children aged <10 years involved proportionally more physical touch than contacts  
151 between older children and adults (Figure 2A and 2B). The proportion of non-physical  
152 contacts, contacts outside the household and contacts of short (<1h) duration was higher  
153 among teenagers and adults than among younger children and infants.

154 There was no difference in contact patterns by sex, for all types of contacts considered.

155 Four hundred and ninety (87%) participants estimated how many casual contacts (i.e. <5  
156 minutes in duration) they had the day before the survey. Over a third (36%) of these  
157 reported 10 or more casual contacts, and 11% reported more than 20 casual contacts.  
158 Among the 13% who could not estimate how many casual contacts they had, there were  
159 proportionally more children with over half (56%) under 10 years. The mean number of  
160 reported ‘ordinary’ contacts was higher among individuals reporting ≥10 casual contacts  
161 than those reporting fewer than 10 contacts (mean 8.9 vs 5.9, P<0.001), as well as among  
162 the 78 individuals who did not know how many casual contacts they had (mean 8.8,  
163 P<0.001).

164 The prevalence of pneumococcal carriage was strongly age dependent, decreasing from 75%  
165 in children <5 years to 46% among 5 – 9 year olds, 17% in 10 – 19 year olds, and further  
166 decreasing to 8% and 7% among 20-39 years and ≥40 years old respectively (Figure 2C).  
167 Overall, 72 (13%) people reported having suffered from ARS in the two weeks prior to the  
168 survey. The prevalence of ARS varied much less with age than that of carriage (Figure 2D),  
169 ranging from 20% among 5 to 9 year olds to 8% among ≥40 years old. There was no sex

170 difference in the prevalence of carriage (age-adjusted RR for males 0.97, P=0.814) or ARS  
171 (age-adjusted RR for males 1.40, P=0.134). Carriage and ARS were poorly correlated  
172 (Pearson's correlation coefficient R=0.09, ranging from -0.08 to 0.29 by age group), and  
173 there was no evidence that the risk of ARS was higher among carriers compared to non-  
174 carriers (age-adjusted relative risk (RR) 1.06 (95%CI 0.85 – 1.33)).

175

176 **Figure 2:** Contact matrices and prevalence of *S.pneumoniae* carriage and Acute Respiratory  
177 Symptoms (ARS), by age. Legend: Contact matrices and prevalence of *S.pneumoniae*  
178 carriage (Panel A) and Acute Respiratory Symptoms (Panel B). In panels C and D, the height  
179 of each bar corresponds to the point prevalence and the error bar represents the 95%  
180 Confidence Interval (CI).

181

182 Social contacts as a risk for pneumococcal carriage or ARS

183 Overall, the mean number of contacts among carriers was significantly higher than non-  
184 carriers, and this observation was consistent across age groups, although most differences  
185 were not statistically significant due to small numbers (Figure 3). In particular, carriers had  
186 more physical contacts (Fig 3A). This pattern was also consistent for ARS, with the exception  
187 of 5 – 9 year olds in whom the mean number of contacts among individuals with ARS was  
188 lower for physical contacts (Fig 3B).

189

190 **Figure 3:** Mean number of contacts by age group and nasopharyngeal carriage status (Panel  
191 A) or ARS status (Panel B). Legend: The graph shows the mean number of contacts among  
192 carriers (red, panel A), non-carriers (blue, panel A), individuals reporting ARS (red, panel B),  
193 and individuals without ARS (blue, panel B). This is shown by age group and for all contacts,  
194 physical contacts, household contacts and contacts lasting over 1 hour.

195 In univariable analysis, the risk of carriage increased with all contacts, household contacts,  
196 contacts  $\geq 1$  hour long and physical contacts. The latter had the largest effect size, with a  
197 13% increased risk for each additional contact reported by participants (Table 1). Physical  
198 contacts and contacts  $\geq 1$  hour were strongly correlated ( $R = 0.76$ ), particularly among  
199 children  $<5$  years ( $R=0.85$ ) and children aged 5 – 9 years ( $R=0.87$ ), hence their effect could  
200 not be disentangled, whereas the correlation between physical contacts and household  
201 contacts was moderate ( $R=0.61$ , ranging from 0.47 to 0.70 between age groups). After age  
202 adjustment, physical contacts or contacts  $\geq 1$  hour remained most significantly associated  
203 with carriage, with a 6% increased risk (95%CI 2 – 10%) for each unit increase in the number  
204 of reported contacts (Table 1). We found good evidence that the number of household  
205 contacts increased the risk as well (Table 1). There was no confounding effect by other  
206 covariates and models were therefore only adjusted for age (Supporting Information S1).

207 An increase in physical, household and long ( $\geq 1h$ ) contacts, were also associated with an  
208 increased risk of ARS in univariable analysis. There was little or no evidence of a  
209 confounding effect of age, given the more constant prevalence of ARS across age groups  
210 (Table 1). Unlike pneumococcal carriage, however, the relative risk was more constant  
211 across types of contacts, and household contacts rather than physical contacts were most

212 strongly associated with a risk for ARS, with a risk increase of 9% (1 – 18%) for each  
213 additional reported contact.

214

215 **Table 1:** Relative risk of *S.pneumoniae* carriage and ARS by frequency of contact

Contact types	Mean contacts	Crude RR (95%CI)	Adjusted RR (95%CI)
<hr/>			
NP carriage	NP+	NP-	
All contacts	8.09	7.53	1.04 (1.00; 1.07)*
Physical contacts	6.83	6.09	1.13 (1.09; 1.17)*
Non-physical contacts	1.25	1.43	0.78 (0.72; 0.86)*
Household contacts	5.64	5.31	1.09 (1.03; 1.14)*
Non-household	2.22	2.45	0.98 (0.93; 1.04)
Contacts ≥ 1hour	7.50	6.56	1.08 (1.05; 1.12)*
Contacts <1h	2.38	2.62	0.81 (0.73; 0.92)*
ARS	ARS+	ARS-	
All contacts	8.18	7.63	1.07 (1.02; 1.12)*
Physical contacts	6.10	6.26	1.09 (1.03; 1.15)*
Non-physical contacts	2.07	1.36	1.00 (0.89; 1.10)
Household contacts	5.70	5.31	1.10 (1.02; 1.19)*
Non-household	2.47	2.31	1.04 (0.98; 1.11)
Contacts ≥ 1hour	7.33	6.74	1.07 (1.01; 1.13)*
Contacts <1 hour	3.14	2.58	1.05 (0.96; 1.15)

216 Legend: \*=associations significant at P<0.05. RR= Risk Ratio. CI= Confidence Interval. NP+=  
217 pneumococcal carrier, NP - = non-carrier, ARS+ = with ARS, ARS- = without ARS

218

219 Next, we analysed whether the number of casual contacts (i.e. contacts lasting < 5 minutes)  
220 was associated with either pneumococcal carriage or ARS. We found no evidence that the  
221 prevalence of pneumococcal carriage or the risk of ARS were associated with reporting  
222 higher levels of casual contacts, as shown in Table 2. Given the small numbers of individuals  
223 reporting ≥30 casual contacts, we pooled the 56 individuals reporting ≥20 casual contacts  
224 into one category.

225 In univariable analysis the risk of pneumococcal carriage was higher in the group of 78  
226 participants who did not know how many casual contacts they may have had. This was due  
227 to the higher proportion of children <5 years and aged 5 – 9 years in that group, however,  
228 after age-adjustment, there was no evidence that the risk of pneumococcal carriage was  
229 higher in that group. Similarly, a higher number of casual contacts was associated with  
230 increased risk of ARS, and with no confounding effect of age or other variables on the  
231 estimates.

232 We found no confounding effect on casual contacts on the relative risk of carriage or ARS as  
233 a function of the frequency of reported 'ordinary' contacts.

234

235

236

237

238 Table 2: Relative risk (RR) of pneumococcal carriage and ARS by level of reported number of  
239 daily casual contacts

Number of casual contacts	Number of individuals	Pneumococcal carriage		Acute Respiratory Symptoms (ARS)	
		Crude RR (95%CI)	Age-adjusted RR (95%CI)	Crude RR (95%CI)	Age-adjusted RR (95%CI)
0 – 9	315	ref	ref	ref	ref
10 – 19	119	0.82 (0.52; 1.27)	0.80 (0.55; 1.75)	1.02 (0.65; 1.49)	1.00 (0.65; 1.54)
≥20	56	0.81 (0.44; 1.49)	1.19 (0.70; 2.04)	0.51 (0.21; 1.27)	0.56 (0.22; 1.39)
Not known	76	1.52 (1.00; 2.31)*	1.07 (0.77; 1.49)	1.01 (0.45; 2.28)	0.88 (0.38; 2.00)

240 Legend: \* = associations significant at P<0.05. RR= Risk Ratio. CI= Confidence Interval

241

242 Finally, we explored the characteristics of age-specific mixing patterns by ARS or carriage  
243 status in greater detail, with a focus on physical contacts and household contacts. We  
244 computed the ratio of the mean number of contacts within and between age groups among  
245 carriers compared to non-carriers, and among individuals with ARS compared to those  
246 without. The average number of physical encounters within and between age groups  
247 tended to be higher for carriers than non-carriers in most instances, albeit with substantial  
248 uncertainty owing to small numbers (Figure 4). Carriers reported more contacts with  
249 children < 5 years, and particularly adult carriers who reported on average more than twice  
250 as many physical contacts with children under five than non-carriers (Figure 4). Given that  
251 most of such contacts occurred within the household, the effect of physical and household  
252 contacts with children <5 years was indistinguishable, whereas for older age groups, the

253 association with physical contact was stronger than that with household contacts (Figure 4).

254 Similar findings were seen for ARS, however, unlike for pneumococcal carriage,  
255 symptomatic adults did not have more contacts with young children than asymptomatic  
256 ones. Results based on absolute differences in the mean number of contacts rather than  
257 relative means are displayed in Figure S1 (Supporting Information), showing the similar  
258 associations than the reported ratios, but providing a quantified difference in mean number  
259 of contacts instead.

260

261 Figure 4: Ratio of the mean reported number of *physical* contacts between pneumococcal  
262 carriers and non-carriers (panel A) and individuals with ARS compared to those without ARS  
263 (Panel B). Legend: Matrices of the ratio of the mean number of contacts among  
264 pneumococcal carriers compared to non-carriers (Panel A) and individuals with ARS  
265 compared to non-ARS (panel B). The numbers represent the point estimate of the ratio and  
266 in brackets the 95% confidence bounds. Blue are cells with a point estimate <1 and in red  
267 with point estimates >1.

268

## 269 **Discussion**

270 Our study provided a unique opportunity to explore whether and how social contact  
271 patterns are associated with someone's risk of acute respiratory infection in a mostly rural  
272 East African setting. Our results show that people who tend to have more frequent close  
273 contacts are more likely to be pneumococcal carriers or to report acute respiratory  
274 symptoms, irrespective of their age. In contrast, we found that less intimate or short casual

275 contacts were not associated with someone's infection risk, suggesting that social contacts  
276 important for transmission are close interpersonal encounters.

277 Existing evidence on the association between contact patterns and risk of infection is  
278 mostly 'ecological', with very few studies based on individual-level data. Using data of  
279 influenza A/H1N1 seroconversion in Hong Kong coupled to social contact data in the same  
280 population, Kwok et al. [16] showed that age rather than social contact patterns were the  
281 main driver of the individual risk of infection in that setting, and further work by Kucharski  
282 et al. [17] on the same data further supported the finding that someone's risk of infection is  
283 related to the average mixing pattern within their age group rather than their reported  
284 number of contacts. However, the validation of the 'social contact hypothesis' with such  
285 data remains difficult, due to challenges in accounting for acquired immunity in the  
286 population, assumptions around stability of behaviour over relatively long time periods as  
287 well as challenges in capturing influenza infection events based on serological data only [16,  
288 17].

289 Studying nasopharyngeal carriage of *S.pneumoniae* as the main biological endpoint enabled  
290 us to address many of these issues, for several reasons. First, the high prevalence of carriage  
291 provided the statistical power required to study individually-matched acquisition risk in our  
292 study. Second, a person's individual behaviour is unlikely to be affected by carriage given  
293 that the vast majority of episodes remain asymptomatic, in contrast to symptomatic  
294 respiratory infections during which social behaviour might change as a result of illness [22].  
295 Most carriage episodes in our study were asymptomatic, and our decision to assess the risk  
296 for ARS and NP carriage separately stemmed from that observation. In addition, given that  
297 duration of carriage is relatively short – at most 3 months in young children and no longer

298 than a few weeks in adults [25] – , and as no difference in our survey was observed in the  
299 mean number of contacts between days of the weeks and between survey weeks, it is  
300 reasonable to assume that contact patterns measured on a given day reflect contact  
301 patterns around the time of pneumococcal acquisition. Finally, as colonisation results in  
302 weak protective immune responses and limited reduction in serotype-specific reacquisition  
303 risk [23, 26], and with over 90 circulating serotypes, natural acquired immunity is unlikely to  
304 distort the association between carriage and social contact patterns, unlike immunizing  
305 infection for which individuals with more frequent risk of infection due to their social  
306 contact patterns are also more likely to be immune.

307 In contrast, self-reported respiratory symptoms may be influenced by factors such as  
308 behaviour change in illness [22] or immunity. The definition itself also presents limitations;  
309 although respiratory viruses such as the Respiratory Syncytial Virus (RSV), adenovirus,  
310 parainfluenza and influenza viruses are likely to account for a large proportion of ARS cases  
311 [27], the definition was used as a non-specific proxy for acute respiratory infection and may  
312 have captured other infectious and non-infectious conditions. Notwithstanding such  
313 caveats, results for ARS were similar to those for nasopharyngeal carriage, but also showed  
314 a more consistent association with all types of non-casual contacts, rather than physical or  
315 contacts of long duration only. This suggests that the definition may have mostly captured  
316 acute infections, and also provides further evidence that the number of close interpersonal  
317 contacts, and particularly household contacts, plays a role in the transmission of acute  
318 respiratory pathogens.

319 One of the striking features of our analysis of the relative number of mean contacts  
320 between age groups is that adults colonised with *S.pneumoniae* reported more than twice

321 as many close contacts with children under five than non-carriers. This is in agreement with  
322 observational and modelling studies showing that pneumococcal carriage risk increases with  
323 household size and with the number of children <5 years in the household [28]. It also  
324 supports the finding that carriage acquisition in adults occurs mostly within the household  
325 as a result of contact with young children who are drivers of infection [29]. In contrast,  
326 however, we found that adults with ARS reported an equal or lower number of contacts  
327 with children <5 years than asymptomatic adults. This likely reflects the epidemiological  
328 differences between carriage and ARS, given the very high prevalence of carriage among <5  
329 year olds and the much less marked difference in age-specific prevalence of ARS across age  
330 groups, in addition to other potential factors such as acquired immunity among adults more  
331 frequently exposed to very young children. Some of the specific details of the contact  
332 patterns and differences between *S.pneumoniae* and ARS are harder to interpret, due in  
333 part to wide statistical uncertainty, particularly among adults in whom the number of  
334 carriers or symptomatic individuals is small. However, overall the results from our analysis  
335 support the general finding that increased close contacts are associated with higher risk of  
336 ARS or pneumococcal carriage.

337 There is still much debate about how respiratory pathogens are transmitted from person to  
338 person; whether through close direct or indirect physical contact, through large droplet  
339 transmission at close range (<1 meter), or through aerosolized particles floating over longer  
340 distances, particularly in poorly ventilated indoor settings [30, 31]. It is likely that many  
341 pathogens can be transmitted through a combination of these routes. Yet the contribution  
342 of each mechanism remains uncertain [30, 31]. It is generally assumed that the main  
343 transmission route of *S.pneumoniae* is through direct contact [32], as well as indirectly

344 through shared glasses or bottles [33]. Analogously, for other colonising bacteria such as  
345 *N.meningitidis*, close contact and intimate kissing are known risk factors among teenagers  
346 and young adults [34]. Similarly, it is believed that influenza and other respiratory viruses  
347 are primarily transmitted through direct contact or contact with large droplet transmission  
348 at close range rather than aerosolized particles [30].

349 Although our objective was not to demonstrate transmission, our findings strongly support  
350 that direct close interpersonal contact is an important mode of transmission for pathogenic  
351 bacteria of the nasopharynx as well as respiratory viruses, and strengthens the scientific  
352 evidence for public health measures based upon these assumptions, such as hand washing  
353 campaigns or chemoprophylaxis of close contacts of cases of meningococcal meningitis to  
354 name a few. The strong association of ARS with household and other close contacts,  
355 independently from physical touch, might suggest that other mechanisms such as indirect  
356 transmission through fomites or aerosol transmission may play a role. Elucidating the  
357 contribution of such factors in this context would be an important question for future  
358 research.

359 Our findings have also several implications for infectious disease research. Contact  
360 structures are central to transmission models, and appropriate assumptions about what  
361 type of contact drives infectious disease transmission are essential. Our results suggest that  
362 the parameterisation of transmission models of *S.pneumoniae* and similar pathogens using  
363 mixing matrices based on physical or another measure of *close* interpersonal contact would  
364 more likely capture relevant contact patterns than those based on any type of social  
365 encounter. This has also implications for the design of contact studies, particularly in low-  
366 income settings given the scarcity of published data currently available and the need to

367 collect additional data from many more settings [7, 8]. In contrast to diary-based  
368 approaches that have been used by many [2, 4, 7, 8], the study design here was relatively  
369 simple and only involved a single face-to-face interview. A drawback of such retrospective  
370 approach is the lack of detailed information about very short (i.e. ‘casual’) contacts, as such  
371 information was deemed unreliable – and this is further supported by evidence that short  
372 contacts tend to be inconsistently recorded even in prospective diary-based approached [1].  
373 However, given that very short contacts may not account for much of the transmission, as  
374 our study suggests, a more simple retrospective design is a potential attractive option in  
375 other settings where contact data are lacking, and in which data collection through more  
376 comprehensive diary-based approaches may be difficult to implement.

377

378 There are some additional limitations to our study. We were unable to explore other  
379 potential confounding factors, such as bed share, ventilation, indoor smoke or hand washing  
380 [35], and the contribution of such factors should be explored further. Moreover, it remains  
381 unclear to what extent our results can be generalised to any acute respiratory infection. For  
382 example, factors such as ventilation and airflow may be of particular importance for  
383 aerosolized transmission of pathogens such as mycobacteria [7], compared to influenza or  
384 *S.pneumoniae* [30, 31], and whether household or physical contacts reflect contact patterns  
385 important for aerosolized transmission remains uncertain [7]. Finally, while we found a  
386 strong association at the individual level, our study does not demonstrate causality.

387 However, in the absence of detailed longitudinal data on acquisition events between all  
388 individuals’ contacts —which would be challenging and possibly unrealistic to obtain— our  
389 findings provide consistent evidence of a ‘dose-response’ association at the individual level  
390 between close social encounters and acquisition risk for respiratory pathogens, and

391 therefore provides robust support both for the social contact hypothesis, and for research  
392 and policy work based upon this hypothesis.

393

394 **Materials and Methods**

395 Data collection

396 The study was conducted in Sheema North Sub-District (Sheema district, South-West  
397 Uganda) between January and March 2014. Sixty clusters were randomly selected from the  
398 215 villages and two small district towns (Kabwohe and Itendero) in the study area,  
399 proportionally to the population size of each village and town. In each cluster 11 or 12  
400 individuals were randomly sampled from different households to both answer questions  
401 about their social contacts and their history of respiratory illness in the last two weeks, as  
402 well as having a nasopharyngeal swab taken. A household was defined as the group of  
403 individuals living under the same roof and sharing the same kitchen on a daily basis.

404 For the social contact questionnaire, participants were first asked to list all the individuals  
405 with whom they had a two-way conversational contact lasting for ≥5 minutes during a  
406 period of approximately 24 hours prior to the survey day (from wake up the previous day  
407 until wake up on the survey day). Such encounters were defined as 'ordinary contacts'. For  
408 each reported ordinary contact, participants (or their parent/guardian) were asked to  
409 estimate the contact's age (or estimated age), how long the encounter lasted for and  
410 whether it involved skin-to-skin touch or utensils passed from mouth to mouth (either of  
411 those defining 'physical contacts'). For very short social encounters (<5 minutes), which

412 were defined as 'casual contacts' (e.g. seeing someone on the way, encounter in a shop  
413 etc.), participants were asked to estimate the number of encounters based on pre-defined  
414 categories (<10 contacts, 10-19 contacts, 20-29 contacts, ≥30 contacts), but not to provide  
415 further details about each contact.

416 Next, participants were asked about respiratory symptoms experienced in the two weeks  
417 prior to the survey, including any of the following: cough, runny nose, sneezing, sore throat,  
418 difficulty breathing.

419 Finally, after the interview was completed, a nasopharyngeal swab was taken from each  
420 participant. NP samples were collected, transported and analysed as per WHO guidelines  
421 [24]. NP swabs (flocked nylon swabs, COPAN, Italy) were inoculated in a skim milk tryptone-  
422 glucose-glycerol (STGG) medium, transported in cool boxes and frozen at the research  
423 laboratory at -20°C within 8 hours of collection. Specimens were inoculated onto a selective  
424 agar plate of 5 mg/L gentamicin-Columbia agar with 5% sheep blood and incubated at 37°C  
425 in 5% CO<sub>2</sub> atmosphere overnight. Pneumococcal identification was based on optochin  
426 susceptibility testing of all alpha-hemolytic colonies and bile solubility testing in case of  
427 intermediate susceptibility to optochin.

428 Statistical analysis

429 We first performed descriptive analyses, with age-specific probability weights to account for  
430 different inclusion probabilities by age at the design stage, and adjusted for the clustering  
431 by village through the use of clustered 'sandwich' variance estimators to account for  
432 possible correlation within each of the sixty clusters [36]. We explored social contact  
433 patterns within and between age groups through contact matrices. In such matrices we

434 report the mean number of ordinary contacts of participants in age group  $j$  with contacts

435 in age group  $i$  ( $m_{ij}$ ), adjusted for reciprocity, as in Melegaro et al. [15].

436 We then analysed whether and how the frequency distribution of contacts was associated

437 with pneumococcal carriage or self-reported ARS.

438 We modelled the effect of 'ordinary' contacts (defined as contacts  $\geq 5$  minutes long) on

439 carriage or respiratory symptoms as a function of contact frequency, through a Poisson

440 model with a robust variance estimator, and inclusion probability weights by age group. We

441 treated contacts as continuous variables, but assessed departure from linearity through

442 likely ratio tests and model comparisons of Bayesian Information Criterion (BIC). In

443 multivariable analysis, we considered for inclusion any covariate significantly associated

444 with the outcome at  $P < 0.10$  in univariable analysis. Model improvement was considered for

445 any decrease in the BIC. Further details are provided in the Supporting Information S1.

446 Next, we used the same analytical approach to assess whether carriage and ARS were

447 associated with the level of self-reported 'casual' contacts (i.e.  $< 5$  minutes long).

448 Finally, we explored differences in the social mixing matrices by status of pneumococcal

449 carriage and ARS. To do so, we computed the ratio of the mean number of reported

450 contacts by participants  $j$  with contacts  $i$  among carriers compared to non-carriers ( $R_{ij}^C$ ) or

451 symptomatic compared to asymptomatic individuals ( $R_{ij}^R$ ), such that  $R_{ij}^C = \frac{m_{ij}^C}{m_{ij}^{N-C}}$  and

452  $R_{ij}^R = \frac{m_{ij}^R}{m_{ij}^{N-R}}$ , where  $C$  = carriers,  $R$  = participants with ARS and  $N$  = the total number of

453 participants. Estimates were not adjusted for reciprocity given that subpopulations were

454 not closed (e.g. contacts of carriers may not be carriers), and that our aim was to compare  
455 reported number of contacts rather than calculate a contact matrix. The uncertainty in  
456 reported values and ratios was obtained through resampling techniques, drawing random  
457 samples from each  $m_{ij}$ , with the number of draws equal to the study population in each age  
458 group  $j$ . Ratios and uncertainty around them was obtained from the ratio of bootstrapped  
459 matrices.

460 The same approach as described above was used to compute the differences in  $m_{ij}$  with  
461  $D_{ij}^C = m_{ij}^C - m_{ij}^{N-C}$  and  $D_{ij}^R = m_{ij}^R - m_{ij}^{N-R}$ .

462 All analyses were performed in STATA 13.1 IC and R version 3.2.

463 **Ethics**

464 Approval was obtained from the Ethical review boards of the London School of Hygiene and  
465 Tropical Medicine, Médecins Sans Frontières, the Faculty of Medicine Research & Ethics  
466 Committee of the Mbarara University of Science and Technology, the Institutional Ethical  
467 Review Board of the MUST, and the Uganda National Council for Science and Technology.

468

469 **Acknowledgments**

470 We thank all the surveyors for their work and all the individuals who took part in this study,  
471 Rinah Arinaitwe for her help in managing the data entry, Dan Nyehangane for his support  
472 with the laboratory analyses, and Dr Yap Boum II for his help overseeing field activities. We

473 are also grateful to Prof Anthony Scott, Dr Fabienne Nackers and Prof Francesco Checchi for  
474 their valuable input in the design and protocol of the field surveys.

475 **Conflict of interest**

476 The authors declare they have no conflict of interest.

477 **Funding**

478 This work was supported by Médecins Sans Frontières International Office, Geneva,  
479 Switzerland. Epicentre received core funding from Médecins Sans Frontières. OLP was  
480 supported by the AXA Research Fund through a Doctoral research fellowship. The funders  
481 had no role in study design, data collection and analysis, decision to publish, or preparation  
482 of the manuscript.

483 **References**

- 484 1. Smieszek T, Burri EU, Scherzinger R, Scholz RW. Collecting close-contact social mixing data  
485 with contact diaries: reporting errors and biases. *Epidemiology and infection*. 2012;140(4):744-52.  
486 doi: 10.1017/S0950268811001130. PubMed PMID: 21733249.
- 487 2. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, et al. Social contacts and  
488 mixing patterns relevant to the spread of infectious diseases. *PLoS medicine*. 2008;5(3):e74. doi:  
489 10.1371/journal.pmed.0050074. PubMed PMID: 18366252; PubMed Central PMCID: PMC2270306.
- 490 3. Eubank S, Guclu H, Kumar VS, Marathe MV, Srinivasan A, Toroczkai Z, et al. Modelling  
491 disease outbreaks in realistic urban social networks. *Nature*. 2004;429(6988):180-4. doi:  
492 10.1038/nature02541. PubMed PMID: 15141212.
- 493 4. Wallinga J, Teunis P, Kretzschmar M. Using data on social contacts to estimate age-specific  
494 transmission parameters for respiratory-spread infectious agents. *American journal of epidemiology*.  
495 2006;164(10):936-44. doi: 10.1093/aje/kwj317. PubMed PMID: 16968863.
- 496 5. Horby P, Pham QT, Hens N, Nguyen TT, Le QM, Dang DT, et al. Social contact patterns in  
497 Vietnam and implications for the control of infectious diseases. *PloS one*. 2011;6(2):e16965. doi:  
498 10.1371/journal.pone.0016965. PubMed PMID: 21347264; PubMed Central PMCID: PMC3038933.
- 499 6. Grijalva CG, Goeyvaerts N, Verastegui H, Edwards KM, Gil AI, Lanata CF, et al. A household-  
500 based study of contact networks relevant for the spread of infectious diseases in the highlands of  
501 Peru. *PloS one*. 2015;10(3):e0118457. doi: 10.1371/journal.pone.0118457. PubMed PMID:  
502 25734772; PubMed Central PMCID: PMC4348542.
- 503 7. Johnstone-Robertson SP, Mark D, Morrow C, Middelkoop K, Chiswell M, Aquino LD, et al.  
504 Social mixing patterns within a South African township community: implications for respiratory

- 505 disease transmission and control. American journal of epidemiology. 2011;174(11):1246-55. doi:  
506 10.1093/aje/kwr251. PubMed PMID: 22071585; PubMed Central PMCID: PMC3224253.  
507 8. Kiti MC, Kinyanjui TM, Koech DC, Munywoki PK, Medley GF, Nokes DJ. Quantifying age-  
508 related rates of social contact using diaries in a rural coastal population of Kenya. PloS one.  
509 2014;9(8):e104786. doi: 10.1371/journal.pone.0104786. PubMed PMID: 25127257; PubMed Central  
510 PMCID: PMC4134222.  
511 9. Fu YC, Wang DW, Chuang JH. Representative contact diaries for modeling the spread of  
512 infectious diseases in Taiwan. PloS one. 2012;7(10):e45113. doi: 10.1371/journal.pone.0045113.  
513 PubMed PMID: 23056193; PubMed Central PMCID: PMC3463600.  
514 10. Read JM, Lessler J, Riley S, Wang S, Tan LJ, Kwok KO, et al. Social mixing patterns in rural and  
515 urban areas of southern China. Proceedings Biological sciences / The Royal Society.  
516 2014;281(1785):20140268. doi: 10.1098/rspb.2014.0268. PubMed PMID: 24789897; PubMed  
517 Central PMCID: PMC4024290.  
518 11. Edmunds WJ, O'Callaghan CJ, Nokes DJ. Who mixes with whom? A method to determine the  
519 contact patterns of adults that may lead to the spread of airborne infections. Proceedings Biological  
520 sciences / The Royal Society. 1997;264(1384):949-57. doi: 10.1098/rspb.1997.0131. PubMed PMID:  
521 9263464; PubMed Central PMCID: PMC1688546.  
522 12. Halloran ME, Ferguson NM, Eubank S, Longini IM, Jr., Cummings DA, Lewis B, et al. Modeling  
523 targeted layered containment of an influenza pandemic in the United States. Proceedings of the  
524 National Academy of Sciences of the United States of America. 2008;105(12):4639-44. doi:  
525 10.1073/pnas.0706849105. PubMed PMID: 18332436; PubMed Central PMCID: PMC2290797.  
526 13. Baguelin M, Flasche S, Camacho A, Demiris N, Miller E, Edmunds WJ. Assessing optimal  
527 target populations for influenza vaccination programmes: an evidence synthesis and modelling  
528 study. PLoS medicine. 2013;10(10):e1001527. doi: 10.1371/journal.pmed.1001527. PubMed PMID:  
529 24115913; PubMed Central PMCID: PMC3793005.  
530 14. Santermans E, Goeyvaerts N, Melegaro A, Edmunds WJ, Faes C, Aerts M, et al. The social  
531 contact hypothesis under the assumption of endemic equilibrium: Elucidating the transmission  
532 potential of VZV in Europe. Epidemics. 2015;11:14-23. doi: 10.1016/j.epidem.2014.12.005. PubMed  
533 PMID: 25979278.  
534 15. Melegaro A, Jit M, Gay N, Zagheni E, Edmunds WJ. What types of contacts are important for  
535 the spread of infections?: using contact survey data to explore European mixing patterns. Epidemics.  
536 2011;3(3-4):143-51. doi: 10.1016/j.epidem.2011.04.001. PubMed PMID: 22094337.  
537 16. Kwok KO, Cowling BJ, Wei VW, Wu KM, Read JM, Lessler J, et al. Social contacts and the  
538 locations in which they occur as risk factors for influenza infection. Proceedings Biological sciences /  
539 The Royal Society. 2014;281(1789):20140709. doi: 10.1098/rspb.2014.0709. PubMed PMID:  
540 25009062; PubMed Central PMCID: PMC4100506.  
541 17. Kucharski AJ, Kwok KO, Wei VW, Cowling BJ, Read JM, Lessler J, et al. The contribution of  
542 social behaviour to the transmission of influenza A in a human population. PLoS pathogens.  
543 2014;10(6):e1004206. doi: 10.1371/journal.ppat.1004206. PubMed PMID: 24968312; PubMed  
544 Central PMCID: PMC4072802.  
545 18. Rohani P, Zhong X, King AA. Contact network structure explains the changing epidemiology  
546 of pertussis. Science. 2010;330(6006):982-5. doi: 10.1126/science.1194134. PubMed PMID:  
547 21071671.  
548 19. O'Brien KL, Wolfson LJ, Watt JP, Henkle E, Deloria-Knoll M, McCall N, et al. Burden of disease  
549 caused by Streptococcus pneumoniae in children younger than 5 years: global estimates. Lancet.  
550 2009;374(9693):893-902. doi: 10.1016/S0140-6736(09)61204-6. PubMed PMID: 19748398.  
551 20. Black RE, Cousens S, Johnson HL, Lawn JE, Rudan I, Bassani DG, et al. Global, regional, and  
552 national causes of child mortality in 2008: a systematic analysis. Lancet. 2010;375(9730):1969-87.  
553 doi: 10.1016/S0140-6736(10)60549-1. PubMed PMID: 20466419.  
554 21. Cohen AL, Hyde TB, Verani J, Watkins M. Integrating pneumonia prevention and treatment  
555 interventions with immunization services in resource-poor countries. Bulletin of the World Health

- 556 Organization. 2012;90(4):289-94. doi: 10.2471/BLT.11.094029. PubMed PMID: 22511825; PubMed  
557 Central PMCID: PMC3324866.
- 558 22. Van Kerckhove K, Hens N, Edmunds WJ, Eames KT. The impact of illness on social networks:  
559 implications for transmission and control of influenza. *American journal of epidemiology*.  
560 2013;178(11):1655-62. doi: 10.1093/aje/kwt196. PubMed PMID: 24100954; PubMed Central PMCID:  
561 PMC3842903.
- 562 23. Prevaes SM, van Wamel WJ, de Vogel CP, Veenhoven RH, van Gils EJ, van Belkum A, et al.  
563 Nasopharyngeal colonization elicits antibody responses to staphylococcal and pneumococcal  
564 proteins that are not associated with a reduced risk of subsequent carriage. *Infection and immunity*.  
565 2012;80(6):2186-93. doi: 10.1128/IAI.00037-12. PubMed PMID: 22451514; PubMed Central PMCID:  
566 PMC3370583.
- 567 24. Satzke C, Turner P, Virolainen-Julkunen A, Adrian PV, Antonio M, Hare KM, et al. Standard  
568 method for detecting upper respiratory carriage of *Streptococcus pneumoniae*: updated  
569 recommendations from the World Health Organization Pneumococcal Carriage Working Group.  
570 *Vaccine*. 2013;32(1):165-79. doi: 10.1016/j.vaccine.2013.08.062. PubMed PMID: 24331112.
- 571 25. Abdullahi O, Nyiro J, Lewa P, Slack M, Scott JA. The descriptive epidemiology of  
572 *Streptococcus pneumoniae* and *Haemophilus influenzae* nasopharyngeal carriage in children and  
573 adults in Kilifi district, Kenya. *The Pediatric infectious disease journal*. 2008;27(1):59-64. doi:  
574 10.1097/INF.0b013e31814da70c. PubMed PMID: 18162940; PubMed Central PMCID: PMC2382474.
- 575 26. Lipsitch M, Abdullahi O, D'Amour A, Xie W, Weinberger DM, Tchetgen Tchetgen E, et al.  
576 Estimating rates of carriage acquisition and clearance and competitive ability for pneumococcal  
577 serotypes in Kenya with a Markov transition model. *Epidemiology*. 2012;23(4):510-9. doi:  
578 10.1097/EDE.0b013e31824f2f32. PubMed PMID: 22441543; PubMed Central PMCID: PMC3670084.
- 579 27. Berkley JA, Munywoki P, Ngama M, Kazungu S, Abwao J, Bett A, et al. Viral etiology of severe  
580 pneumonia among Kenyan infants and children. *Jama*. 2010;303(20):2051-7. doi:  
581 10.1001/jama.2010.675. PubMed PMID: 20501927; PubMed Central PMCID: PMC2968755.
- 582 28. Farida H, Severin JA, Gasem MH, Keuter M, Wahyono H, van den Broek P, et al.  
583 Nasopharyngeal carriage of *Streptococcus pneumoniae* in pneumonia-prone age groups in Semarang,  
584 Java Island, Indonesia. *PloS one*. 2014;9(1):e87431. doi: 10.1371/journal.pone.0087431. PubMed  
585 PMID: 24498104; PubMed Central PMCID: PMC3909120.
- 586 29. Melegaro A, Gay NJ, Medley GF. Estimating the transmission parameters of pneumococcal  
587 carriage in households. *Epidemiology and infection*. 2004;132(3):433-41. PubMed PMID: 15188713;  
588 PubMed Central PMCID: PMC2870123.
- 589 30. Brankston G, Gitterman L, Hirji Z, Lemieux C, Gardam M. Transmission of influenza A in  
590 human beings. *The Lancet Infectious diseases*. 2007;7(4):257-65. doi: 10.1016/S1473-  
591 3099(07)70029-4. PubMed PMID: 17376383.
- 592 31. In: Larson EL, Liverman CT, editors. *Preventing Transmission of Pandemic Influenza and*  
593 *Other Viral Respiratory Diseases: Personal Protective Equipment for Healthcare Personnel: Update*  
594 *2010*. Washington (DC)2011.
- 595 32. Ferreira DM, Neill DR, Bangert M, Gritzfeld JF, Green N, Wright AK, et al. Controlled human  
596 infection and rechallenge with *Streptococcus pneumoniae* reveals the protective efficacy of carriage  
597 in healthy adults. *American journal of respiratory and critical care medicine*. 2013;187(8):855-64.  
598 doi: 10.1164/rccm.201212-2277OC. PubMed PMID: 23370916; PubMed Central PMCID:  
599 PMC3707375.
- 600 33. Levine H, Balicer RD, Zarka S, Sela T, Rozhavski V, Cohen D, et al. Dynamics of pneumococcal  
601 acquisition and carriage in young adults during training in confined settings in Israel. *PloS one*.  
602 2012;7(10):e46491. doi: 10.1371/journal.pone.0046491. PubMed PMID: 23056322; PubMed Central  
603 PMCID: PMC3466294.
- 604 34. MacLennan J, Kafatos G, Neal K, Andrews N, Cameron JC, Roberts R, et al. Social behavior  
605 and meningococcal carriage in British teenagers. *Emerging infectious diseases*. 2006;12(6):950-7.  
606 PubMed PMID: 16707051; PubMed Central PMCID: PMC3373034.

- 607 35. Tumwesigire SG, Barton T. Environmental risk factors for acute respiratory infections among  
608 children of military personnel in Uganda. East African medical journal. 1995;72(5):290-4. PubMed  
609 PMID: 7555884.  
610 36. Rogers WH. Regression standard errors in clustered samples. Stata Technical Bulletin  
611 1993;13:19–23.  
612 .  
613

614 **Supporting Information Captions**

615 **S1 Text.** Analytical approach and model comparison

616  
617 **S1 Figure:** Difference in the mean number of *physical* contact between pneumococcal  
618 carriers and non-carriers (Panel A) and between individuals suffering and not suffering from  
619 respiratory symptoms (Panel B).Legend: Matrices of the mean difference between the mean  
620 number of contacts among pneumococcal carriers and non-carriers (panel A), and  
621 individuals with ARS compared to non-ARS (panel B). The numbers represent the point  
622 estimate and the lower and upper bound of the 95% credible interval are shown inside the  
623 brackets.

624

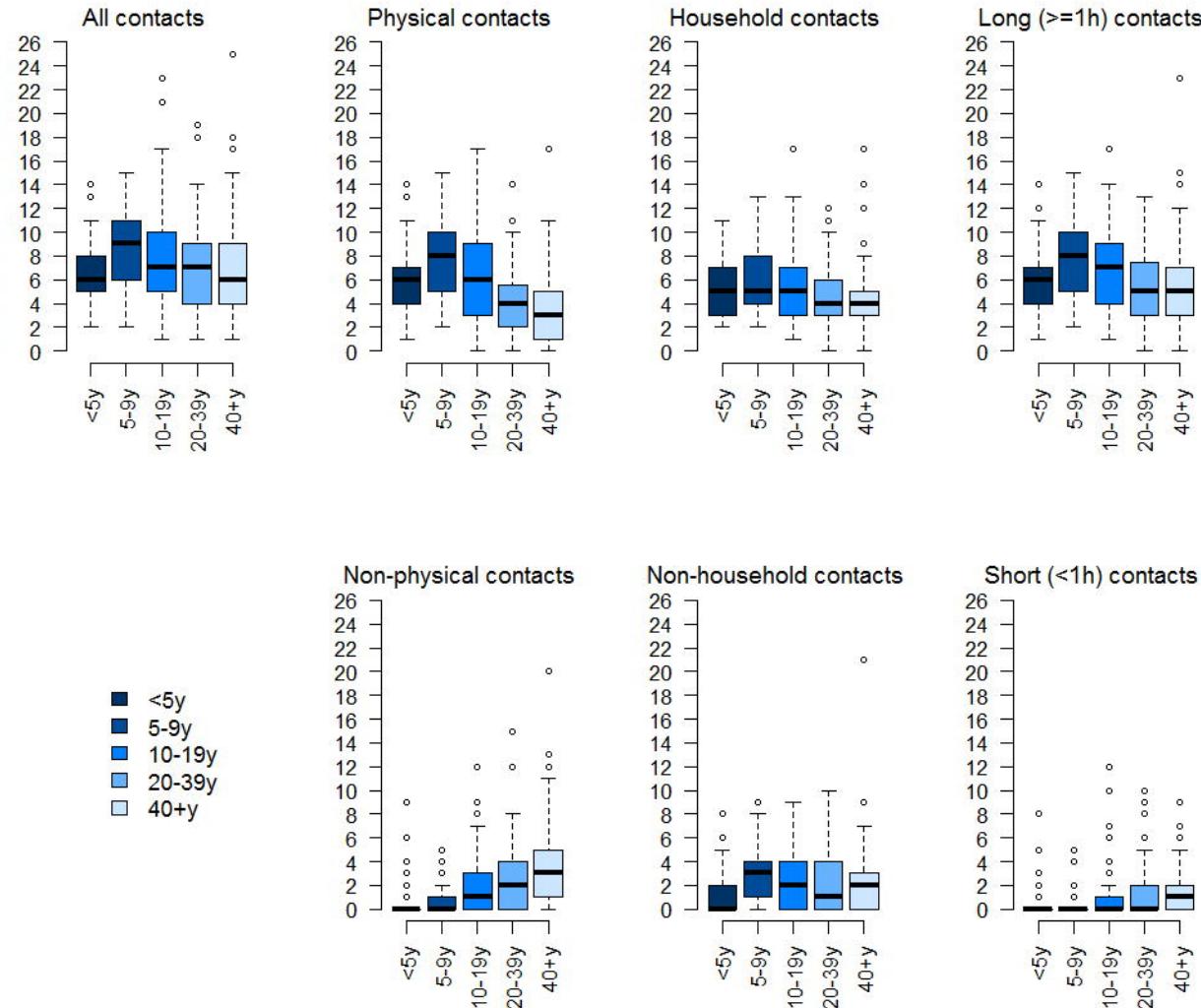
625

626 **S2 Text:** Data dictionary

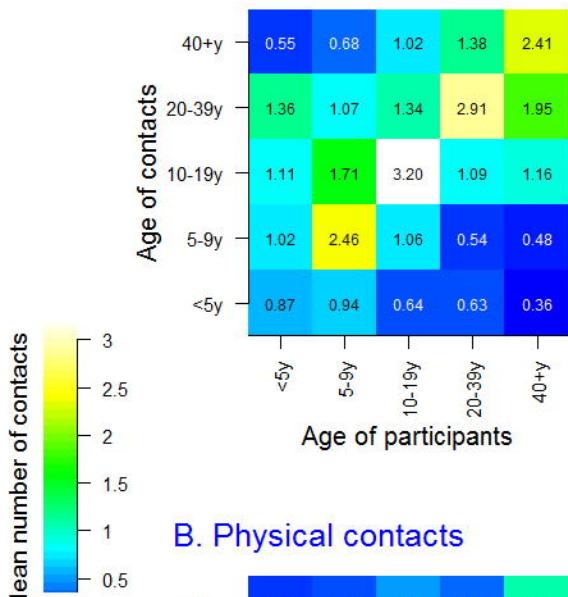
627

628 **S3 Data:** Dataset

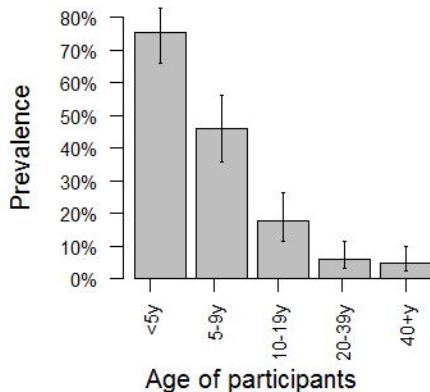
Number of reported contacts



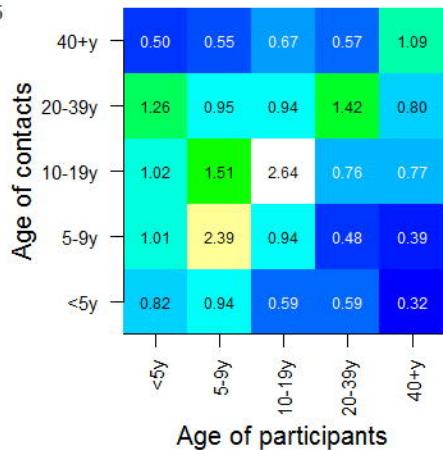
### A. All contacts



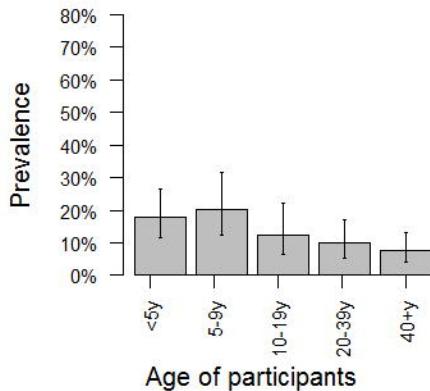
### C. Pneumococcal carriage prevalence



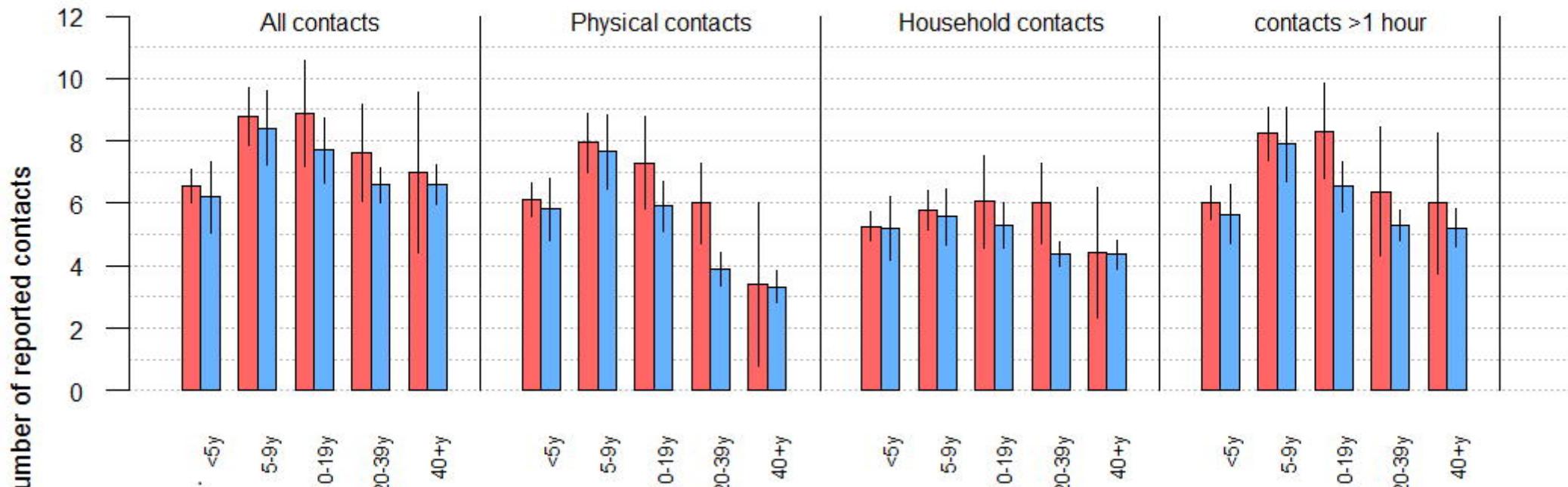
### B. Physical contacts



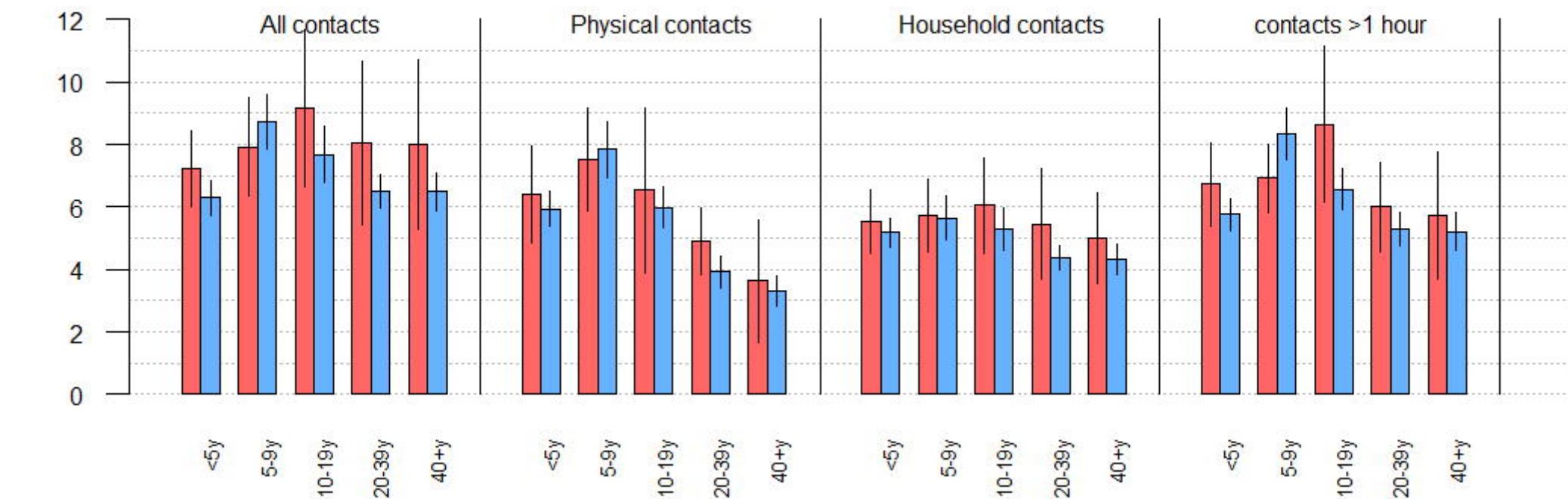
### D. Prevalence of respiratory symptoms



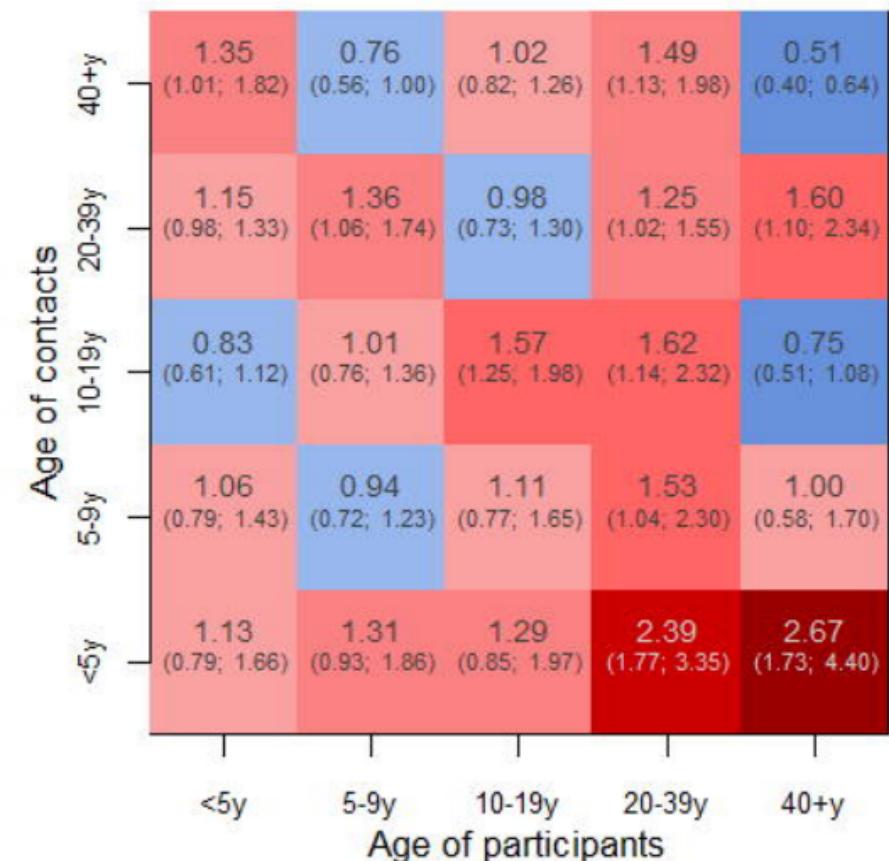
### A: Pneumococcal carriage



### B: ARS



### A. Pneumococcal carriage



### B. Respiratory symptoms

