nature communications

Article

https://doi.org/10.1038/s41467-025-64850-9

Characterizing social behavior relevant for infectious disease transmission in four low- and middle-income countries, 2021-2023

Received: 13 February 2025

Accepted: 26 September 2025

Published online: 30 October 2025

Check for updates

Kristin N. Nelson[1,13] ✉, Moses C. Kiti[1,13], Machi Shiiba[1], Charfudin Sacoor[2], Azucena Bardají[2,3], Ivalda Macicame[4], Edgar Jamisse[2], Corssino Tchavana[2], Americo José[4], Nilzio Cavele[2], Herberth Maldonado[5], Claudia Jarquin[5], H. María Ajsivinac[5], Venkata Raghava[6], Prasanna Samuel[6], Rajan Srinivasan[6], Momin Kazi[7], Raheel Allana[7], Sara S. Kim[1], Pragati V. Prasad[1], Dehao Chen[1], Carol Liu[1], Samuel M. Jenness[1], Noureen Ahmed[8], Obianuju Aguolu[9], Maria A. Sundaram[10], Inci Yildirim[11], Fauzia Malik[8], Alessia Melegaro[12], Benjamin A. Lopman[1,14] & Saad B. Omer[8,14]

Infectious diseases account for nearly half of all child mortality worldwide, with most of the burden concentrated in low and middle-income countries (LMIC). Person-to-person interactions, or 'contacts', facilitate the spread of respiratory and enteric pathogens. The number and nature of contacts likely vary across countries along with social and cultural norms, but few studies have compared behaviors across countries and none have done so with a focus on children. Here we present data from a population-based study conducted from 2021 to 2023 in Guatemala, India, Mozambique, and Pakistan. Across four countries, 5085 participants reported a total of 84,829 contacts across two days. Mean contact rates were highest among 10- to 19-year-olds except in Pakistan, where contacts were highest among 5- to 9-year-olds. Non-home locations which presented high risk for transmission were schools in India, workplaces in Pakistan, and 'other' social / leisure locations in Mozambique and Guatemala. Among children under 5 years of age, the proportion of contacts with non-household members was highest in Mozambique and lowest in India; most of these were reported at home. Contact patterns by age diverge from prior projections that are extrapolated from contact data from high-income countries, underscoring the value of local data collection.

Infectious diseases account for nearly a quarter of all mortality and half of all child mortality worldwide[1,2]. Half of deaths among children and adolescents caused by infectious diseases are due to enteric and lower respiratory infections[3].

Person-to-person interactions, or 'contacts', facilitate the spread of pathogens which are transmitted through the air and by direct deposition onto surfaces. Patterns of contact are likely to vary across countries along with social and cultural norms that influence the composition and size of households, and the interactions that occur in community settings such as schools, workplaces, and on public transport. Within countries, variation in the structure and importance of social institutions, employment opportunities, and population density in rural compared to urban communities may lead to differences in individual behaviors which may influence the how quickly and

A full list of affiliations appears at the end of the paper. ✉e-mail: knbratt@emory.edu

to whom infections are likely to spread[4,5]. Children under five years of age play important roles in the transmission of many pathogens[6–8] and are more vulnerable to severe disease from enteric and respiratory infections. Differences in caregiving practices and school attendance for young children that vary by setting may lead to different patterns of exposure during the vulnerable period of early childhood, shaping the epidemiology of enteric and respiratory infections in this age group[9].

Past efforts to collect data on person-to-person interactions across countries have focused primarily on high-income countries. The POLYMOD study collected data on contact patterns in eight European countries in the early 2000s[9]; more recently, the CoMix study documented changing behaviors in Europe during the COVID-19 pandemic[10]. In contrast, relatively little data has been available directly from low- and middle-income countries (LMIC), which make up 85% of the global population[11] and in which 89% of mortality from infectious disease among children occurs[12]. Social contact data collected from young children is available from some high-income countries but sparse in LMIC[13,14]. Among the data available from LMIC, key study design elements including the mode of data collection and sampling strategy have varied[15,16], making comparisons across studies difficult. A study conducted during the COVID-19 pandemic collected data on social contacts in nineteen African countries, but recruited and interviewed participants by phone, so its generalizability to the general population and to non-pandemic time periods is unclear[17].

The paucity of data from LMIC presents a challenge to disease modeling efforts, since estimating community-level impact of an intervention relies upon assumptions about the extent to which those directly receiving the intervention may interact with and therefore indirectly protect those who do not. Specifically, information on contact rates disaggregated by age, sex, and location of contact can be used to investigate the potential impact of targeted interventions, such as vaccination of a specific age group, or implemented in specific locations, such as social distancing or masking policies enacted in schools or workplaces. Efforts to generalize the POLYMOD data to other countries[18,19] have been cited over 3000 times since they were published, demonstrating the persistent use of these data in modeling efforts despite known differences between contact patterns between high-income and LMIC[20]. Models supported by robust data on human behavior may facilitate more accurate predictions of the intensity and length of an outbreak, a clearer understanding of entrenched patterns of transmission of endemic infections, and more precise forecasts of intervention effectiveness.

To fill these gaps, we conducted a study across eight sites in four countries: Guatemala, India, Mozambique, and Pakistan. We compared person-to-person interactions relevant for the spread of enteric and respiratory infections within and across sites, with a special focus on children under 5 years of age.

Results

Across four countries, 5085 participants ($n = 1141$ in Guatemala, $n = 1244$ in India, $n = 1363$ in Mozambique, and $n = 1337$ in Pakistan) reported a total of 84,829 contacts across both days (Table 1 and Supplementary Table 1). Of those participants, 29% ($n = 1480$) were under five years of age. The daily number of contacts was lowest (mean: 6.5 contacts, SD: 1.9) in Guatemala and highest (12.5 contacts, SD: 6.2) in Pakistan. In Mozambique and Pakistan, more contacts were reported in the rural than in the urban site (8.3 contacts compared to 5.4 contacts, $p < 1e{-5}$ in Mozambique and 15 contacts compared to 10.2 contacts, $p < 1e{-5}$, in Pakistan). In India and Guatemala, mean daily contacts were more similar in the urban and rural sites ($p = 0.040$ and 0.043, respectively). (Table 2) Comparisons of medians, rather than means, gave qualitatively similar results (Supplementary Table 2).

Mean daily contact rates were highest among 10- to 19-year-olds (7.0 contacts in Guatemala, 10.2 contacts in India, and 8.4 contacts in Mozambique) except in Pakistan, where mean contacts were highest among 5- to 9-year-olds (14.3 contacts). Qualitatively, patterns of contact by age were similar across countries and sites. In every site, assortative contacts, or contacts within the same age group, were highest among 10- to 19-year-olds (Fig. 1, Supplementary Fig. 1).

Household members accounted for one-third to one-half (36–44%) of contacts in all countries. In Pakistan, 86% of contacts were physical, compared to 66–73% of contacts in India, Mozambique, and Guatemala. Participants from the urban site in Pakistan reported an especially high proportion of physical contacts (93%). Outdoor contacts made up 18–22% of contacts in India, Guatemala, and Pakistan, but nearly half of all contacts (48%) in Mozambique (Supplementary Table 1).

In India, school contacts made up a larger proportion of children's contacts than in other countries, peaking at 64% among 10- to 14-year-olds. Work contacts made up the highest proportion of contacts in India (6%) and the lowest proportion in Guatemala (2%). Work contacts were higher among older adults in India (10% among 60+-year-olds) than in other countries. Contacts on transit were more common in Guatemala (11% of all contacts) and India (11%) than in Mozambique (2%) or Pakistan (6%) (Fig. 1). Overall, the location at which contacts were reported were not substantially different by site within a country (Supplementary Fig. 2). Contacts greatest in duration were reported at home (mean of 2.5 h) and typically were shortest on transit (mean 1.2 h) and in markets or when performing other essential activities (mean 52 min) (Supplementary Fig. 3). In all sites and age groups, exposure-hours of contact were greatest at home (Supplementary Fig. 4).

Among children under 5 years of age, daily contact rates were highest in Pakistan (daily mean of 9.2, 9.2 and 10.7 contacts for children <6 months, 6–11 months, and 1–4-year-olds, respectively) and lowest in Guatemala (mean of 6, 6.2, and 6.5 contacts, respectively). The proportion of contacts with non-household members was highest in Mozambique (69% among 6- to 11-month-olds) and lowest in India (45% among 6- to 11-month-olds). To identify locations posing the highest risk to young children, we calculated contact exposure-hours (summing the length of each contact) by location with non-household members. Contact exposure-hours with non-household members were highest in Mozambique and occurred primarily at home in all countries except Guatemala, where the most contacts among <6-month-olds and 6- to 11-month-olds were reported in Other locations (Fig. 2).

When we compared contact rates estimated from our study to those projected for each country using the POLYMOD data[18], we found that projections tended to overestimate absolute contact rates, especially among children. However, we did not find substantial differences in age-specific contact patterns overall. In all countries, our data showed more contacts at home and fewer contacts at work than in the projections (Fig. 3, Supplementary Fig. 5).

In a multivariable model including sex, household size, education level, and occupation, age remained a strong predictor of total daily contacts: school-age children and adolescents (5–9- or 10–19-year-olds) had the highest number of contacts in most sites. (Supplementary Tables 3–6) In the same model, sex was weakly associated with contact rates, except in Pakistan, where men reported a higher number of contacts than women in the urban site (OR: 1.52, 95% CI: 1.43–1.62) and a lower number in the rural site (OR: 0.82, 95% CI: 0.77, 0.88). In urban sites, participants that reported at least some college had more contacts than participants with primary school education (ORs ranging from 1.08 to 1.75). In a model with the same variables evaluating the relationship between number of non-household contacts reported and demographic characteristics, results were similar, but associations with age were generally weaker and differences by sex in Pakistan were more pronounced. In most sites, fewer non-household contacts were reported as household size increased.

We used a transmission model to simulate the introduction and spread of an infection under the age-structured contact patterns in each site. Cumulative incidence of infection by age group varied by country, with the highest overall incidence in rural Pakistan (96%) and

 https://doi.org/10.1038/s41467-025-64850-9

Table 1 | Characteristics of study participants by site

	Guatemala				India				Mozambique				Pakistan			
	Urban		Rural		Urban		Rural		Urban		Rural		Urban		Rural	
N	566		575		622		622		687		676		679		658	
	n	(%, col)	*n*	(%, col)	*n*	(%, col)	*n*	(%, col)	*n*	(%, col)	*n*	(%, col)	*n*	(%, col)	*n*	(%, col)
Age																
<6mo	52	9.2%	59	10.3%	55	8.8%	58	9.3%	66	9.6%	62	9.2%	69	10.2%	64	9.7%
6–11mo	58	10.2%	45	7.8%	55	8.8%	54	8.7%	64	9.3%	82	12.1%	68	10.0%	62	9.4%
1–4 y	63	11.3%	55	9.6%	65	10.5%	61	9.8%	72	10.5%	63	9.3%	66	9.7%	62	9.4%
5–9 y	52	9.2%	65	11.3%	63	10.1%	63	10.1%	58	8.4%	64	9.5%	63	9.3%	66	10.0%
10–19 y	102	18.0%	100	17.4%	132	21.2%	127	20.4%	124	18.0%	125	18.5%	130	19.1%	128	19.5%
20–29 y	61	10.8%	64	11.1%	63	10.1%	62	10.0%	61	8.9%	64	9.5%	66	9.7%	63	9.6%
30–39 y	47	8.3%	62	10.8%	63	10.1%	63	10.1%	61	8.9%	64	9.5%	74	10.9%	63	9.6%
40–59 y	64	11.3%	61	10.6%	64	10.3%	71	11.4%	120	17.5%	89	13.2%	67	9.9%	79	12.0%
60+ y	67	11.8%	64	11.1%	62	10.0%	63	10.1%	61	8.9%	63	9.3%	70	10.3%	63	9.6%
Missing	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	6	0.9%	8	1.2%
Sex																
Female	363	64.0%	372	64.7%	326	52.4%	317	51.0%	334	48.6%	332	49.1%	334	49.2%	322	48.9%
Male	202	35.8%	203	35.3%	296	47.6%	305	49.0%	353	51.4%	343	50.7%	343	50.5%	331	50.3%
Missing	1	0.2%	0	0.0%	0	0%	0	0%	0	0.0%	1	0.1%	2	0.3%	5	0.8%
Highest education completed[a]																
Currently in school	128	32.6%	138	33.2%	163	36.5%	176	39.2%	195	40.2%	173	36.9%	83	17.4%	107	22.8%
No formal education	113	28.8%	133	32.0%	46	10.3%	42	9.4%	14	2.9%	65	13.9%	188	39.5%	260	55.3%
Primary	73	18.6%	90	21.6%	41	9.2%	41	9.1%	152	31.3%	141	30.1%	50	10.5%	71	15.1%
Secondary	39	9.9%	28	6.7%	155	34.7%	137	30.5%	99	20.4%	64	13.6%	152	31.9%	29	6.2%
Some college or higher	19	4.8%	3	0.7%	33	7.4%	45	10.0%	4	0.8%	4	0.9%	2	0.4%	3	0.6%
Missing	21	5.3%	24	5.8%	9	2.0%	8	1.8%	21	4.3%	22	4.7%	1	0.2%	0	0.0%
Occupation[b]																
Student	98	28.7%	92	26.2%	112	29.2%	121	31.3%	132	19.2%	115	28.4%	67	16.2%	48	11.9%
Semiprofessional/ professional	24	7.0%	4	1.1%	20	5.2%	12	3.1%	56	8.2%	27	6.7%	48	11.6%	6	1.5%
Semiskilled/skilled laborer	81	23.8%	53	15.1%	70	18.2%	85	22.0%	121	17.6%	95	23.5%	65	15.7%	157	38.9%
Other	1	0.3%	7	2.0%	34	8.9%	42	10.9%	14	2.0%	58	14.3%	29	7.0%	5	1.2%
Unemployed out-side home	137	40.2%	195	55.6%	148	38.5%	126	32.6%	104	15.1%	110	27.2%	204	49.4%	188	46.5%
Household size (categories)																
0–2	148	26.3%	145	25.2%	82	13.2%	78	12.5%	56	8.2%	70	10.3%	27	4.0%	8	1.2%
3–5	356	62.8%	348	60.5%	393	63.1%	383	61.6%	273	39.7%	277	41.0%	386	56.8%	184	28.0%
6–10	58	10.2%	81	14.1%	142	22.8%	155	24.9%	261	38.0%	249	36.8%	259	38.1%	347	52.7%
11+	3	0.5%	1	0.2%	5	0.8%	6	1.0%	32	4.7%	31	4.6%	7	1.0%	119	18.1%
Missing	1	0.2%	0	0.0%	0	0.0%	0	0.0%	65	9.5%	49	7.2%	0	0.0%	0	0.0%
Median household size (IQR)	3 (2, 4)		3 (2, 5)		4 (4, 5)		5 (4, 6)		5 (4, 7)		5 (4, 7)		5 (4, 6)		7 (5, 10)	
Generations in household																
1	126	22.2%	121	21.0%	75	12.1%	73	11.7%	158	23.0%	145	21.4%	18	2.7%	21	3.2%
2	405	71.4%	422	73.4%	363	58.4%	310	49.8%	308	44.8%	322	47.6%	610	89.8%	518	78.7%
3	35	6.2%	32	5.6%	184	29.6%	239	38.4%	211	30.7%	208	30.8%	50	7.4%	118	17.9%
Missing	0	0.2%	0	0.0%	0	0.0%	0	0.0%	10	1.5%	1	0.1%	1	0.1%	1	0.2%
Multi-family household																
No	503	88.7%	521	90.6%	523	84.1%	495	79.6%	515	75.0%	478	70.7%	660	97.2%	557	84.7%
Yes	63	11.1%	54	9.4%	99	15.9%	127	26.0%	162	23.6%	197	29.1%	18	2.7%	100	15.2%
Missing	0	0.2%	0	0.0%	0	0.0%	0	0.0%	10	1.5%	1	0.1%	1	0.1%	1	0.2%

[a]Among participants over 5 years of age; in Mozambique, missing category includes some 5- to 9-year-olds who have not yet entered school.

[b]Among participants 10 years of age and older.

https://doi.org/10.1038/s41467-025-64850-9

Table 2 | Mean daily contacts by participant characteristics by country and site

	Guatemala		India		Mozambique		Pakistan	
Overall, mean (SD)	6.5 (1.9)		7.9 (3.1)		6.8 (4.2)		12.5 (6.2)	
	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
	6.4 (1.7)	6.6 (2.0)	8.1 (2.9)	7.7 (3.2)	5.4 (3.0)	8.3 (4.7)	10.2 (5.4)	15.0 (6.0)
Age								
<6mo	6 (1.4)	6.1 (1.4)	6.7 (1.7)	6.2 (1.8)	3 (2.0)	6.3 (3.1)	7 (2.1)	11.6 (5.7)
6–11mo	6.3 (1.2)	6 (0.9)	7.4 (2.0)	6.6 (2.0)	4.4 (2.4)	7.2 (3.5)	6.7 (2.2)	11.8 (5.2)
1–4 y	6.5 (1.8)	6.7 (1.3)	8 (2.3)	8.3 (3.0)	5.4 (3.7)	7.3 (3.1)	6.6 (2.0)	15 (6.7)
5–9 y	6.5 (1.7)	7.1 (1.7)	9.8 (3.3)	10.1 (3.6)	5.1 (3.4)	8.4 (4.9)	9.5 (3.7)	18.8 (5.7)
10–19 y	6.8 (1.5)	7.2 (1.9)	9.7 (2.5)	10.7 (3.7)	7.2 (3.4)	9.6 (4.7)	10.8 (4.6)	16 (5.9)
20–29 y	6.8 (2.1)	6.6 (2.6)	7.9 (2.7)	7.8 (3.1)	5 (2.0)	8.8 (4.7)	10.2 (5.6)	13.3 (4.9)
30–39 y	6.1 (1.6)	6.4 (1.8)	7.8 (2.9)	7.7 (2.8)	5.3 (2.8)	7.8 (4.7)	13.2 (7.4)	15.2 (6.2)
40–59 y	6.2 (1.7)	6.7 (1.9)	7.8 (2.8)	6.9 (2.3)	4.7 (2.7)	8.4 (5.3)	9.8 (6.0)	14 (5.8)
60+ y	5.3 (1.5)	5.8 (1.7)	6.7 (3.0)	5.9 (2.3)	4 (2.5)	5.7 (4.1)	10.6 (4.3)	14.8 (5.4)
Sex								
Female	6.5 (1.7)	6.7 (2.1)	8.0 (3.1)	7.5 (3.3)	5.3 (2.9)	8.2 (4.8)	6.9 (2.3)	16.2 (6.4)
Male	6.3 (1.7)	6.6 (1.7)	8.3 (2.8)	7.8 (3.0)	5.5 (3.1)	8.4 (4.7)	13.3 (5.6)	13.7 (5.3)
Household size								
0–2	5.5 (1.6)	5.6 (1.5)	6.2 (2.5)	5.9 (2.8)	4.1 (2.5)	6.9 (3.9)	8.5 (6.1)	12.2 (4.9)
3–5	6.4 (1.4)	6.5 (1.5)	8.0 (2.7)	7.5 (3.2)	5.1 (2.7)	8.0 (4.8)	9.6 (5.7)	13.5 (6.0)
6–10	8.3 (1.8)	8.7 (2.4)	9.4 (2.9)	9.0 (2.7)	5.7 (3.1)	9.0 (5.1)	11.1 (4.6)	15.0 (5.9)
11+	10.7 (5.2)	–[a]	14.6 (2.8)	13.4 (0.6)	7.2 (3.5)	9.1 (3.5)	13.6 (2.3)	17.4 (5.6)
Highest education level attained[b]								
Currently enrolled in school	6.8 (1.7)	7.2 (2.0)	9.9 (2.7)	10.7 (3.7)	6.1 (3.3)	8.8 (4.7)	11.6 (4.5)	17.3 (5.8)
No formal education	6.2 (1.5)	6.5 (1.8)	7.0 (3.1)	6.4 (2.5)	4.3 (2.9)	6.9 (3.8)	8.8 (4.4)	14.4 (6.1)
Primary school	6.3 (1.6)	6.2 (1.6)	7.7 (2.4)	5.6 (2.0)	4.9 (2.6)	7.9 (4.9)	8.7 (5.0)	15.1 (4.9)
Secondary school	6.1 (1.7)	6.2 (3.0)	7.8 (2.9)	7.5 (2.6)	5.1 (2.5)	9.5 (5.0)	11.8 (6.3)	12.0 (4.9)
Some college/higher education	7.6 (3.0)	8.4 (3.5)	7.9 (2.7)	7.4 (2.4)	8.0 (4.1)	7.3 (4.0)	19.0 (14.5)	22.0 (5.8)
Occupation[c]								
Student	6.7 (1.6)	7.2 (1.9)	9.8 (2.7)	10.7 (3.8)	6.2 (3.3)	9.2 (4.7)	10.7 (4.2)	17.7 (5.9)
Semiprofessional/professional	5.9 (1.6)	8.3 (3.8)	8.6 (3.2)	8.7 (3.1)	5.3 (2.3)	8.7 (4.7)	16.4 (4.9)	14.4 (4.6)
Skilled/semiskilled	6.9 (2.2)	6.5 (1.8)	8.4 (2.8)	7.6 (2.5)	4.6 (2.2)	8.5 (5.4)	13.5 (6.8)	13.5 (5.1)
Unemployed outside home	6.0 (1.5)	6.4 (2.1)	7.1 (2.8)	6.2 (2.3)	5.0 (3.1)	6.9 (4.4)	8.1 (4.4)	15.5 (6.2)
Other	–[a]	5.9 (2.3)	7.9 (2.9)	6.9 (2.5)	9.1 (3.4)	9.0 (4.8)	8.4 (2.3)	12.2 (3.2)
Day of week								
Weekday	6.5 (1.9)	6.8 (2.1)	8.3 (3.3)	7.8 (3.5)	6.0 (3.8)	8.6 (5.7)	10.1 (5.6)	14.7 (6.1)
Weekend	6.3 (1.7)	6.3 (2.0)	7.6 (2.6)	7.5 (3.2)	5.6 (3.4)	8.5 (5.7)	10.5 (5.4)	15.7 (6.6)

[a]$n = 1$.
[b]Among participants over 5 years of age; 'Missing' category excluded.
[c]Among participants 10 years of age and older.

the lowest in urban Mozambique (41%). Cumulative incidence was higher in the rural compared to the urban sites in Mozambique and Pakistan, and more similar across sites in Guatemala and India. Cumulative incidence was highest among adolescents (10- to 19-year-olds) in Mozambique but highest among adults in all other sites (Fig. 4, Supplementary Fig. 6). Cumulative incidence under projections from Prem et al.[18] was higher than under GlobalMix contact patterns across all sites and age groups except in rural Pakistan. Likewise, Prem et al. contact patterns resulted in earlier epidemic peaks and more volatile effective reproduction numbers, indicating more rapid spread and resolution of the outbreak than under GlobalMix contact patterns. (Supplementary Figs. 7, 8)

Discussion

We describe social behaviors relevant for transmission of enteric and respiratory infections in four low- and middle-income countries across three continents, with an emphasis on children under 5 years of age. Mean daily contact rates were highest among school-aged children in every country and tended to be higher in rural sites. Important non-home locations for contact varied across countries. Among children under 5 years of age, contacts with non-household members occurred primarily at home and 'other' locations. Overall, we show that patterns of social interaction are broadly similar across countries, but key differences, including within-country rural-urban variation and the relative importance of non-home locations, could have implications for transmission dynamics of enteric and respiratory infections and the effectiveness of interventions aiming to limit their spread.

Daily contact rates reported in this study are within range of some reported in the UK[21] and Zimbabwe[22], but overall lower than previous estimates. In a recent review and meta-analysis, the mean number of contacts per day reported across studies conducted from 2005 to 2018 in 22 countries, was 14.5 contacts[20]. It is notable that our estimates are

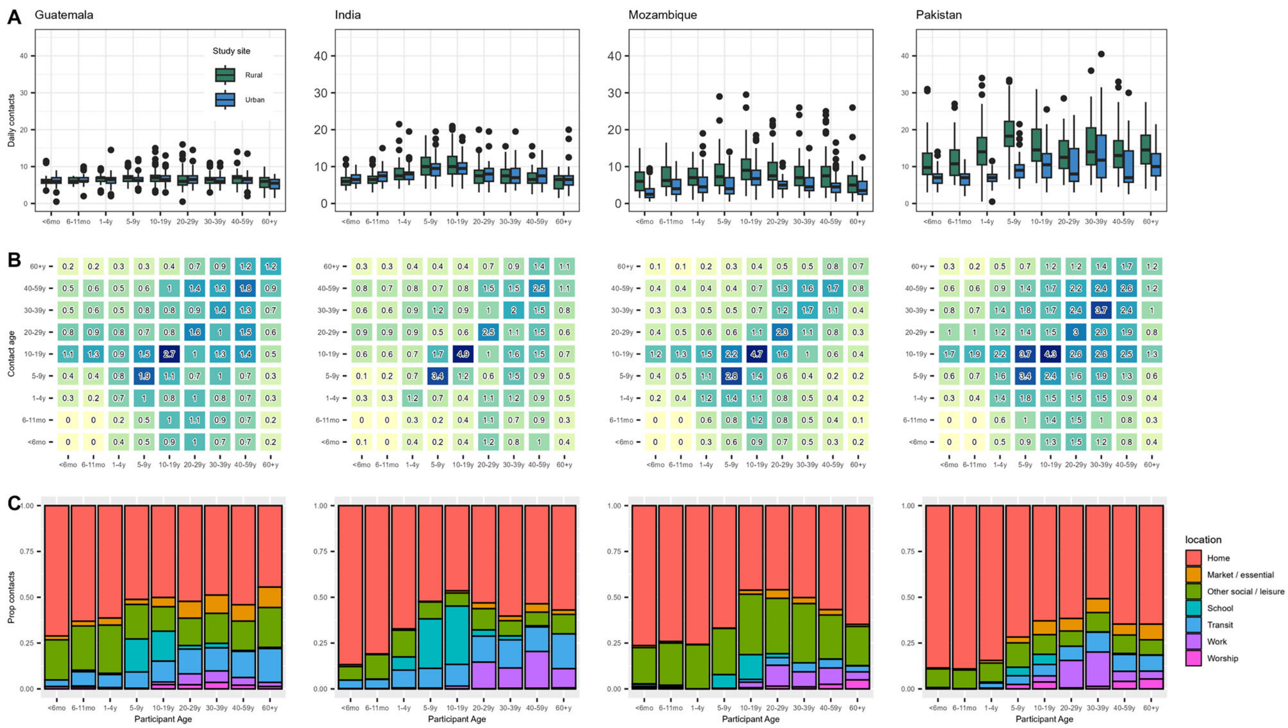

Fig. 1 | Daily contact rates by participant age and site, contact age, and location of contact. Patterns of contact by country and age group. Panel **A** shows the number of contacts reported on day 1 by age group of participants in rural and urban sites. The center line of the box shows the median, and edges of the box are the 25th percentile (Q1) and 75th percentile (Q3). The minima of the whisker extends to Q1–1.5 *IQR, and the maxima of the whisker extends to Q3 + 1.5*IQR. Dots indicate outliers that exceeded minima or maxima of the whiskers. The number of participants is n = 1143 in Mozambique (Rural: n = 576, Urban: n = 567), n = 1244 in India (Rural: n = 622, Urban: n = 622), n = 1363 in Guatemala (Rural: n = 676, Urban: n = 687), n = 1323 in Pakistan (Rural: n = 650, Urban: n = 673). Panel **B** shows contact rates by participant age and contact age. Contact rates are made symmetric to adjust for differences in the size of age groups included and lack of reciprocity in contact reporting. Colors in each cell reflect the mean number of contacts reported between age groups, with higher numbers in dark blue and lower numbers in yellow. (The color scale is country-specific to visually demonstrate patterns.) Panel **C** shows the proportion of contact occurring in each location by age group. Contacts for which the location was not reported are not shown.

low despite using prospective data collection and use of paper questionnaires, which have been shown to more fully ascertain contacts than other survey methods[23]. The timing of our study may partially account for fewer reported contacts, given the protracted impact of the COVID-19 pandemic on social behavior. Mozambique and Guatemala were likely most affected given the overlap of pandemic mitigation measures with data collection periods in 2021. However, despite official changes in pandemic mitigation policies that occurred during these periods, we show that the number of contacts reported was relatively stable through multiple months of data collection in each site. (Supplementary Figs. 9–12) This is consistent with anecdotal reports from field staff that while some policies to limit dense public gatherings were still in place through 2021, these restrictions were no longer meaningfully impacting daily interactions of most community members. Lower rates of contact could also stem from our household-based method of recruitment, which could have led us to recruit individuals who were more likely to be unemployed or otherwise less engaged in communal activities. Finally, it is also possible that contact rates have truly declined in recent years, given the widespread use of mobile smartphones and other technologies that have weakened social connections and reduced the frequency of in-person interactions.

Variations in the nature and number of contacts by country may reflect social, cultural, economic and even climactic differences. While the proportion of physical contacts reported in our study overall was similar to that reported in other low- and middle-income countries[20], we report a particularly high proportion of physical contacts in Pakistan. This likely reflects customs of greeting friends and acquaintances that often involve touch, especially among men. (Men in Pakistan also reported substantially more contacts than women, which may reflect gendered religious and social norms that govern common interaction.) In Mozambique, more interactions were reported outdoors than in other countries. This may reflect the influence of climate, which in southern Mozambique features mild temperatures that fluctuate little through the year, and differences in the built environment that encourage congregating outside. All of these differences may influence the speed and patterns of spread of infections and can be used to better understand behaviors to target for intervention that may vary across settings.

In two of four countries, rural contact rates were substantially higher than in urban areas, a counterintuitive result that may reflect the relative closeness of community structure in rural areas despite lower population density. Our findings are generally consistent with other studies, some of which show higher contact rates in rural vs peri-urban participants[16,24], and others[4,25] which do not show differences. Understanding variation in contact patterns in rural and urban settings is particularly important when modeling the spread of new infections, which are often initially introduced into well-connected urban centers and diffuse more slowly to rural areas.

While we found similar rates of contact as reported previously for infants and young children[26,27], our study provides additional detail on changes of children's contact patterns from birth to age 5, a critical period of vulnerability to infectious disease. Across countries, the majority of children's contacts are reported at home. However, the proportion of contacts occurring at home falls steadily from birth to age 5 in Guatemala and India and remains more stable in Mozambique and Pakistan. Patterns in biological susceptibility to infection change during this period, as maternally derived immunity wanes and the

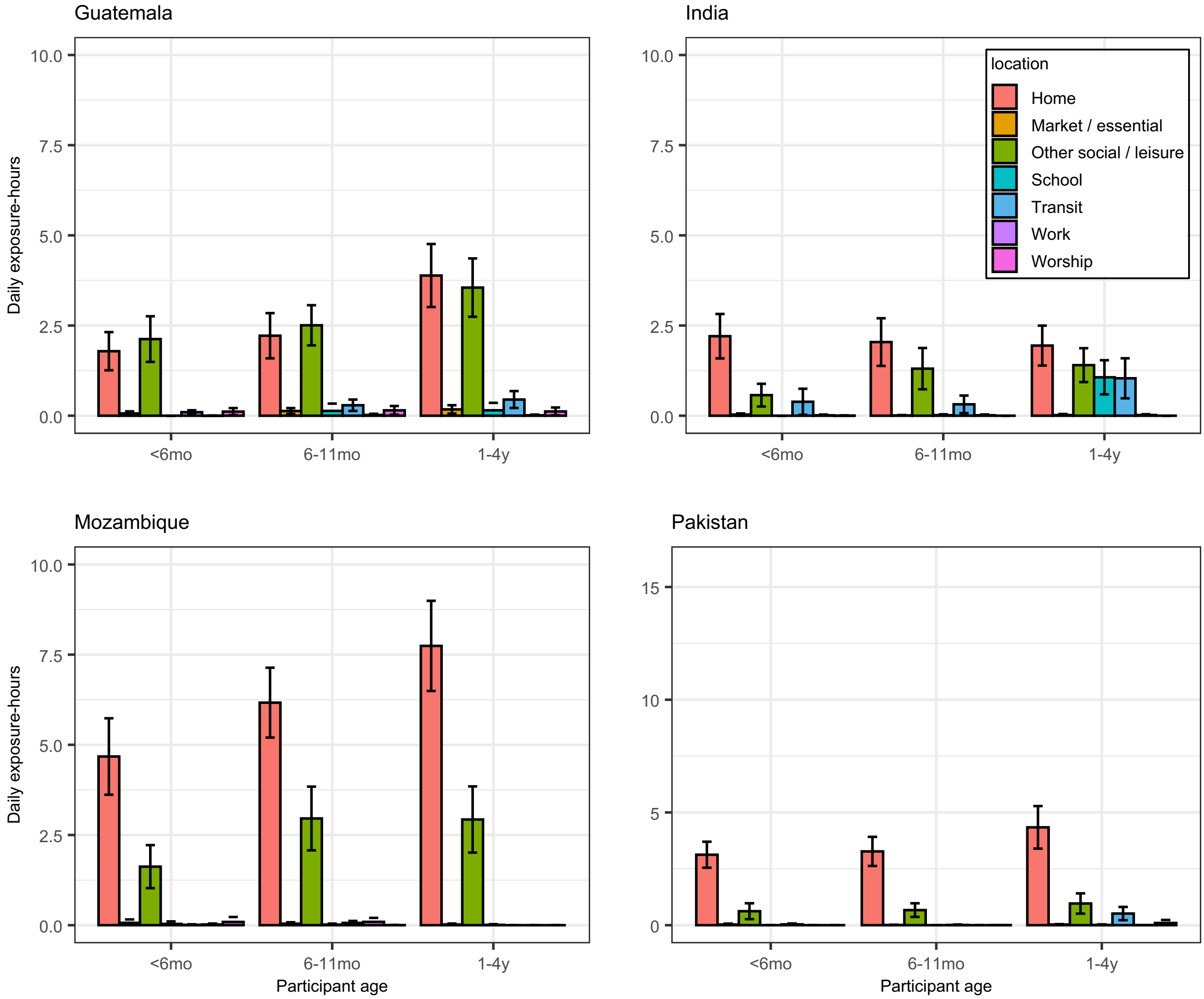

Fig. 2 | Mean daily exposure-hours of contact with non-household members by age group and location among children under 5 years of age. Mean daily exposure-hours were calculated by summing the product of the duration of contact and the number of contacts reported in each location. Contacts for which the location was not reported are not shown. Error bars show 95% confidence intervals of the mean. The number of participants in each age group in each country is as follows; $n = 111, 103, 118$ in Guatemala, $n = 113, 109, 126$ in India, $n = 128, 146, 135$ in Mozambique, $n = 133, 130, 128$ in Pakistan for <6mo, 6–11mo, and 1–4 y groups, respectively.

innate immune system matures[28]. Our data demonstrate the social changes that occur in parallel and shape exposure to pathogens in early childhood, and thereby risk of infection and illness in this age group[28]. Importantly, our study also differentiates contacts that occur at home and those that occur with household members. For children under 5 years of age, the two most important locations of exposure to non-household contacts were in their own homes and in Other locations. These patterns likely reflect sociocultural differences in caregiving practices for young children across countries, which can shape transmission patterns among children or between children and caregivers.

Similar to previous surveys[18], we find the highest rates of contact among school-aged children and the highest rates of assortative mixing among adolescents (10–19-year-olds). The relative importance of contacts in schools across settings can inform context-specific predictions of the effectiveness of school closings and other policies that may limit school attendance to prevent spread of infection, especially given the great social costs of such interventions on children and communities.

While POLYMOD-based projections show broadly similar patterns of contact by age across the four countries; our estimates show differences in age patterns by country. POLYMOD-based projections tend to overestimate contacts at work and underestimate contacts at home, consistent with previous literature that has shown that people in LMIC have more contacts at home and fewer contacts at work than in high-income countries[20]. When we used our empirical data to model the spread of an emerging pathogen in each site, we found that cumulative incidence varied substantially, with sites reporting lower contact rates displaying lower incidence. In Pakistan, qualitative differences in contact patterns by age led to higher incidence among young children in the rural site as compared to the urban site. Compared to the epidemic simulated using POLYMOD-based projections of contact patterns, we saw qualitatively different patterns by age in cumulative incidence of infection, with substantially more variation by age in simulations using the GlobalMix data and a generally slower progression of the epidemic. Though this modeling exercise was based on simple assumptions which would need to be adjusted to better capture the more complex natural

https://doi.org/10.1038/s41467-025-64850-9

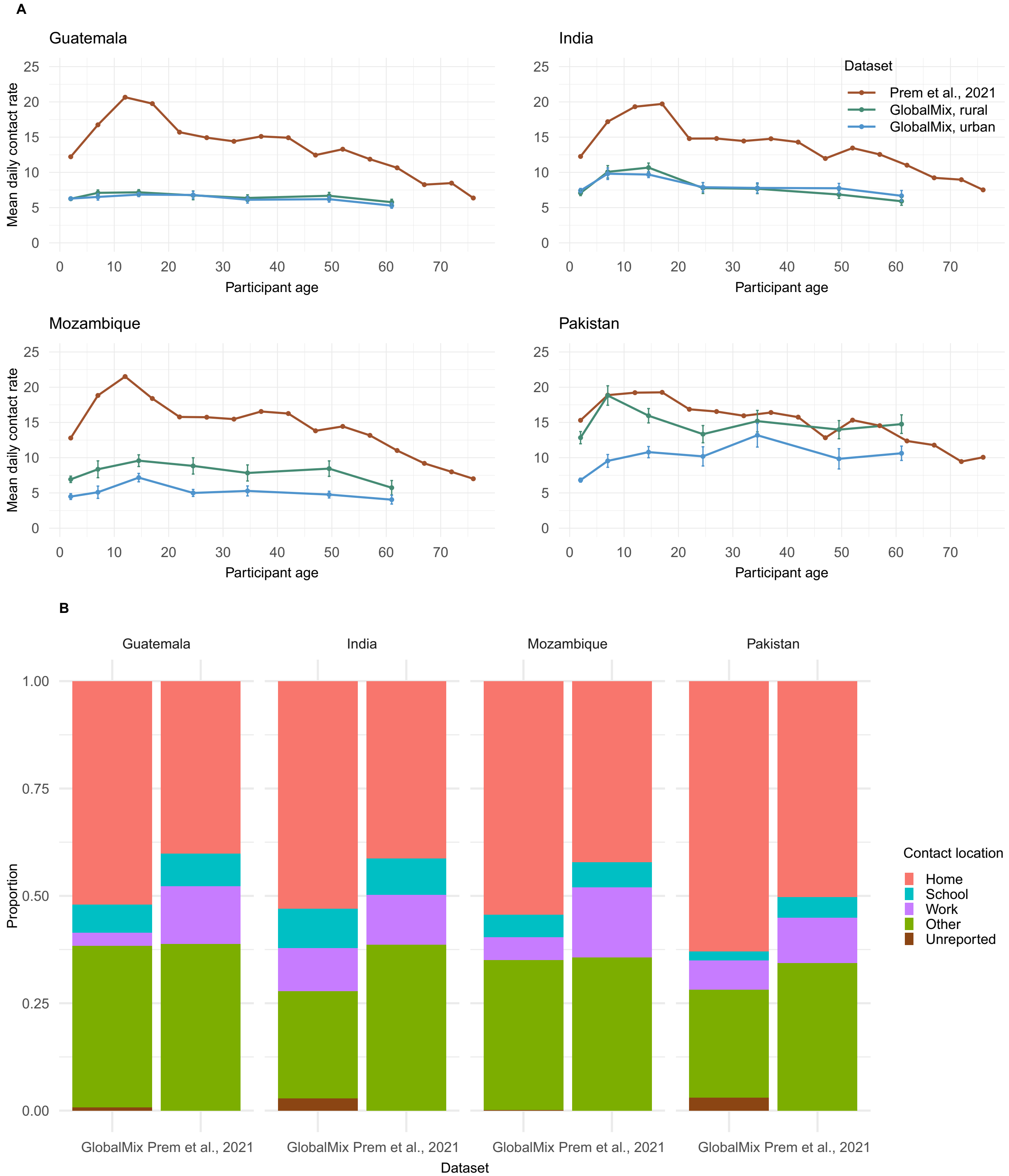

Fig. 3 | Comparison of contact rate estimates by age and location of contact with POLYMOD-based projections. A Mean daily contact rates by age estimated by Prem et al.[2] based on the POLYMOD data and the mean daily contact rates estimated in our study, GlobalMix. 95% confidence intervals are shown with vertical error bars for GlobalMix data. The number of participants is $n = 1141$ in Guatemala (Rural: $n = 575$, Urban: $n = 566$), $n = 1244$ in India (Rural: $n = 622$, Urban: $n = 622$), $n = 1352$ in Mozambique (Rural: $n = 675$, Urban: $n = 677$, $n = 1321$ in Pakistan (Rural: $n = 649$, Urban: $n = 672$). **B** Proportions of contacts reported in each location, categorized using the POLYMOD scheme. Contacts in GlobalMix for which the location was unreported are shown in brown.

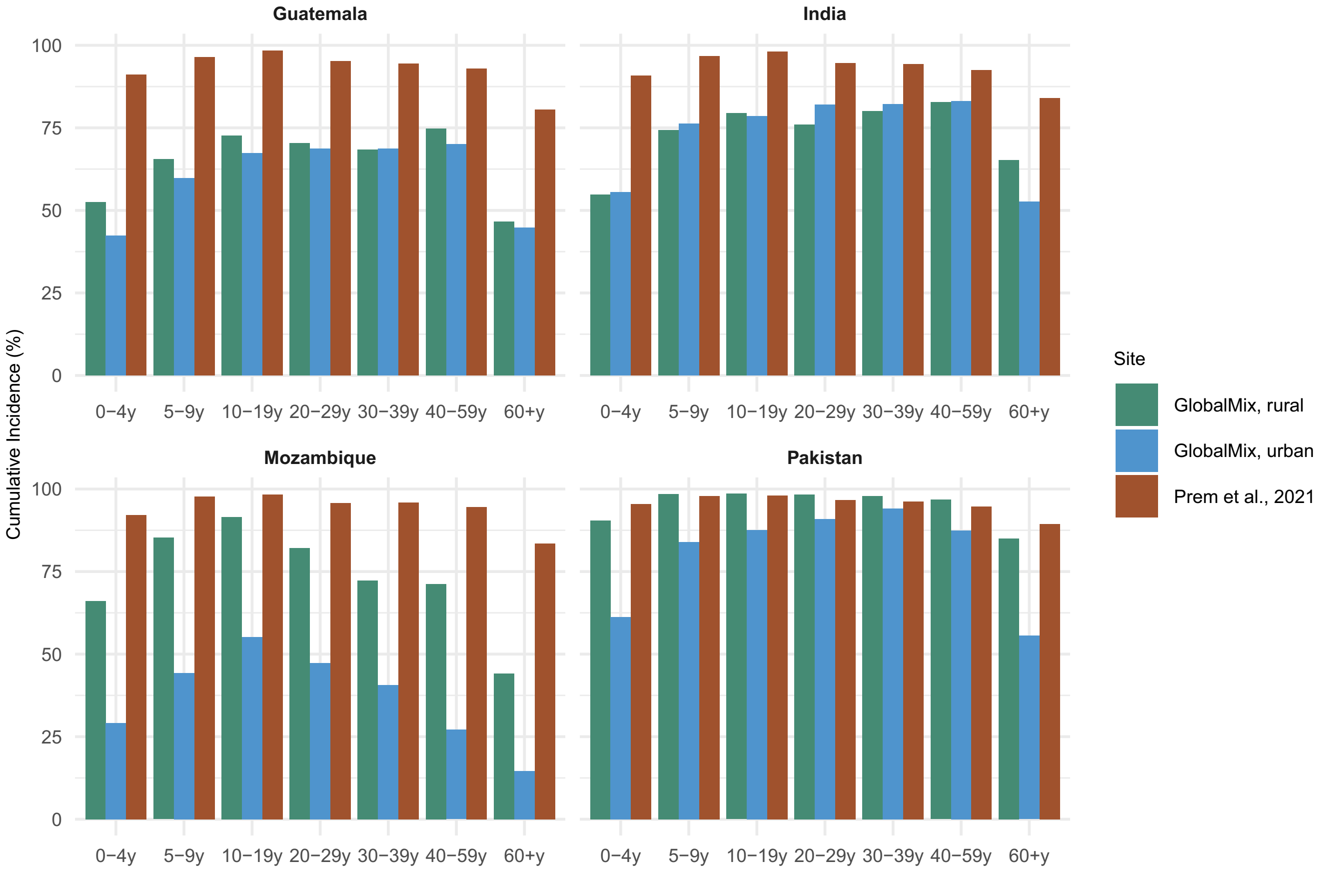

Fig. 4 | Effect of contact rate assumptions on epidemic dynamics by country and site. Brown bars show cumulative incidence of infection estimates under contact patterns based on those reported in Prem et al.[2] which are based on projections using the POLYMOD data. Green and blue bars show cumulative incidence estimates under contact patterns estimates from our study, GlobalMix, in the rural and urban site in each country, respectively.

history of most enteric and respiratory pathogens, it illustrates that using context-specific data can lead to different inferences about how pathogens will spread and affect demographic subpopulations.

There were some limitations to this study. Information on children's contacts were recorded using 'shadows', or individuals who recorded contacts on behalf of the participant. This may have led to inaccuracies since we do not know how completely shadows observed children's contacts. Second, we defined contact as an interaction involving three or more words with another person at distance of an arm's length. More transient interactions, which include being near a person in a crowded public place without speaking to them, are not captured by this definition. However, closer contacts are likely more important for transmission and allowed for comparison with previous work[18]. Third, these data should not be considered representative of each country. They can be considered representative of each site, since we adjusted overall contact rates for age and sex population distributions at each site. We recruited at multiple sites per country to capture at least some variation in contact patterns associated with varying population density within each country. Finally, we recruited participants over a limited time frame at each site, so could not capture seasonal variations in contact patterns that may be important for transmission (i.e., holidays, school start and end). One previous study in India has documented a higher number of contacts in the winter season relative to the monsoon or summer seasons; however, these differences were slight[32].

This study represents the first effort to collect and compare social contact data using standardized tools across multiple LMIC, with a special focus on contact patterns among children under 5 years of age. These data can inform future modeling efforts aiming to accurately capture behavioral patterns that drive the speed and patterns of spread of infections in different contexts, and to estimate the impact of interventions targeting a broad range of pathogens.

Methods

Data collection procedures

The GlobalMix study aimed to characterize social contact patterns across eight sites – one rural and one urban per country – in four countries. We chose countries spanning three continents, three World Bank income level classifications, and a range of social, cultural, and economic features. (Supplementary Table 7) The study design is described in detail elsewhere and involved a qualitative data collection phase, which included focus group discussions and cognitive interviews to adapt the data collection instruments and study procedures to each site[29]. Households, defined as a group of people who shared space for cooking, were enumerated and sampling was guided by existing demographic surveillance systems (India, Mozambique, Pakistan) and satellite imaging of households (Guatemala). In India, Mozambique and Pakistan, we used demographic surveillance data to generate age-stratified lists of community members and randomly select individuals within each age group for participation using a random number generator. Selected individuals were approached by field staff in the order they were selected until age-specific quotas were reached. In Guatemala, where demographic surveillance system data were not available, households within communities were randomly numbered and visited sequentially.

https://doi.org/10.1038/s41467-025-64850-9

Field staff randomly selected one member from each household and offered participation in the study until age-specific quotas were reached. Informed consent was obtained from adults over 18 years old and from parents/guardians for minors under 18 years old. Participants completed an enrollment questionnaire which included information on their age, gender, school enrollment/occupation, and characteristics of their household.

Field workers trained participants to complete a paper-based contact diary. Participants were randomly assigned a day of the week to begin two consecutive days of reporting contacts. Field workers returned to the household after this period and entered the contacts recorded by the participant into an electronic survey platform. Participants reported characteristics of the person with whom they had contact, such as their age, gender, and relationship to the participant, as well as the location (home, store/market, school, work, place of worship, and other country-specific options), duration (<5 min, 5–15 min, 16–30 min, 31 min–1 h, 1–4 h, and 4+ hours), environment (indoors/outdoors), and proximity (physical/non-physical) of the contact. For adults or children unable to complete the diary themselves, study staff assisted in selecting another person (a 'shadow') to record the participant's contacts. Both participants and shadows received training in how to record contacts and field staff periodically checked in with participants over the two-day contact reporting period to ensure complete and accurate reporting by shadows.

We defined contact as a two-way conversation with three or more words exchanged at a physical distance of an arm's length or less from another person. A subset of contacts were classified as physical, which we defined as a contact that involved directly touching someone's skin or the clothes they were wearing.

Statistical analysis

We applied post-stratification weights representative of the population age distribution at each site. Using these weights, we calculated the mean, standard deviation, median, and interquartile range of daily contacts, averaged across two days of reporting, for each site and country overall. We generated symmetric age-stratified contact matrices by age. Symmetric contact matrices adjust for differences in the size of age groups included in the sample and lack of reciprocity in contact reporting (i.e., 10–14-year-olds reported more contacts with 60+ year-olds than 60+ year-olds reported with 10–14-year-olds). We calculated contact exposure-hours by summing the length of each contact, using the midpoint of each duration category (e.g., 2.5 min for <5 min), and using 4 h for contacts reported as longer than 4 h.

We defined physical contacts that lasted longer than one hour as high risk for transmission of most infections. For infections that transmit via deposition onto surfaces (hands, faces, fomites), physical contact is more likely to lead to transmission; for infections that transmit through the air, exposures longer than one hour are usually deemed sufficient for spread[30–33].

To assess the robustness of relationships between demographic characteristics with contact rates, we evaluated associations between participant characteristics and (1) overall contacts and (2) non-household contacts using negative binomial multivariable regression models. We report site-specific adjusted odds ratios and 95% confidence intervals for each model.

Transmission modeling

To explore the impact of incorporating context-specific, empirical data on contact patterns on transmission dynamics, we built a simple compartmental model representing the transmission of an acute respiratory or enteric infection. We simulated an epidemic of an infection to which everyone is initially susceptible, with a mean duration of 14 days and which confers lifelong immunity. (Supplementary Table 8) We created a synthetic population representative of the age distribution in each site and separately modeled the spread of the infection in rural and urban areas based on (1) the GlobalMix contact patterns and (2) projected contact patterns for each country using POLYMOD data from Prem et al.[18]. We calculated and compared age-specific cumulative incidence, epidemic peak timing, and the effective reproduction number under different contact patterns. We simulated from the compartmental model using the 'EpiModel' R package version 2.4.0. and effective reproduction numbers were estimated using the parametric serial interval method, assuming a mean serial interval of 2.6 and standard deviation of 1.5, in the 'EpiEstim' R package version 2.2.4.

All data analysis was conducted in R version 4.1.2.

Ethical approval

This study was approved by Emory University Institutional Review Board (approval number 00105630), Yale University Institutional Review Board (reliance agreement approval number 2000026911), and review boards at Christian Medical College Vellore (Approval #12065), Universidad del Valle de Guatemala (Protocol #223-12-2020) and the National Ministry for Publicparticipants 10 years of age Health and Social Assistance in Guatemala (Protocol #31-2020), Manhiça Health Research Institute Internal Scientific Committee (CCI/03/2020) and Internal Ethical Review Board (initial approval CIBS-CISM/011/2020), and The Aga Khan University (Approval #2022-7912-23648).

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The questionnaires and data for this analysis are available at https://github.com/lopmanlab/GlobalMix and can also be found on Zenodo. (https://doi.org/10.5281/zenodo.16754436)[34].

Code availability

The code to reproduce this analysis is available at https://github.com/lopmanlab/GlobalMix and can also be found on Zenodo. (https://doi.org/10.5281/zenodo.16754436)[34].

References

1. Global health estimates: Leading causes of death. https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death (2000–2021).
2. Perin, J. et al. Global, regional, and national causes of under-5 mortality in 2000–19: an updated systematic analysis with implications for the Sustainable Development Goals. *Lancet Child Adolesc. Health* **6**, 106–115 (2022).
3. The unfinished agenda of communicable diseases among children and adolescents before the COVID-19 pandemic, 1990–2019: a systematic analysis of the Global Burden of Disease Study 2019. *Lancet* **402**, 313–335 (2023).
4. Read, J. M. et al. Social mixing patterns in rural and urban areas of southern China. *Proc. R. Soc. B Biol. Sci.* **281**, 20140268 (2014).
5. Kiti, M. C., Melegaro, A., Cattuto, C. & Nokes, D. J. Study design and protocol for investigating social network patterns in rural and urban schools and households in a coastal setting in Kenya using wearable proximity sensors. *Wellcome Open Res* **4**, 84 (2019).

6. Bennett, A. et al. Community transmission of rotavirus infection in a vaccinated population in Blantyre, Malawi: a prospective household cohort study. *Lancet Infect. Dis.* **21**, 731–740 (2021).
7. Worby, C. J. et al. On the relative role of different age groups in influenza epidemics. *Epidemics* **13**, 10–16 (2015).
8. Cohen, C. et al. Incidence and transmission of respiratory syncytial virus in urban and rural South Africa, 2017-2018. *Nat. Commun.* **15**, 116 (2024).
9. Mossong, J. et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLOS Med.* **5**, e74 (2008).
10. Wong, K. L. M. et al. Social contact patterns during the COVID-19 pandemic in 21 European countries – evidence from a two-year study. *BMC Infect. Dis.* **23**, 268 (2023).
11. World Bank Open Data. *World Bank Open Data* https://data.worldbank.org.
12. Jamison, D. T. et al. Global health 2035: a world converging within a generation. *Lancet* **382**, 1898–1955 (2013).
13. Campbell, P. T., McVernon, J., Shrestha, N., Nathan, P. M. & Geard, N. Who's holding the baby? A prospective diary study of the contact patterns of mothers with an infant. *BMC Infect. Dis.* **17**, 634 (2017).
14. Hoek et al. The social life of infants in the context of infectious disease transmission; social contacts and mixing patterns of the very young. *PLOS One* **8**, e76180 (2013).
15. Nagpal, S. et al. Seasonal variations in social contact patterns in a rural population in north India: Implications for pandemic control. *PLOS One* **19**, e0296483 (2024).
16. Leung, W. T. M. et al. Social contact patterns relevant for infectious disease transmission in Cambodia. *Sci. Rep.* **13**, 5542 (2023).
17. Dobreva, Z. et al. Characterising social contacts under COVID-19 control measures in Africa. *BMC Med.* **20**, 344 (2022).
18. Prem, K. et al. Projecting contact matrices in 177 geographical regions: An update and comparison with empirical data for the COVID-19 era. *PLOS Comput. Biol.* **17**, e1009098 (2021).
19. Prem, K., Cook, A. R. & Jit, M. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLOS Comput. Biol.* **13**, e1005697 (2017).
20. Mousa, A. et al. Social contact patterns and implications for infectious disease transmission – a systematic review and meta-analysis of contact surveys. *eLife* **10**, e70294 (2021).
21. Klepac, P. et al. Contacts in context: large-scale setting-specific social mixing matrices from the BBC Pandemic project. Preprint at https://doi.org/10.1101/2020.02.16.20023754 (2020).
22. Melegaro, A. et al. Social contact structures and time use patterns in the Manicaland Province of Zimbabwe. *PloS One* **12**, e0170459 (2017).
23. Leung, K., Jit, M., Lau, E. H. Y. & Wu, J. T. Social contact patterns relevant to the spread of respiratory infectious diseases in Hong Kong. *Sci. Rep.* **7**, 7974 (2017).
24. Kiti, M. C. et al. Quantifying age-related rates of social contact using diaries in a rural coastal population of Kenya. *PLOS ONE* **9**, e104786 (2014).
25. Melegaro, A., Jit, M., Gay, N., Zagheni, E. & Edmunds, W. J. What types of contacts are important for the spread of infections? Using contact survey data to explore European mixing patterns. *Epidemics* **3**, 143–151 (2011).
26. Oguz, M. M., Camurdan, A. D., Aksakal, F. N., Akcaboy, M. & Acoglu, E. A. Social contact patterns of infants in deciding vaccination strategy: a prospective, cross-sectional, single-centre study. *Epidemiol. Infect.* **146**, 1157–1166 (2018).
27. Neal, E. F. G. et al. Associations between ethnicity, social contact, and pneumococcal carriage three years post-PCV10 in Fiji. *Vaccine* **38**, 202–211 (2020).
28. Simon, A. K., Hollander, G. A. & McMichael, A. Evolution of the immune system in humans from infancy to old age. *Proc. R. Soc. B Biol. Sci.* **282**, 20143085 (2015).
29. Aguolu, O. G. et al. Comprehensive profiling of social mixing patterns in resource poor countries: a mixed methods research protocol. *PLOS One* **19**, e0301638 (2024).
30. Kenyon, T. A., Valway, S. E., Ihle, W. W., Onorato, I. M. & Castro, K. G. Transmission of multidrug-resistant Mycobacterium tuberculosis during a long airplane flight. *N. Engl. J. Med.* **334**, 933–938 (1996).
31. Driver, C. R., Valway, S. E., Morgan, W. M., Onorato, I. M. & Castro, K. G. Transmission of Mycobacterium tuberculosis associated with air travel. *JAMA* **272**, 1031–1035 (1994).
32. Butler, A. J., Thomas, M. K. & Pintar, K. D. M. Expert elicitation as a means to attribute 28 enteric pathogens to foodborne, waterborne, animal contact, and person-to-person transmission routes in Canada. *Foodborne Pathog. Dis.* **12**, 335–344 (2015).
33. DuPont, H. L., Levine, M. M., Hornick, R. B. & Formal, S. B. Inoculum size in shigellosis and implications for expected mode of transmission. *J. Infect. Dis.* **159**, 1126–1128 (1989).
34. Nelson, K., Shiiba, M., Kim, S. & Prasad, P. GlobalMix. *Zenodo* https://doi.org/10.5281/zenodo.17094566 (2025).

Acknowledgements

This study was funded by NIH/NICHD R01HD097175-01 and the lead author was supported by NIH/NIAID K01AI166093-01A1. The Manhiça Health Research Center (CISM) is supported by the Government of Mozambique and the Spanish Agency for International Development Cooperation (AECID). Polana Caniço Health Research Center (CISPOC) is supported by the Government of Mozambique. The funders of this study had no role in its design and conduct, nor the analysis and presentation of results. This study would not have been possible without the generous time and support of the individuals who participated. We also thank those who provided essential support of this project at the sites, including the social sciences teams in Mozambique and Guatemala (Clara Secaira, Liza Muñoz, María Fernanda González, among others), Consejos Comunitarios de Desarrollo Urbano y Rural (COCODES) and Ministry of Health Centers and Posts from San Juan Ostuncalco and Concepción Chiquirichapa, Quetzaltenango, Guatemala. We are grateful for the tireless work of field teams and data managers that were involved in data collection and database management at the sites. Finally, we are indebted to Gayathri S., a friend and colleague, who made valuable contributions to the study in India and to whom this manuscript is dedicated.

Author contributions

B.A.L., S.B.O., F.M., A.M., S.M.J., V.R., A.B., C.S., H.M., and M.K. conceived of and designed the study, M.C.K. led the cross-country data collection efforts, K.N.N., I.M., E.J., C.T., A.J., N.C., C.J., H.M.A., P.S., R.S., R.A., C.L., N.A., O.A., M.A.S., and I.Y. supported specific aspects of the data collection and/or led data collection at the sites, K.N.N., M.C.K., M.S., S.S.K., P.V.P., and D.C. conducted data analysis, K.N.N. wrote the initial draft of the manuscript, and all authors critically reviewed the manuscript for intellectual content.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-64850-9.

Correspondence and requests for materials should be addressed to Kristin N. Nelson.

Peer review information *Nature Communications* thanks Juanjuan Zhang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

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

[1]Emory University Rollins School of Public Health, Atlanta, GA, USA. [2]Manhiça Health Research Institute, Manhiça, Mozambique. [3]Barcelona Institute for Global Health ISGlobal, Barcelona, Spain. [4]National Institute of Health, Maputo, Mozambique. [5]Universidad del Valle de Guatemala, Guatemala City, Guatemala. [6]Christian Medical College, Vellore, Tamil Nadu, India. [7]The Aga Khan University, Karachi, Pakistan. [8]Peter O'Donnell Jr. School of Public Health, UT Southwestern Medical Center, Dallas, TX, USA. [9]The Ohio State University, Columbus, OH, USA. [10]Marshfield Clinic Research Institute, Marshfield, WI, USA. [11]Yale University, New Haven, CT, USA. [12]University of Milan, Milan, Italy. [13]These authors contributed equally: Kristin N. Nelson, Moses C. Kiti. [14]These authors jointly supervised this work. Benjamin A Lopman, Saad B Omer. ✉e-mail: knbratt@emory.edu