Social Contact Patterns and Implications for Infectious Disease Transmission: A Systematic Review and Meta-Analysis of Contact Surveys

Andria Mousa[1*], Peter Winskill[1], Oliver J Watson[1], Oliver Ratmann[2], Mélodie Monod[2], Marco Ajelli[3,4], Aldiouma Diallo[5], Peter J Dodd[6], Carlos G Grijalva[7], Moses Chapa Kiti[8], Anand Krishnan[9], Rakesh Kumar[9], Supriya Kumar[10], Kin On Kwok[11,12,13], Claudio F Lanata[14,15], Olivier Le Polain de Waroux[16], Kathy Leung[17,18], Wiriya Mahikul[19], Alessia Melegaro[20], Carl D Morrow[21,22], Joël Mossong[23], Eleanor FG Neal[24,25], David J Nokes[8,26], Wirichada Pan-ngum[27], Gail E Potter[28,29], Fiona M Russell[24,25], Siddhartha Saha[30], Jonathan D Sugimoto[31,32,33], Wan In Wei[11], Robin R Wood[21], Joseph T Wu[17,18], Juanjuan Zhang[34], Patrick GT Walker[1] & Charles Whittaker[1*]

*For correspondence: a.mousa17@imperial.ac.uk; charles.whittaker16@imperial.ac.uk

[1] MRC Centre for Global Infectious Disease Analysis; and the Abdul Latif Jameel Institute for Disease and Emergency Analytics (J-IDEA), School of Public Health, Imperial College London, London, UK.

[2] Department of Mathematics, Imperial College London, London, UK.

[3]Department of Epidemiology and Biostatistics, Indiana University School of Public Health, Bloomington, IN, USA.

[4] Laboratory for the Modeling of Biological and Socio-technical Systems, Northeastern University, Boston, MA.

[5] VITROME, Institut de Recherche pour le Developpement, Senegal.

[6] School of Health and Related Research, University of Sheffield, UK.

[7] Division of Pharmacoepidemiology, Department of Health Policy. Vanderbilt University Medical Center. Nashville, TN, USA.

[8] KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya.

[9] Centre for Community Medicine, All India Institute of Medical Sciences, New Delhi, India.

[10] Bill & Melinda Gates Foundation, Seattle, USA.

[11] JC School of Public Health and Primary Care, The Chinese University of Hong Kong, Hong Kong Special Administrative Region, China.

[12] Stanley Ho Centre for Emerging Infectious Diseases, The Chinese University of Hong Kong, Hong Kong Special Administrative Region, China.

[13] Shenzhen Research Institute of The Chinese University of Hong Kong, Shenzhen, China.

[14] Instituto de Investigación Nutricional, Lima, Peru.

[15] Department of Medicine, Vanderbilt University, Nashville, TN, USA.

[16] London School of Hygiene and Tropical Medicine, London, UK.

[17] WHO Collaborating Centre for Infectious Disease Epidemiology and Control, School of Public Health, LKS Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China.

[18] Laboratory of Data Discovery for Health (D24H), Hong Kong Science Park, New Territories, Hong Kong SAR, China.

[19] Faculty of Medicine and Public Health, HRH Princess Chulabhorn College of Medical Science, Chulabhorn Royal Academy, Bangkok 10210, Thailand.

[20] Dondena Centre for Research on Social Dynamics and Public Policy, Department of Social and Political Sciences, Bocconi University, Milan, Italy.

[21] Desmond Tutu HIV Centre, Department of Medicine, Faculty of Health Sciences, University of Cape Town, South Africa.

[22] Centre for Infectious Disease Epidemiology and Research (CIDER), School of Public Health and Family Medicine, Faculty of Health Sciences, University of Cape Town South Africa.

[23] Health Directorate, Luxembourg.

[24] Infection & Immunity, Murdoch Children's Research Institute, Parkville, Victoria, Australia

[25] Department of Paediatrics, University of Melbourne, Parkville, Victoria, Australia.

[26] School of Life Sciences, University of Warwick, Coventry UK.

[27] Department of Tropical Hygiene, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand

[28]National Institute for Allergies and Infectious Diseases, National Institutes of Health, Rockville MD, USA.

[29] The Emmes Company, Rockville MD, USA.

[30] Influenza Programme, US Centers for Disease Control and Prevention, India Office, US Embassy, New Delhi.

[31] Seattle Epidemiologic Research and Information Center, Cooperative Studies Program, Office of Research and Development, United States Department of Veterans Affairs, USA.

[32] Department of Epidemiology, University of Washington, USA.

[33] Fred Hutchinson Cancer Research Center, Seattle, WA, USA.

[34] School of Public Health, Fudan University, Key Laboratory of Public Health Safety, Ministry of Education, Shanghai, China.

Abstract

Background: Transmission of respiratory pathogens such as SARS-CoV-2 depends on patterns of contact and mixing across populations. Understanding this is crucial to predict pathogen spread and the effectiveness of control efforts. Most analyses of contact patterns to date have focussed on high-income settings.

Methods: Here, we conduct a systematic review and individual-participant meta-analysis of surveys carried out in low- and middle-income countries and compare patterns of contact in these settings to surveys previously carried out in high-income countries. Using individual-level data from 28,503 participants and 413,069 contacts across 27 surveys we explored how contact characteristics (number, location, duration and whether physical) vary across income settings.

Results: Contact rates declined with age in high- and upper-middle-income settings, but not in low-income settings, where adults aged 65+ made similar numbers of contacts as younger individuals and mixed with all age-groups. Across all settings, increasing household size was a key determinant of contact frequency and characteristics, with low-income settings characterised by the largest, most intergenerational households. A higher proportion of contacts were made at home in low-income settings, and work/school contacts were more frequent in high-income strata. We also observed contrasting effects of gender across income-strata on the frequency, duration and type of contacts individuals made.

Conclusions: These differences in contact patterns between settings have material consequences for both spread of respiratory pathogens, as well as the effectiveness of different non-pharmaceutical interventions.

Funding: This work is primarily being funded by joint Centre funding from the UK Medical Research Council and DFID (MR/R015600/1).

Introduction

Previous outbreaks of Ebola(Mbala-Kingebeni et al., 2019), influenza(Khan et al., 2009), and the ongoing COVID-19 pandemic have highlighted the importance of understanding the transmission dynamics and spread of infectious diseases, which depend fundamentally on the underlying patterns of social contact between individuals. Together, these patterns give rise to complex social networks that influence disease dynamics(Eubank et al., 2004; Ferrari et al., 2006; Firth et al., 2020; Zhang et al., 2020), including the capacity for emergent pathogens to become endemic(Ghani and Aral, 2005; Jacquez et al., 1988), the overdispersion of the offspring distribution underlying the reproduction number(Delamater et al., 2019) and the threshold at which herd-immunity is reached(Fontanet and Cauchemez, 2020; Mistry et al., 2021). They can similarly modulate the effectiveness of non-pharmaceutical interventions (NPIs), such as school closures and workplace restrictions, that are typically deployed to control and contain the spread of infectious diseases (Prem et al., 2020).

Social contact surveys provide insight into the features of these networks, which is typically achieved through incorporating survey results into mathematical models of infectious disease transmission frequently used to guide decision making in response to outbreaks(Chang et al., 2021; Davies et al., 2020). Such inputs are necessary for models to have sufficient realism to evaluate relevant policy questions. However, despite the known importance of contact patterns as determinants of the infectious disease dynamics, our understanding of how they vary globally remains far from complete. Reviews of contact patterns to date have focussed on High-Income countries (HICs)(Hoang et al., 2019). This is despite evidence that social contact patterns differ systematically across settings in ways that have material consequences for the dynamics of infectious disease transmission and the evolution of epidemic trajectories(Prem et al., 2017; Walker et al., 2020). Previous reviews has also primarily explored the total number of contacts made by individuals(Hoang et al., 2019) and/or how these contacts are distributed across different age/sex groups(Horton et al., 2020). Whilst these factors are a vital component underpinning disease spread, recent work has also underscored the importance of the characteristics of contacts (such as the location, duration and extent of physical contact) in determining transmission risk(Thompson et al., 2021).

Here, we carry out a systematic review of contact surveys (conducted prior to the emergence of COVID-19) in Lower-Income, Lower-Middle and Upper-Middle-Income countries (LICs, LMICs and UMICs, respectively). Alongside previously published data from HICs(Kwok et al., 2018, 2014; Leung et al., 2017; Mossong et al., 2008), we collate individual participant data (IPD) on social contacts from published work spanning 27 surveys from 22 countries and over 28,000 individuals. We use a Bayesian framework to explore drivers and determinants of contact patterns across a wider range of settings and at a more granular scale than has previously been possible. Specifically, we assess the influence of key factors such as age, gender and household structure on both the total number and characteristics (such as duration, location and type) of contact made by an individual, and explore how the comparative importance of different factors varies across different settings. We additionally evaluate the extent and degree of assortativity in contact patterns between different groups, and how this varies across settings.

Results

Systematic Review and Individual-Participant-Data (IPD) Meta-analysis

A total of 3,409 titles and abstracts were retrieved from the databases, and 313 full-text articles were screened for eligibility (Supplementary Figure 1). This search identified 19 studies with suitable contact data from LIC, LMIC and UMIC settings– individual-level data were obtained from 16 of these studies, including one study from a LIC, six studies from a LMIC and nine studies from an UMIC. These were analysed alongside four HIC studies from Hong Kong and Europe. The majority of the studies collected data representative of the general population, through random sampling and included a combination of both rural and urban sites (see Supplementary Text 1 for further details). Although most studies included respondents of all ages, one study restricted their participants to ages over 18 years (Dodd et al., 2015), one to ages over 15 years(Mahikul et al., 2020), one to ages over 6 months(Huang et al., 2020), one study only collected contact data on infants under 6 months(Oguz et al., 2018) and another on contacts of children under 6 years and their caregivers(Neal et al., 2020). The distribution of participant age groups in each study was also dependent on the sampling method. For instance, two studies focused on school and university students and their contacts, thereby oversampling older children and young adults (Ajelli and Litvinova, 2017; Stein et al., 2014). Details of the identified studies and a full description of the systematic review findings can be found in Supplementary Text 1 and Supplementary Table 1.

In total, this meta-analysis yielded 28,503 participants reporting on 413,069 contacts. All studies contained information on main demographic variables such as age and gender. Availability of other variables analysed here for each study are listed in Supplementary Table 2. All studies reported the number of contacts made in the past 24 hours of (or day preceding) the survey. The definitions of contacts were broadly similar across studies (Supplementary Table 1). Specifically, contacts were defined as skin-to-skin (physical) contact or a two-way conversation in the physical presence of another person. All studies scored above 65% of the items on the AXIS risk of bias tool, suggesting good or fair quality (Supplementary Table 3). Among all participants 47.5% were male, 30.1% were aged under 15 years and 7.2% were aged over 65 years. The majority (83.4%) of participants were asked to report the number of contacts they made on a weekday. A large proportion (34.1%) of respondents lived in large households of 6 or more people but this was largely dependent on income setting (LIC/LMIC=63.2%, UMIC=35.9%, HIC=4.9%). Among school-aged children (5 to 18 years), 88.1% were students, and 59.1% of adults aged over 18 were employed.

Total number of contacts and contact location

The median number of contacts made per day across all the studies was 9 (IQR= 5-17), and was similar across income strata (LIC/LMIC=10[5-17], UMIC=8[5-16], HIC=9[5-17]; Table 1). There was a large variation in contact rates across different studies, with the median number of daily contacts ranging from 4 in a Zambian setting(Dodd et al., 2015) to 24 in an online Thai survey(Stein et al., 2014). When stratifying by study methodology, median daily contacts was higher in diary-based surveys compared to interview-/questionnaire- based surveys, which was true across all income strata (Table 1, Supplementary Figure 2).

Overall, children aged 5 to 15 had the highest number of daily contacts (Figure 1A-C), although there was substantial variation between studies and across income-strata in how the number of daily contacts varied with age (Figure 1A-C). Across UMICs and HICs, the number of daily contacts made by participants decreased with age, with this decrease most notable in the oldest age-groups (adjCRR for 65+ vs. <15 years [95%CrI]: UMIC=0.67[0.63-0.71] and HIC=0.57[0.54-0.60]). By contrast, there was no evidence of contact rates declining in the oldest age-groups in LICs/LMICs (adjCRR for 65+ vs. <15 years [95%CrI]=0.94[0.89-1.00]). We observed contrasting effects of gender on the number of daily contacts, with men making more daily contacts compared to women in LICs/LMICs after accounting for age (adjCRR=1.17, 95%CrI:1.15-1.20; Figure 1D), but no effect of gender on total daily contacts for other income strata (CRR[95%CrI]: UMIC=1.01[0.98-1.04], HIC=0.99[0.97-1.02]). There were also differences in the number of daily contacts made according to the methodology used and whether the survey was carried out on a weekday or over the weekend – in both instances, contrasting effects of these factors on the number of daily contacts according to income strata were observed (Figures 1D-1F).

We also examined the influence of factors that might influence both the total number and location (home, work, school and other) of the contacts individuals make. Across all income-strata, students (defined as those currently in education, attending school and aged between 5 and 18 years) made more daily contacts than non-students aged between 5 and 18 (adjCRR [95%CrI]: LIC/LMIC=1.26[1.16-1.37], UMIC=1.18[1.03-1.35] and HIC=1.54[1.42-1.66]; Figure 1D-F). Similarly, we observed strong and significant effects of employment in all income strata, with adults who were employed having a higher number of total daily contacts compared to those not in employment (adjCRR [95%CrI]: LIC/LMIC= 1.17[1.12-1.23], UMIC= 1.07[1.03-1.13], HIC= 1.60[1.54-1.65]; Figure 1D-F). The number of daily contacts made at home were proportional to the participant's household size (Supplementary figure 3). Total daily contacts increased with household size (Figure 2A, Supplementary Figure 2) across all income-strata; individuals living in large households (6+ members) had 1.47 (95%CrI:1.32-1.64) (LIC/LMICs), 2.58 (95%CrI:2.37-2.80) (UMICs) and 1.51 (95%CrI:1.40-1.63) (HICs) times more daily contacts than those living alone, after accounting for age and gender (Figure 1E-F). Sensitivity analyses excluding additional contacts (as defined in Methods), showed little difference in effect sizes for total daily contacts, and were strongly correlated with the effect sizes shown in Figure 1D-F (Supplementary Figure 4).

Motivated by this suggestion of strong, location-related (school, work and household) effects on total daily contact rates, we further explored the locations in which contacts were made. Contact location was known for 314,235 contacts, 42.7% of which occurred at home (13.1% at work, 12.5% at school and 31.7% in other locations). Across income-strata, there was significant variation in the proportion of contacts made at home – being highest in LICs/LMICs (68.3%) and lowest in HICs (37.0%) (Figure 2B). Age differences were also observed in the number of contacts made at home, particularly for LICs/LMICs (Figure 2C-2D). Relatedly, a higher proportion of contacts occurred at work and school (14.6 % and 11.3%) in HICs compared to LICs/LMICs (3.9% and 5.2%, respectively; Supplementary Figure 5). Strong, gender specific patterns of contact location were also observed. Across all income strata males made a higher proportion of their contacts at work compared to females, although this difference was largest for LICs/LMICs (Supplementary Figure 5). Further, we found significant variation between income strata in median household size (7 in LICs/LMICs, 5 in UMICs and 3 in HICs). This trend of decreasing household size with increasing country income was consistent with global data (Figure 2E). The larger households observed for LIC/LMIC settings were also more likely to be intergenerational – in LICs/LMICs, 59.4% of participants aged over 65 lived in households of at least 6 members compared to 17.5% in UMICs and only 2.2% in HICs.

Type and duration of contact

Data on the type of contacts (physical and non-physical) were recorded for 20,910 participants. The mean percentage of physical contacts across participants was 56.0% and was the highest for LICs/LMICs (64.5%). At the study level, the highest mean percentage of physical contacts was observed for a survey of young children and their caregivers conducted in Fiji(Neal et al., 2020) (84.0%) and the lowest in a Hong Kong contact survey(Leung et al., 2017)(18.9%). Physical contact was significantly less common among adults compared to children under 15 years in all settings (ORs ranged between 0.22 to 0.48) (Figure 3A-F). Despite the proportion of physical contacts generally decreasing with age, there was a higher proportion observed for adults aged 80 or over (Figure 3A-C). Contacts made by male participants were more likely to be physical compared to female participants in UMICs (adjOR= 1.13, 95%CrI=1.10-1.16) and HICs (adjOR= 1.09, 95%CrI=1.07-1.12), but in LICs/LMICs men had a lower proportion of physical contacts than women (adjOR= 0.81, 95%CrI=0.79-0.83; Figure 3D-F). Most physical contacts made by women in LICs were made at home (73.5%), whilst for HICs this was just 41.4% - similar differences across income-strata were observed for men, although the proportions were always lower than observed for women (62.4% for LIC/LMICs and 36.4% for HICs). Increasing household size was generally associated with a higher proportion of contacts being physical (for households of 6+ members compared to 1 member: adjCRR[95%CrI]: LIC/LMIC=1.73[1.48-2.02], UMIC= 1.30[1.12-1.52], HIC= 1.57[1.48-1.67]; Figure 3D-F). Employment was associated with having a significantly lower proportion of physical contacts in LICs/LMICs (adjOR=0.83, 95%CrI:0.79-0.87) and HICs (adjOR=0.71, 95%CrI:0.69-0.73), but not in UMICs (adjOR=1.11, 95%CrI:1.03-1.19). The proportion of physical contacts among all contacts was the highest for households (70.4%), followed by schools (58.5%), community (55.7%) and work (33.6%) (Supplementary Figure 6).

Data on the duration of contact (<1 or ≥1hr) were available for 22,822 participants. The percentage of contacts lasting at least 1 hour was 63.2% and was highest for UMICs (76.0%) and lowest for LICs/LMICs (53.1%). Across both UMICs and HICs, duration of contacts was lower in individuals aged over 15 years compared to those aged 0-15, with the extent of this disparity most stark for HICs (for ages 65+ compared to <15 years: adjCRR [95%CrI]: LIC/LMIC= 0.61[0.57-0.64], UMIC= 0.61[0.58-0.65], HIC= 0.35[0.33-0.37]; Figure 4A-F). We observed contrasting effects of gender across income-strata: males made longer-lasting contacts than females in UMICs (adjOR=1.11, 95%CrI=1.08-1.14); Figure 4D-F), but not in LIC/LMICs (adjOR=0.92, 95%CrI=0.90-0.95) or HICs (adjOR=0.98, 95%CrI=0.97-1.00). Participants reported shorter contacts on weekends compared to weekdays in LICs/LMICs (adjOR=0.91, 95%CrI: 0.88-0.95), and HICs (adjOR=0.95, 95%CrI: 0.92-0.97), but not in UMICs (adjOR=1.12, 95%CrI=1.03-1.21). Contacts lasting over an hour as a proportion of all contacts was highest for households (72.7%), followed by schools (67.9%), community (47.0%) and work (44.0%). However, it was only in HICs that there was a significant effect of being a student (adjOR=1.18, 95%CrI: 1.09-1.27; Figure 4D-F) on the proportion of contacts lasting ≥1 hour. For all income strata, the proportion of contacts >1h increased with increasing household size (Figure 4D-F). The sensitivity analysis weighing all studies equally within an income group yielded similar results to those from the main analysis (range of Pearson's correlation coefficients between main analysis and sensitivity analysis effect sizes: 0.92-1.00), and any differences are discussed in Supplementary Text 2.

Assortativity by age and gender

Twelve studies collected information on the gender of the contact and eight studies contained information on age allowing assignment of contacts to one of the three age-groups described in Methods (Supplementary Table 2, Supplementary Text 3). We found evidence to suggest that contacts were assortative by gender for all income strata, as participants were more likely to mix with their own gender (Supplementary Text 3). Mixing was also assortative by age, with participants more likely to contact individuals who belonged to the same age group this degree of age-assortativity was lowest for LICs/LMICs, where only 29% of contacts made by adults were with individuals of the same age group. By contrast, in HICs we observed a higher degree of assortative mixing, with most contacts (51.4%) made by older adults occurring with individuals belonging to the same age group.

Discussion

Understanding patterns of contact across populations is vital to predicting the dynamics and spread of infectious diseases, as well understanding the control interventions likely to have the greatest impact. Here, using a systematic review and individual-participant data meta-analysis of contact surveys, we summarise research exploring these patterns across a range of populations spanning 28,503 individuals and 22 countries. Our findings highlight substantial differences in contact patterns between income settings. These differences are driven by setting-specific sociodemographic factors such as age, gender, household structure and patterns of employment, which all have material consequences for transmission and spread of respiratory pathogens.

Across the collated studies, the total number of contacts was highest for school-aged children. This is consistent with previous results from HICs(Béraud et al., 2015; Fu et al., 2012; Hoang et al., 2019; Ibuka et al., 2016; Lapidus et al., 2013) and shown here to be generally true for LICs/LMICs and UMICs also. Interestingly however, we observed differences in patterns of contact in adults across income strata. Whilst contact rates in HICs declined in older adults, this was not observed in LICs/LMICs, where contact rates did not differ in the oldest age-group compared to younger ages. This is consistent with variation in household structure and size across settings, with nearly two thirds of participants aged 65+ in included LIC/LMIC surveys living in large, likely intergenerational, households (6+ members), compared to only 2% in HICs. HICs were also characterised by more assortative mixing between age-groups, with older adults in LICs/LMICs more likely to mix with individuals of younger ages, again consistent with the observed differences between household structures across the two settings. These results have important consequences for the viability and efficacy of protective policies centred around shielding of elderly individuals (i.e. those most at risk from COVID-19 or influenza. In these settings other strategies may be required to effectively shield vulnerable populations, as has been previously suggested (Dahab et al., 2020).Our results support the idea of households as a key site for transmission of respiratory pathogens(Thompson et al., 2021), with the majority of contacts made at home. Our analysis highlights that the number of contacts made at home is mainly driven by household size. However, the relative importance of households compared to other locations is likely to vary across settings. We observed significant differences across income settings in the distribution of contacts made at home, work and school. The proportion of contacts made at home was highest for LIC/LMICs, where larger average household sizes were associated with more contacts, more physical contacts, and longer lasting contacts. By contrast, participants in HICs tended to report more contacts occurring at work and school. The lower number of contacts at work in LIC/LMIC may be explained by the types of employment (e.g agriculture in rural surveys) and a selection bias (women at home/homemakers more likely to be surveyed in questionnaire-based surveys). Our analyses similarly highlighted significant variation in the duration and nature of contacts across settings. Contacts made by female participants in LICs/LMICs were more likely to be physical compared to men, whilst the opposite effect was observed for HICs and UMICs, potentially reflecting context-specific gender roles. In all settings, we observed a general decline of physical contacts with age, except in the very old(Mossong et al., 2008), potentially reflecting higher levels of dependency and the need for physical care.

Altogether, these results suggest differences between settings in the comparative importance of different locations (such as the household or the workplace) to transmission of SARS-CoV-2, a finding which would likely modulate the impact of different NPIs (such as workplace or school closures, stay at home orders etc). Moreover, it suggests that previous estimates of NPI effectiveness (primarily derived from European data and settings (Brauner et al., 2021) may be of limited generalisability to non-European settings characterised by different structures and patterns of social contact. However, beyond highlighting heterogeneity in where and how transmission is likely to occur, it remains challenging to disentangle exactly how these differences in contact patterns would shape patterns of transmission. Whilst the collated data provide a cross-sectional snapshot into the networks of social contact underpinning transmission, they remain insufficient to completely resolve this network or its temporal dynamics. Our results therefore do not consider key features relevant to population-level spread and transmission (such as overall network structure or the extent of repeated contacts, which would be most likely to occur with household members) which previous work has demonstrated can have a significant impact on infectious disease dynamics, both in general terms (Bansal et al., 2010; Keeling and Eames, 2005) as well as with COVID-19 (Rader et al., 2020). It is in this context that recent results generating complete social networks (including both the frequency and identity of an individual's contacts) from high-resolution GPS data represent promising developments in understanding social contact networks and how they shape transmission (Firth et al., 2020).

There are important caveats to these findings. Data constraints limited the numbers of factors we were able to explore – for example, despite evidence(Kiti et al., 2014) suggesting that contact patterns differ across rural and urban settings, only 3 studies(Kiti et al., 2014; O. le Polain de Waroux et al., 2018; Neal et al., 2020) contained information from both rural and urban sites, allowing classification. Similarly, we were unable to examine the impact of socioeconomic factors such as household wealth, despite experiences with COVID-19 having highlighted strong socio-economic disparities in both transmission and burden of disease(De Negri et al., 2021; Routledge et al., 2021; Ward et al., 2021; Winskill et al., 2020) and previous work suggesting that poorer individuals are less likely to be employed in occupations amenable to remote working(Loayza, 2020). A lack of suitably detailed information in the studies conducted precludes analysis of these factors but highlights the importance of incorporating economic questions into future contact surveys, such as household wealth and house square footage. Other factors also not controlled for here, but that may similarly shape contact patterns include school holidays or seasonal variations in population movement and composition that we are unable to capture given the cross-sectional nature of these studies.

Another important limitation to these results is that we are only able to consider a limited set of contact characteristics (the location and duration of the contact and whether it was physical). Previous work has highlighted the importance of these factors in determining the risk of respiratory pathogen transmission(Chang et al., 2021; Dunne et al., 2018; Olivier le Polain de Waroux et al., 2018; Neal et al., 2020; Thompson et al., 2021), but only a limited number of studies reported whether a contact was “close” or “casual”(Kwok et al., 2018, 2014; O. le Polain de Waroux et al., 2018) and whether the contact was made indoors or outdoors(Wood et al., 2012); both factors likely to influence transmission risk(Bulfone et al., 2021; Chu et al., 2020). More generally, the relevance and comparative importance of different contacts to transmission likely varies according to the specific pathogen and its predominant transmission modality (e.g. aerosol, droplet, fomite etc). It is therefore important to note that these results do not provide a direct indication of explicit transmission risk, but rather an indicator of factors likely to be relevant to transmission.

Relatedly, it is also important to note that the studies collated here were conducted over a wide time-period (2005-2018). In conjunction with the cross-sectional nature of the included studies, this precludes us from being able to examine for potential time-related trends in contact patterns. Additionally, the collated surveys were all carried out prior to the onset of the SARS-CoV-2 pandemic. Previous work has documented significant alterations to patterns of social contact in response to individual-level behaviour changes or government implemented NPIs aimed at controlling SARS-CoV-2 spread, and that these changes are dynamic and time-varying (Gimma et al., 2021; McCreesh et al., 2021). A detailed understanding of the impact of changing contact patterns on disease spread necessarily requires both an understanding of baseline contact patterns (as detailed in the studies collated here), and what changes have occurred as a result of control measures – however this latter data remains sparse and is available for only a limited number of settings(Jarvis et al., 2021, 2020; Liu et al., 2021). Description of contact location was also coarse and precluded more granular analyses of specific settings, such as markets, which have previously been shown to be important locations for transmission in rural areas(Grijalva et al., 2015).

Heterogeneity between studies was larger for LICs/LMICs and UMICs, which we partly accounted for, through fitting random study effects. These study differences may be attributed to the way individual contact surveys were conducted, making comparisons of contact patterns among surveys more difficult (e.g. prospective/retrospective diary surveys, online/paper questionnaires, face-to-face/phone interviews, and different contact definitions). For instance, there is evidence suggesting that prospective reporting, which is less affected by recall bias, can often lead to a higher number of contacts being reported(Mikolajczyk and Kretzschmar, 2008) and a lower probability of casual or short-lasting contacts being missed. The relatively high contact rates observed in HICs may be explained by the fact that all but two HIC surveys used diary methods. Our study highlights that a unified definition of "contact" and standard practice in data collection could help increase the quality of collected data, leading to more robust and reliable conclusions about contact patterns. Whilst we aggregate results by income strata due to the limited availability of data (particularly in lower- and middle-income countries), it is important to note that the outcomes considered here are likely to be shaped by several different factors other than country-level income. Whilst some of these factors will be correlated with a country's income status (e.g. household size(Walker et al., 2020)), many others will be unique to a particular setting or geographical area or correlate only weakly with country-level data. Examples include patterns of employment, the role of women, and other contextual factors. These analyses are therefore intended primarily to provide indications of prevailing patterns, rather than a definitive description of contact patterns in a specific context and highlight the significant need for further studies to by carried out in a diversity of different locations.

Despite these limitations however, our results highlight significant differences in the structure and nature of contact patterns across settings. These differences suggest that the comparative importance of different locations and age-groups to transmission will likely vary across settings and have critical consequences for the efficacy and suitability of strategies aimed at controlling the spread of respiratory pathogens such as SARS-CoV-2. Most importantly, our study highlights the limited amount of work that has been undertaken to date to better understand and quantify patterns of contact across a range of settings, particularly in lower- and middle-income countries, which is vital in informing control strategies reducing the spread of such pathogens.

Methods

Systematic Review

Data sources and search strategy: Two databases (Ovid MEDLINE and Embase) were searched on 26th May 2020 to identify studies reporting on contact patterns in LICs, LMICs and UMICs (Supplementary Table 4). Collated records underwent title and abstract screening for relevance, before full-text screening using pre-determined criteria. Studies were included if they reported on any type of face-to-face or close contact with humans and were carried out in LICs, LMICs or UMICs only. No restrictions on collection method (e.g. prospective diary-based surveys or retrospective surveys based on a face-to-face/phone interview or questionnaire) were applied. Studies were excluded if they did not report contacts relevant to air-borne diseases (e.g. sexual contacts), were conducted in HICs, were contact tracing studies of infected cases, or were conference abstracts. All studies were screened independently by two reviewers (AM and CW). Differences were resolved through consensus and discussion. The study protocol can be accessed through PROSPERO (registration number: CRD42020191197). Income group classification (LIC/LMIC, UMIC, or HIC) was based on 2019 World Bank data (fiscal year 2021)(World Bank Group, 2020).

Data extraction: Individual-level data were obtained from publication supplementary data, as well as online data repositories such as Zenodo, figshare and OSF. When not publicly available, study authors were contacted to request data. Extracted data included the participant's age, gender, employment, student status, household size and total number of contacts, as well as the day of the week for which contacts were reported. Some studies reported information at the level of individual contacts and included the age, gender, location and duration of the contact, as well whether it involved physical contact. Individual-level data from HICs, not systematically identified, were used for comparison, and included three studies from Hong Kong(Kwok et al., 2018, 2014; Leung et al., 2017) and the 8 European countries from the POLYMOD study(Mossong et al., 2008). Data were collated, cleaned and standardised using Stata version 14. Country-specific average household size were obtained from the United Nations Database on Household Size and Composition(United Nations Department of Economic and Social Affairs Population Division, 2019). Gross domestic product based on purchasing power parity (GDP PPP) was obtained from the World Data Bank database(World Bank International Comparison Programme, 2021). Findings are reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist of items specific to IPD meta-analyses (Supplementary Table 5). Risk of bias was assessed using the AXIS critical appraisal tool used to evaluate quality of cross-sectional studies(Downes et al., 2016), modified to this study's objectives (Supplementary Table 3). Each item was attributed a zero or a one, and a quality score was assigned to each study, ranging from 0% ("poor" quality) to 100% ("good" quality). The individual-level data across all studies and analysis code are available at https://github.com/mrc-ide/contact_patterns (see Supplementary Text 4 for data dictionary).

Statistical analysis

The mean, median and interquartile range of total daily unique contacts were calculated for subgroups including country income status, individual study, survey methodology (diary-based or questionnaire/interview-based), survey day (weekday/weekend), and respondent characteristics such as age, sex, employment/student status and household size. Detailed description of data assumptions for each study can be found in Supplementary Text 4.

A negative binomial regression model was used to explore the association between the total number of daily contacts and the participant’s age, sex, employment/student status and household size, as well as methodology and survey day. Incidence rate ratios from these regressions are referred to as “Contact Rate Ratios” (CRRs). A sensitivity analysis was carried out that excluded additional contacts (such as additional work contacts, group contacts, and number missed out, which were recorded separately and in less detail by participants compared to their other contacts (Ajelli and Litvinova, 2017; Kumar et al., 2018; Leung et al., 2017; Zhang et al., 2020)). Logistic regressions were used to explore determinants of contact duration (<1hr/1hr+) and type (physical/non-physical), using the same explanatory variables as in the total contacts analyses. There were differences in the contact duration categories defined by studies, and the threshold of 1 hour for longer durations was used to maximise sample size, by allowing inclusion of all available data. An additional sensitivity analysis, weighing all studies equally within an income stratum, explored the impact of study size on the estimated CRRs and ORs for all main outcomes (total contacts, duration and whether physical). The proportion of contacts made at each location (home, school, work and other) was explored descriptively and contacts made with the same individual in separate locations/instances were considered as separate contacts.

All analyses were done in a Bayesian framework using the probabilistic programming language Stan, using uninformative priors in all analyses and implemented in R via the package *brms*(Bürkner, 2018, 2017). All analyses were stratified by three income strata (LICs and LMICs were combined to preserve statistical power) and included random effects by study, to account for heterogeneity between studies. The only exceptions to this were any models adjusting for methodology which did not vary by study. The effect of each factor was explored in an age- and gender-adjusted model. All models exploring the effect of student status or employment status were restricted to children aged between 5 and 18 years and adults over 18, respectively. In the remaining models including all ages, age was adjusted as a categorical variable (<15, 15 to 65 and over 65 years). CRRs, Odds Ratios (ORs) and their associated 95% Credible Intervals are presented for all regression models. Here, we report estimates adjusted for age and gender (referred to as adjCRR or adjOR). Studies which collated contact-level data were used to assess assortativity of mixing by age and gender for different country-income strata by calculating the proportions of contacts made by participants that are male or female and those that belong to three broad age groups (children, adults, and older adults; Supplementary Text 3).

Ethics statement

All original studies included were approved by an institutional ethics review committee. Ethics approval was not required for the present study.

Acknowledgements

We would like to acknowledge the Fiji Ministry of Health and Medical Services for their contribution to the study set in Fiji, M. Elizabeth Halloran for sharing the Senegal data, and Nickson Murunga for processing the data request for the Kenyan survey.

Competing interests

M.A. has received research funding from Seqirus outside the submitted work. G.E.P. was employed by the Emmes Company while analyzing the Niakhar Senegal social contact network data included in this study. The Emmes Company was contracted to perform data cleaning and data analysis of the Niakhar, Senegal clinical trial data (but not the social contact network data) for this study before G.E.P. joined the Emmes Company (in October 2015). After G.E.P. joined the Emmes Company, the sole support from Emmes for this manuscript was in the form of salary support for G.E.P. All other

authors declare no conflicts of interest. Outside of the submitted work C.G.G. has received grants, contracts, or consulting fees from the following bodies: CDC, AHRQ, FDA, NIH, Campbell Alliance/Syneos Health, Sanofi, Pfizer and Merck.

Financial disclosures

A.M., P.W., P.G.T.W. and C.W acknowledge joint Centre funding from the UK Medical Research Council and DFID (MR/R015600/1). O.J.W. acknowledges funding from the UK Foreign Commonwealth and Development Office. K.O.K acknowledges support by CUHK Direct grant for research (2019.020), Health and Medical Research Fund (reference number: INF-CUHK-1, 17160302, 18170312), General Research Fund (reference number: 14112818), Early Career Scheme (reference number: 24104920) and Wellcome Trust (UK, 200861/Z/16/Z). P.J.D. was supported by a fellowship from the UK Medical Research Council (MR/P022081/1); this UK-funded award is part of the European and Developing Countries Clinical Trials Partnership 2 (EDCTP2) programme supported by the EU. E.F.G.N. holds an Australian Government Research Training Program Scholarship. F.M.R. receives funding from the Australian National Health and Medical Research Council, WHO, the Bill & Melinda Gates Foundation; Wellcome Trust, DFAT. M.M. acknowledges funding from the EPSRC through the EPSRC Centre for Doctoral Training in Modern Statistics and Statistical Machine Learning. J.D.S received funding for this work from the University of Washington and a grant from US National Institutes of Health, NIAID. C.G.G. declares funding from NIH (K24AI148459). G.E.P. was supported previously by General Medical Sciences / National Institute of Health U01-GM070749. G.E.P was employed by the Emmes Company while analyzing the Niakhar Senegal social contact network data included in this study. The Emmes Company was contracted to perform data cleaning and data analysis of the Niakhar, Senegal clinical trial data (but not the social contact network data) for this study before G.E.P. joined the Emmes Company (in October 2015). After G.E.P. joined the Emmes Company, the sole support from Emmes for this manuscript was in the form of salary support for G.E.P. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

Ajelli M, Litvinova M. 2017. Estimating contact patterns relevant to the spread of infectious diseases in Russia. *J Theor Biol* **419**:1–7. doi:10.1016/j.jtbi.2017.01.041

Bansal S, Read J, Pourbohloul B, Meyers LA. 2010. The dynamic nature of contact networks in infectious disease epidemiology. *http://mc.manuscriptcentral.com/tjbd* **4**:478–489. doi:10.1080/17513758.2010.503376

Béraud G, Kazmercziak S, Beutels P, Levy-Bruhl D, Lenne X, Mielcarek N, Yazdanpanah Y, Boëlle P-Y, Hens N, Dervaux B. 2015. The French Connection: The First Large Population-Based Contact Survey in France Relevant for the Spread of Infectious Diseases. *PLoS One* **10**:e0133203. doi:10.1371/journal.pone.0133203

Brauner JM, Mindermann S, Sharma M, Johnston D, Salvatier J, Gavenčiak T, Stephenson AB, Leech G, Altman G, Mikulik V, Norman AJ, Monrad JT, Besiroglu T, Ge H, Hartwick MA, Teh YW, Chindelevitch L, Gal Y, Kulveit J. 2021. Inferring the effectiveness of government interventions against COVID-19. *Science (80-)* **371**. doi:10.1126/SCIENCE.ABD9338

Bulfone TC, Malekinejad M, Rutherford GW, Razani N. 2021. Outdoor Transmission of SARS-CoV-2 and Other Respiratory Viruses: A Systematic Review. *J Infect Dis* **223**:550–561. doi:10.1093/infdis/jiaa742

Bürkner PC. 2018. Advanced Bayesian multilevel modeling with the R package brms. *R J* **10**:395–411. doi:10.32614/rj-2018-017

Bürkner PC. 2017. brms: An R package for Bayesian multilevel models using Stan. *J Stat Softw* **80**:1–

28. doi:10.18637/jss.v080.i01

Chang S, Pierson E, Koh PW, Gerardin J, Redbird B, Grusky D, Leskovec J. 2021. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* **589**:82–87. doi:10.1038/s41586-020-2923-3

Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, El-harakeh A, Bognanni A, Lotfi T, Loeb M, Hajizadeh A, Bak A, Izcovich A, Cuello-Garcia CA, Chen C, Harris DJ, Borowiack E, Chamseddine F, Schünemann F, Morgano GP, Muti Schünemann GEU, Chen G, Zhao H, Neumann I, Chan J, Khabsa J, Hneiny L, Harrison L, Smith M, Rizk N, Giorgi Rossi P, AbiHanna P, El-khoury R, Stalteri R, Baldeh T, Piggott T, Zhang Y, Saad Z, Khamis A, Reinap M. 2020. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *Lancet* **395**:1973–1987. doi:10.1016/S0140-6736(20)31142-9

Dahab M, van Zandvoort K, Flasche S, Warsame A, Ratnayake R, Favas C, Spiegel PB, Waldman RJ, Checchi F. 2020. COVID-19 control in low-income settings and displaced populations: what can realistically be done? *Confl Heal 2020 141* **14**:1–6. doi:10.1186/S13031-020-00296-8

Davies NG, Kucharski AJ, Eggo RM, Gimma A, Edmunds WJ, Jombart T, O'Reilly K, Endo A, Hellewell J, Nightingale ES, Quilty BJ, Jarvis CI, Russell TW, Klepac P, Bosse NI, Funk S, Abbott S, Medley GF, Gibbs H, Pearson CAB, Flasche S, Jit M, Clifford S, Prem K, Diamond C, Emery J, Deol AK, Procter SR, van Zandvoort K, Sun YF, Munday JD, Rosello A, Auzenbergs M, Knight G, Houben RMGJ, Liu Y. 2020. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. *Lancet Public Heal* **5**:e375–e385. doi:10.1016/S2468-2667(20)30133-X

De Negri F, Galiezz R, Miranda P, Koeller P, Zucoloto G, Costa J, Farias CM, Travassos GH, Medronho RA. 2021. Socioeconomic factors and the probability of death by Covid-19 in Brazil. *J Public*

Health (Bangkok) 1–6. doi:10.1093/pubmed/fdaa279

Delamater PL, Street EJ, Leslie TF, Yang YT, Jacobsen KH. 2019. Complexity of the basic reproduction number (R0). *Emerg Infect Dis* **25**:1–4. doi:10.3201/eid2501.171901

Dodd PJ, Looker C, Plumb ID, Bond V, Schaap A, Shanaube K, Muyoyeta M, Vynnycky E, Godfrey-Faussett P, Corbett EL, Beyers N, Ayles H, White RG. 2015. Age- and Sex-Specific Social Contact Patterns and Incidence of *Mycobacterium tuberculosis* Infection. *Am J Epidemiol* **183**:kwv160. doi:10.1093/aje/kwv160

Downes MJ, Brennan ML, Williams HC, Dean RS. 2016. Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open* **6**. doi:10.1136/bmjopen-2016

Dunne EM, Satzke C, Ratu FT, Neal EFG, Boelsen LK, Matanitobua S, Pell CL, Nation ML, Ortika BD, Reyburn R, Jenkins K, Nguyen C, Gould K, Hinds J, Tikoduadua L, Kado J, Rafai E, Kama M, Mulholland EK, Russell FM. 2018. Effect of ten-valent pneumococcal conjugate vaccine introduction on pneumococcal carriage in Fiji: results from four annual cross-sectional carriage surveys. *Lancet Glob Heal* **6**:e1375–e1385. doi:10.1016/S2214-109X(18)30383-8

Eubank S, Guclu H, Kumar VSA, Marathe M V., Srinivasan A, Toroczkai Z, Wang N. 2004. Modelling disease outbreaks in realistic urban social networks. *Nature* **429**:180–184. doi:10.1038/nature02541

Ferrari MJ, Bansal S, Meyers LA, Björnstad ON. 2006. Network frailty and the geometry of herd immunity. *Proc R Soc B Biol Sci* **273**:2743–2748. doi:10.1098/rspb.2006.3636

Firth JA, Hellewell J, Klepac P, Kissler S, Jit M, Atkins KE, Clifford S, Villabona-Arenas CJ, Meakin SR, Diamond C, Bosse NI, Munday JD, Prem K, Foss AM, Nightingale ES, Zandvoort K van, Davies NG, Gibbs HP, Medley G, Gimma A, Flasche S, Simons D, Auzenbergs M, Russell TW, Quilty BJ, Rees EM, Leclerc QJ, Edmunds WJ, Funk S, Houben RMGJ, Knight GM, Abbott S, Sun FY, Lowe R, Tully DC, Procter SR, Jarvis CI, Endo A, O'Reilly K, Emery JC, Jombart T, Rosello A, Deol AK, Quaife M, Hué S, Liu Y, Eggo RM, Pearson CAB, Kucharski AJ, Spurgin LG. 2020. Using a real-world network to model localized COVID-19 control strategies. *Nat Med* **26**:1616–1622. doi:10.1038/s41591-020-1036-8

Fontanet A, Cauchemez S. 2020. COVID-19 herd immunity: where are we? *Nat Rev Immunol*. doi:10.1038/s41577-020-00451-5

Fu Y, Wang D-W, Chuang J-H. 2012. Representative Contact Diaries for Modeling the Spread of Infectious Diseases in Taiwan. *PLoS One* **7**:e45113. doi:10.1371/journal.pone.0045113

Ghani AC, Aral SO. 2005. Patterns of sex worker - Client contacts and their implications for the persistence of sexually transmitted infections. *J Infect Dis* **191**:S34–S41. doi:10.1086/425276

Gimma A, Munday JD, Wong KL, Coletti P, Zandvoort K van, Prem K, group CC-19 working, Klepac P, Rubin GJ, Funk S, Edmunds WJ, Jarvis CI. 2021. CoMix: Changes in social contacts as measured by the contact survey during the COVID-19 pandemic in England between March 2020 and March 2021. *medRxiv* 2021.05.28.21257973. doi:10.1101/2021.05.28.21257973

Grijalva CG, Goeyvaerts N, Verastegui H, Edwards KM, Gil AI, Lanata CF, Hens N. 2015. A Household-Based Study of Contact Networks Relevant for the Spread of Infectious Diseases in the Highlands of Peru. *PLoS One* **10**:e0118457. doi:10.1371/journal.pone.0118457

Hoang T, Coletti P, Melegaro A, Wallinga J, Grijalva CG, Edmunds JW, Beutels P, Hens N. 2019. A Systematic Review of Social Contact Surveys to Inform Transmission Models of Close-contact Infections. *Epidemiology* **30**:723–736. doi:10.1097/EDE.0000000000001047

Horton KC, Hoey AL, Béraud G, Corbett EL, White RG. 2020. Systematic review and meta-analysis of sex differences in social contact patterns and implications for tuberculosis transmission and control. *Emerg Infect Dis*. doi:10.3201/eid2605.190574

Huang Y, Cai X, Zhang B, Zhu G, Liu T, Guo P, Xiao J, Li X, Zeng W, Hu J, Ma W. 2020. Spatiotemporal heterogeneity of social contact patterns related to infectious diseases in the Guangdong

Province, China. *Sci Rep* **10**:1–10. doi:10.1038/s41598-020-63383-z

Ibuka Y, Ohkusa Y, Sugawara T, Chapman GB, Yamin D, Atkins KE, Taniguchi K, Okabe N, Galvani AP. 2016. Social contacts, vaccination decisions and influenza in Japan. *J Epidemiol Community Health* **70**:162–167. doi:10.1136/jech-2015-205777

Jacquez JA, Simon CP, Koopman J, Sattenspiel L, Perry T. 1988. Modeling and analyzing HIV transmission: the effect of contact patterns. *Math Biosci* **92**:119–199. doi:10.1016/0025-5564(88)90031-4

Jarvis CI, Gimma A, van Zandvoort K, Wong KLM, Abbas K, Villabona-Arenas CJ, O'Reilly K, Quaife M, Rosello A, Kucharski AJ, Gibbs HP, Atkins KE, Barnard RC, Bosse NI, Procter SR, Meakin SR, Sun FY, Abbott S, Munday JD, Russell TW, Flasche S, Sherratt K, Eggo RM, Davies NG, Quilty BJ, Auzenbergs M, Hellewell J, Jombart T, Jafari Y, Leclerc QJ, Lowe R, Foss AM, Jit M, Deol AK, Hué S, Knight GM, Endo A, Prem K, Emery JC, Clifford S, Medley G, Funk S, Sandmann FG, Tully DC, Pearson CAB, Gore-Langton GR, Showering A, Houben RMGJ, Nightingale ES, Klepac P, Waterlow NR, Chan YWD, Rudge JW, Simons D, Diamond C, Williams J, Brady O, Liu Y, Edmunds WJ. 2021. The impact of local and national restrictions in response to COVID-19 on social contacts in England: a longitudinal natural experiment. *BMC Med* **19**:1–12. doi:10.1186/s12916-021-01924-7

Jarvis CI, Van Zandvoort K, Gimma A, Prem K, Auzenbergs M, O'Reilly K, Medley G, Emery JC, Houben RMGJ, Davies N, Nightingale ES, Flasche S, Jombart T, Hellewell J, Abbott S, Munday JD, Bosse NI, Funk S, Sun F, Endo A, Rosello A, Procter SR, Kucharski AJ, Russell TW, Knight G, Gibbs H, Leclerc Q, Quilty BJ, Diamond C, Liu Y, Jit M, Clifford S, Pearson CAB, Eggo RM, Deol AK, Klepac P, Rubin GJ, Edmunds WJ. 2020. Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. *BMC Med* **18**:1–10. doi:10.1186/s12916-020-01597-8

Keeling MJ, Eames KT. 2005. Networks and epidemic models. *J R Soc Interface* **2**:295–307.

doi:10.1098/RSIF.2005.0051

Khan K, Arino J, Hu W, Raposo P, Sears J, Calderon F, Heidebrecht C, Macdonald M, Liauw J, Chan A, Gardam M. 2009. Spread of a Novel Influenza A (H1N1) Virus via Global Airline Transportation. *N Engl J Med* **361**:212–214. doi:10.1056/nejmc0904559

Kiti MC, Kinyanjui TM, Koech DC, Munywoki PK, Medley GF, Nokes DJ. 2014. Quantifying Age-Related Rates of Social Contact Using Diaries in a Rural Coastal Population of Kenya. *PLoS One* **9**:e104786. doi:10.1371/journal.pone.0104786

Kumar S, Gosain M, Sharma H, Swetts E, Amarchand R, Kumar R, Lafond KE, Dawood FS, Jain S, Widdowson M-A, Read JM, Krishnan A. 2018. Who interacts with whom? Social mixing insights from a rural population in India. *PLoS One* **13**:e0209039. doi:10.1371/journal.pone.0209039

Kwok KO, Cowling B, Wei V, Riley S, Read JM. 2018. Temporal variation of human encounters and the number of locations in which they occur: A longitudinal study of Hong Kong residents. *J R Soc Interface* **15**. doi:10.1098/rsif.2017.0838

Kwok KO, Cowling BJ, Wei VWI, Wu KM, Read JM, Lessler J, Cummings DA, Malik Peiris JS, Riley S. 2014. Social contacts and the locations in which they occur as risk factors for influenza infection. *Proc R Soc B Biol Sci* **281**. doi:10.1098/rspb.2014.0709

Lapidus N, de Lamballerie X, Salez N, Setbon M, Delabre RM, Ferrari P, Moyen N, Gougeon M-L, Vely F, Leruez-Ville M, Andreoletti L, Cauchemez S, Boëlle P-Y, Vivier É, Abel L, Schwarzinger M, Legeas M, Le Cann P, Flahault A, Carrat F. 2013. Factors Associated with Post-Seasonal Serological Titer and Risk Factors for Infection with the Pandemic A/H1N1 Virus in the French General Population. *PLoS One* **8**:e60127. doi:10.1371/journal.pone.0060127

le Polain de Waroux O., Cohuet S, Ndazima D, Kucharski AJ, Juan-Giner A, Flasche S, Tumwesigye E, Arinaitwe R, Mwanga-Amumpaire J, Boum Y, Nackers F, Checchi F, Grais RF, Edmunds WJ. 2018. Characteristics of human encounters and social mixing patterns relevant to infectious diseases spread by close contact: A survey in Southwest Uganda. *BMC Infect Dis* **18**:172. doi:10.1186/s12879-018-3073-1

le Polain de Waroux Olivier, Flasche S, Kucharski AJ, Langendorf C, Ndazima D, Mwanga-Amumpaire J, Grais RF, Cohuet S, Edmunds WJ. 2018. Identifying human encounters that shape the transmission of Streptococcus pneumoniae and other acute respiratory infections. *Epidemics* **25**:72–79. doi:10.1016/j.epidem.2018.05.008

Leung K, Jit M, Lau EHY, Wu JT. 2017. Social contact patterns relevant to the spread of respiratory infectious diseases in Hong Kong. *Sci Rep* **7**:4–8. doi:10.1038/s41598-017-08241-1

Liu CY, Berlin J, Kiti MC, Fava E Del, Grow A, Zagheni E, Melegaro A, Jenness SM, Omer S, Lopman B, Nelson K. 2021. Rapid review of social contact patterns during the COVID-19 pandemic. *medRxiv* 2021.03.12.21253410. doi:10.1101/2021.03.12.21253410

Loayza N V. 2020. Costs and Trade-Offs in the Fight Against the COVID-19 Pandemic, Costs and Trade-Offs in the Fight Against the COVID-19 Pandemic. World Bank, Washington, DC. doi:10.1596/33764

Mahikul W, Kripattanapong S, Hanvoravongchai P, Meeyai A, Iamsirithaworn S, Auewarakul P, Pan-ngum W. 2020. Contact Mixing Patterns and Population Movement among Migrant Workers in an Urban Setting in Thailand. *Int J Environ Res Public Health* **17**:2237. doi:10.3390/ijerph17072237

Mbala-Kingebeni P, Aziza A, Di Paola N, Wiley MR, Makiala-Mandanda S, Caviness K, Pratt CB, Ladner JT, Kugelman JR, Prieto K, Chitty JA, Larson PA, Beitzel B, Ayouba A, Vidal N, Karhemere S, Diop M, Diagne MM, Faye M, Faye O, Aruna A, Nsio J, Mulangu F, Mukadi D, Mukadi P, Kombe J, Mulumba A, Villabona-Arenas CJ, Pukuta E, Gonzalez J, Bartlett ML, Sozhamannan S, Gross SM, Schroth GP, Tim R, Zhao JJ, Kuhn JH, Diallo B, Yao M, Fall IS, Ndjoloko B, Mossoko M, Lacroix A, Delaporte E, Sanchez-Lockhart M, Sall AA, Muyembe-Tamfum JJ, Peeters M, Palacios G, Ahuka-

Mundeke S. 2019. Medical countermeasures during the 2018 Ebola virus disease outbreak in the North Kivu and Ituri Provinces of the Democratic Republic of the Congo: a rapid genomic assessment. *Lancet Infect Dis* **19**:648–657. doi:10.1016/S1473-3099(19)30118-5

McCreesh N, Dlamini V, Edwards A, Olivier S, Dayi N, Dikgale K, Nxumalo S, Dreyer J, Baisley K, Siedner MJ, White RG, Herbst K, Grant AD, Harling G. 2021. Impact of the Covid-19 epidemic and related social distancing regulations on social contact and SARS-CoV-2 transmission potential in rural South Africa: analysis of repeated cross-sectional surveys. *BMC Infect Dis 2021 211* **21**:1–11. doi:10.1186/S12879-021-06604-8

Mikolajczyk RT, Kretzschmar M. 2008. Collecting social contact data in the context of disease transmission: Prospective and retrospective study designs. *Soc Networks* **30**:127–135. doi:10.1016/j.socnet.2007.09.002

Mistry D, Litvinova M, Pastore y Piontti A, Chinazzi M, Fumanelli L, Gomes MFC, Haque SA, Liu QH, Mu K, Xiong X, Halloran ME, Longini IM, Merler S, Ajelli M, Vespignani A. 2021. Inferring high-resolution human mixing patterns for disease modeling. *Nat Commun* **12**:1–12. doi:10.1038/s41467-020-20544-y

Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, Massari M, Salmaso S, Tomba GS, Wallinga J, Heijne J, Sadkowska-Todys M, Rosinska M, Edmunds WJ. 2008. Social Contacts and Mixing Patterns Relevant to the Spread of Infectious Diseases. *PLoS Med* **5**:e74. doi:10.1371/journal.pmed.0050074

Neal EFG, Flasche S, Nguyen CD, Ratu FT, Dunne EM, Koyamaibole L, Reyburn R, Rafai E, Kama M, Ortika BD, Boelsen LK, Kado J, Tikoduadua L, Devi R, Tuivaga E, Satzke C, Mulholland EK, Edmunds WJ, Russell FM. 2020. Associations between ethnicity, social contact, and pneumococcal carriage three years post-PCV10 in Fiji. *Vaccine* **38**:202–211. doi:10.1016/j.vaccine.2019.10.030

Oguz MM, Camurdan AD, Aksakal FN, Akcaboy M, Altinel Acoglu E. 2018. Social contact patterns of infants in deciding vaccination strategy: A prospective, cross-sectional, single-centre study. *Epidemiol Infect* **146**:1157–1166. doi:10.1017/S0950268818001048

Prem K, Cook AR, Jit M. 2017. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Comput Biol* **13**:e1005697. doi:10.1371/journal.pcbi.1005697

Prem K, Liu Y, Russell TW, Kucharski AJ, Eggo RM, Davies N, Flasche S, Clifford S, Pearson CAB, Munday JD, Abbott S, Gibbs H, Rosello A, Quilty BJ, Jombart T, Sun F, Diamond C, Gimma A, van Zandvoort K, Funk S, Jarvis CI, Edmunds WJ, Bosse NI, Hellewell J, Jit M, Klepac P. 2020. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. *Lancet Public Heal* **5**:e261–e270. doi:10.1016/S2468-2667(20)30073-6

Rader B, Scarpino S V., Nande A, Hill AL, Adlam B, Reiner RC, Pigott DM, Gutierrez B, Zarebski AE, Shrestha M, Brownstein JS, Castro MC, Dye C, Tian H, Pybus OG, Kraemer MUG. 2020. Crowding and the shape of COVID-19 epidemics. *Nat Med 2020 2612* **26**:1829–1834. doi:10.1038/s41591-020-1104-0

Routledge I, Epstein A, Takahashi S, Hakim J, Janson O, Duarte E, Turcios K, Vinden J, Sujishi K, Rangel J, Coh M, Besana L, Ho W-K, Oon C-Y, Ong CM, Yun C, Lynch K, Wu A, Wu W, Karlon W, Thornborrow E, Peluso M, Henrich T, Pak J, Briggs J, Greenhouse B, Rodriguez-Barraquer I. 2021. Citywide serosurveillance of the initial SARS-CoV-2 outbreak in San Francisco. *Res Sq*. doi:10.21203/rs.3.rs-180966/v1

Stein ML, van Steenbergen JE, Buskens V, van der Heijden PGM, Chanyasanha C, Tipayamongkholgul M, Thorson AE, Bengtsson L, Lu X, Kretzschmar MEE. 2014. Comparison of Contact Patterns Relevant for Transmission of Respiratory Pathogens in Thailand and the Netherlands Using

Respondent-Driven Sampling. *PLoS One* **9**:e113711. doi:10.1371/journal.pone.0113711

Thompson HA, Mousa A, Dighe A, Fu H, Arnedo-Pena A, Barrett P, Bellido-Blasco J, Bi Q, Caputi A, Chaw L, De Maria L, Hoffmann M, Mahapure K, Ng K, Raghuram J, Singh G, Soman B, Soriano V, Valent F, Vimercati L, Wee LE, Wong J, Ghani AC, Ferguson NM. 2021. SARS-CoV-2 setting-specific transmission rates: a systematic review and meta-analysis. *Clin Infect Dis*. doi:10.1093/cid/ciab100

United Nations Department of Economic and Social Affairs Population Division. 2019. Database on Household Size and Composition 2019. https://population.un.org/Household/index.html#/countries/840

Walker PGT, Whittaker C, Watson OJ, Baguelin M, Winskill P, Hamlet A, Djafaara BA, Cucunubá Z, Mesa DO, Green W, Thompson H, Nayagam S, Ainslie KEC, Bhatia S, Bhatt S, Boonyasiri A, Boyd O, Brazeau NF, Cattarino L, Cuomo-Dannenburg G, Dighe A, Donnelly CA, Dorigatti I, Van Elsland SL, FitzJohn R, Fu H, Gaythorpe KAM, Geidelberg L, Grassly N, Haw D, Hayes S, Hinsley W, Imai N, Jorgensen D, Knock E, Laydon D, Mishra S, Nedjati-Gilani G, Okell LC, Unwin HJ, Verity R, Vollmer M, Walters CE, Wang H, Wang Y, Xi X, Lalloo DG, Ferguson NM, Ghani AC. 2020. The impact of COVID-19 and strategies for mitigation and suppression in low- And middle-income countries. *Science (80-)* **369**:413–422. doi:10.1126/science.abc0035

Ward H, Cooke G, Whitaker M, Redd R, Eales O, Brown JC, Collet K, Cooper E, Daunt A, Jones K, Moshe M, Willicombe M, Day S, Atchison C, Darzi A, Donnelly CA, Riley S, Ashby D, Barclay WS, Elliott P. 2021. REACT-2 Round 5: increasing prevalence of SARS-CoV-2 antibodies demonstrate impact of the second wave and of vaccine roll-out in England. *medRxiv* 2021.02.26.21252512. doi:10.1101/2021.02.26.21252512

Winskill P, Whittaker C, Walker P, Watson O, Laydon D, Imai N, Cuomo-Dannenburg G, Ainslie K, Baguelin M, Bhatt S, Boonyasiri A, Cattarino L, Ciavarella C, Cooper L V, Coupland H, Cucunuba

Z, Van Elsland SL, Fitzjohn R, Flaxman S, Gaythorpe K, Green W, Hallett T, Hamlet A, Hinsley W, Knock E, Lees J, Mellan T, Mishra S, Nedjati-Gilani G, Nouvellet P, Okell L, Parag K V, Thompson HA, Juliette H, Unwin T, Vollmer M, Wang Y, Whittles L, Xi X, Ferguson N, Donnelly C, Ghani A. 2020. Report 22: Equity in response to the COVID-19 pandemic: an assessment of the direct and indirect impacts on disadvantaged and vulnerable populations in low-and lower middle-income countries. doi:10.25561/78965

Wood R, Racow K, Bekker L-G, Morrow C, Middelkoop K, Mark D, Lawn SD. 2012. Indoor Social Networks in a South African Township: Potential Contribution of Location to Tuberculosis Transmission. *PLoS One* **7**:e39246. doi:10.1371/journal.pone.0039246

World Bank Group. 2020. World Bank Country and Lending Groups – World Bank Data Help Desk. https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

World Bank International Comparison Programme. 2021. World Development Indicators database Eurostat-OECD PPP Programme.

Zhang J, Litvinova M, Liang Y, Wang Y, Wang W, Zhao S, Wu Q, Merler S, Viboud C, Vespignani A, Ajelli M, Yu H. 2020. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science (80-)* **368**:1481–1486. doi:10.1126/science.abb8001

Table 1- Summary table of total daily contacts. The total number of observations, as well as the mean, median and interquartile range (p25 and p75) of total daily contacts shown by participant and study characteristics.

			N	Mean	p25	Median	p75
Overall			28,503	14.5	5	9	17
Gender		**Male**	13,218	15.3	5	9	18
		Female	14,598	13.7	5	9	16
Age		**<15**	8,561	14.6	6	10	19
		15 to 65	17,841	14.9	5	9	17
		>65	2,047	10.4	3	6	12
Income status		**LIC/LMIC**	9,906	15.4	5	10	17
		UMIC	8,330	14.4	5	8	16
		HIC	10,267	13.7	5	9	17
Survey Methodology		**Diary**	12,226	13.9	6	10	18
		Interview/Survey	16,227	15.0	4	8	16
Day type		**Weekend**	4,308	14.7	5	9	16
		Weekday	21,579	14.1	5	9	17
Employment *(in those aged >18)*		**Yes**	8,879	15.4	5	9	17
		No	6,158	9.8	4	7	12
Student *(in those aged 5 to 18)*		**Yes**	4,438	18.4	8	14	24
		No	600	10.4	5	8	14
Household size		**1**	1,479	10.4	3	6	12
		2	3,220	11.8	4	7	14
		3	4,130	12.0	4	7	14
		4	5,240	13.4	5	8	17
		5	3,109	12.5	4	8	14
		6+	8,873	17.7	7	11	20
Study	**Belgium**	**Mossong**	750	11.8	5	9	15
	China	**Read**	1,821	18.6	7	13	22
	China	**Zhang**	965	18.8	4	10	30
	Fiji	**Neal**	2,019	6.4	4	6	8
	Finland	**Mossong**	1,006	11.1	5	9	15
	Germany	**Mossong**	1,341	7.9	4	6	10
	Hong Kong	**Kwok (2014)**	762	18.3	5	9	18
	Hong Kong	**Kwok (2018)**	1,066	11.9	3	7	13
	Hong Kong	**Leung**	1,149	14.4	3	7	15
	India	**Kumar**	2,943	27.0	12	17	26
	Italy	**Mossong**	849	19.8	10	17	27
	Kenya	**Kiti**	568	17.7	10	15	23
	Luxembourg	**Mossong**	1,051	17.5	8	14	24
	Netherlands	**Mossong**	269	13.9	6	11	19
	Peru	**Grijalva**	588	15.3	8	12	20
	Poland	**Mossong**	1,012	16.3	7	13	22.5
	Russia	**Ajelli**	502	18.0	6	11	19
	South Africa	**Dodd**	1,276	5.2	4	5	7
	South Africa	**Wood**	571	15.6	9	14	20
	Senegal	**Potter**	1,417	19.7	10	15	25
	Thailand	**Mahikul**	369	22.6	13	20	31
	Thailand	**Stein**	219	58.5	15	24	55
	Uganda	**Le Polain de Waroux**	568	7.0	5	7	9
	United	**Mossong**	1,012	11.7	6	10	16
	Vietnam	**Horby**	865	7.7	5	7	9
	Zambia	**Dodd**	2,300	4.8	3	4	6
	Zimbabwe	**Melegaro**	1,245	10.7	6	9	14

Figure 1 – Total number of contacts. Sample median total number of contacts shown by gender (right) and 5-year age groups up to ages 80+ shown for A) LICs/LMICs, B) UMICs and C) HICs. Grey lines denote individual studies, and the solid black line is the median across all studies of within that income group. Studies with a diary-based methodology are represented by a solid grey line and those with a questionnaire or interview design are shown as a dashed line. For UMICs, one study outlier with extremely high number of contacts is excluded (online Thai survey with a "snowball" design by Stein et al., 2014). Contact Rate Ratios and associated 95% Credible intervals from a negative binomial model with random study effects are shown in D (LICs/LMICs), E (UMICs) and F (HICs). All models were adjusted for age and gender and were ran separately for each key variable (weekday/weekend, household size, survey methodology, student/employment status).

Figure 2- Contact location and household size. A) Sample median number of contacts by household size in review data, stratified by income strata. Shaded area denotes the interquartile range. B) sample mean % of contacts made at each location (home, school, work, other) by income group. C) total daily contacts (sample mean number) made at each location by 5-year age group. D) Sample median number of contacts made at home by 5-year age groups and income strata. Shaded area denotes the interquartile range. E) Average household size and GDP; red circles represent median household size in single studies from the review. GDP information was obtained from the World Bank Group and global household size data from the Department of Economic and Social Affairs, Population Division, United Nations.

Figure 3- Physical contacts. Mean proportion of contacts that are physical shown by gender (right) and 5-year age groups up to ages 80+ shown for A) LICs/LMICs, B) UMICs and C) HICs. Grey lines denote individual studies, and the solid black line is the mean across all studies of within that income group. Studies with a diary-based methodology are represented by a solid grey line and those with a questionnaire or interview design are shown as a dashed line. Odds Ratios and associated 95% Credible intervals from a logistic regression model with random study effects are shown in D (LICs/LMICs), E (UMICs) and F (HICs). All models were adjusted for age and gender and were ran separately for each key variable (weekday/weekend, household size, survey methodology, student/employment status).

Figure 4- Contact duration. Mean proportion of contacts that last at least an hour shown by gender (right) and 5-year age groups up to ages 80+ shown for A) LICs/LMICs, B) UMICs and C) HICs. Grey lines denote individual studies and the solid black line is the mean across all studies of within that income group. Studies with a diary-based methodology are represented by a solid grey line and those with a questionnaire or interview design are shown as a dashed line. Odds Ratios and associated 95% Credible intervals from a logistic regression model with random study effects are shown in D (LICs/LMICs), E (UMICs) and F (HICs). All models were adjusted for age and gender and were ran separately for each key variable (weekday/weekend, household size, survey methodology, student/employment status).

Supplementary files

Supplementary Figure 1- PRISMA flow diagram of the screening process and selection of eligible studies.

Supplementary Figure 2- Total number of contacts boxplots and violin plots by participant/study characteristics

Supplementary Figure 3- The relationship between household size and median daily contacts made at home divided by a participant's household size, stratified by income strata

Supplementary Figure 4- Comparison of estimated regression coefficients for predicting total contacts with and without the inclusion of additional contacts.

Supplementary Figure 5- Location of contact by participant/study characteristics: A) Age, B) Gender, C) Day of the week, D) Household size, E) employment (in participants aged 18 or over) and F) Student status (in participants aged 5 to <20 years)

Supplementary Figure 6- Contact location and A) Type of contacts and B) Duration of contact

Supplementary Table 1 - Extraction table of study characteristics

Supplementary Table 2- Data availability by study

Supplementary Table 3- Risk of bias table (AXIS critical appraisal tool)

Supplementary Table 4 - Search string

Supplementary Table 5- PRISMA-IPD Checklist of items to include when reporting a systematic review and meta-analysis of individual participant data (IPD)

Supplementary Text 1- Systematic Review Findings

Supplementary Text 2 - Comparison of estimated regression coefficients in the main analysis and sensitivity analysis weighing each study equally within an income group.

Supplementary Text 3 - Assortativity by age and sex

Supplementary Text 4-Data assumptions and data dictionary.