OPEN

A Systematic Review of Social Contact Surveys to Inform Transmission Models of Close-contact Infections

Thang Hoang[a], Pietro Coletti[a], Alessia Melegaro[b], Jacco Wallinga[c,d], Carlos G. Grijalva[e], John W. Edmunds[f], Philippe Beutels[g], and Niel Hens[a,g]

Background: Researchers increasingly use social contact data to inform models for infectious disease spread with the aim of guiding effective policies about disease prevention and control. In this article, we undertake a systematic review of the study design, statistical analyses, and outcomes of the many social contact surveys that have been published.
Methods: We systematically searched PubMed and Web of Science for articles regarding social contact surveys. We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines as closely as possible.
Results: In total, we identified 64 social contact surveys, with more than 80% of the surveys conducted in high-income countries. Study settings included general population (58%), schools or universities (37%), and health care/conference/research institutes (5%). The largest number of studies did not focus on a specific age group (38%), whereas others focused on adults (32%) or children (19%). Retrospective (45%) and prospective (41%) designs were used most often with 6% using both for comparison purposes. The definition of a contact varied among surveys, e.g., a nonphysical contact may require conversation, close proximity, or both. We identified age, time schedule (e.g., weekday/weekend), and household size as relevant determinants of contact patterns across a large number of studies.
Conclusions: We found that the overall features of the contact patterns were remarkably robust across several countries, and irrespective of the study details. By considering the most common approach in each aspect of design (e.g., sampling schemes, data collection, definition of contact), we could identify recommendations for future contact data surveys that may be used to facilitate comparison between studies.

Keywords: Behavioral change; Contact data; Contact pattern; Contact surveys; Infectious diseases

(*Epidemiology* 2019;30: 723–736)

Submitted September 30, 2018; accepted May 24, 2019.

From the [a]Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Hasselt University, Agoralaan Gebouw D, Diepenbeek, Belgium; [b]Carlo F. Dondena Centre for Research on Social Dynamics and Public Policy, Bocconi University, Milano, Italy; [c]Centre for Infectious Disease Control, National Institute for Public Health, Bilthoven, The Netherlands; [d]Department of Biomedical Data Sciences, Leiden University, Leiden, The Netherlands; [e]Department of Health Policy, Vanderbilt University School of Medicine, Nashville, TN; [f]Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, United Kingdom; and [g]Centre for Health Economics Research and Modelling Infectious Diseases, Vaccine and Infectious Disease Institute, University of Antwerp, Antwerp, Belgium.

The results reported herein correspond to specific aims of grant 682540—TransMID to investigator N.H. from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation program.
The authors report no conflicts of interest.
T.V. Hoang and P. Coletti have equal contribution.
Reproducibility: All results reported in the manuscript can be reproduced by using searching queries provided in eAppendix; http://links.lww.com/EDE/B552.
SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Correspondence: Niel Hens, Interuniversity Institute for Biostatistics and statistical Bioinformatics, Hasselt University, Agoralaan Gebouw D, Diepenbeek, Belgium, 3590. E-mail: niel.hens@uhasselt.be.

ISSN: 1044-3983/19/3005-0723
DOI: 10.1097/EDE.0000000000001047

Despite the great progress in infectious disease control and prevention that were initiated during the last century, infectious pathogens continue to pose a threat to humanity, as illustrated by severe acute respiratory syndrome, influenza, antimicrobial resistant bacteria, Ebola, and resurgent measles, potentially disrupting everyday life, burdening public health, and occasionally dominating media headlines well into the 21st century.

Many infectious diseases can spread rapidly between people within and between age groups, households, schools, workplaces, cities, regions, and countries through a diversity of social contacts.[1]. Understanding and quantifying social mixing patterns is therefore of critical importance to establishing appropriate simulation models of the spread of infectious diseases. Such mathematical transmission models have become indispensable to guide health policy. Which interventions should be offered to which people in which circumstances? How would such interventions affect transmission chains and the disease burden throughout the population? What would be the population effectiveness and cost-effectiveness of such interventions? Well-informed answers to these questions require mathematical models. The validity of such models depends heavily on the appropriateness of their structure and their parameters, including what they assume about how people interact.

Indeed, a transmission model's integrated mixing patterns (i.e., who mixes with whom?) have a strong influence on the transmission parameters (i.e., who infects whom?). The latter are the most influential drivers for the outputs of such models. Whereas 20th-century models made strong assumptions about mixing patterns, it has become increasingly common to use empirical data on social interactions as a direct model input over the last decade.[2,3] For sexually transmitted infections, data from surveys on sexual behavior were available for use as an input for models. On the other hand, for infectious diseases that are transmitted by direct contact, minimal data on relevant social contacts was available. Edmunds et al[4] conducted a first study aimed to collect precisely this information using a convenience sample. This study was followed by a study that reported on relevant social contacts in a representative sample of the population that covered all ages in a city.[5] The landmark study that reported on relevant social contacts in representative samples for eight different European countries using contact diaries was the POLYMOD study.[6] Numerous other studies have been reported since. Several of these studies report on social mixing patterns as obtained through direct observation, contact diaries, or electronic proximity sensors. The strengths and weaknesses of these methods have been discussed.[7] Nevertheless, to our knowledge, a comprehensive review of the study designs for contact diaries and of major determinants of mixing patterns is lacking for this rapidly growing field of research, a gap which we aim to fill here.

In the current article, we systematically retrieve and review the literature on social contact surveys. First, we provide an overview of the literature to help identify a standard. Second, we present the different approaches for data collection and identify strengths and limitations. Third, we report on the main determinants of contact. We use these findings to guide future studies.

METHODS

We conducted a systematic review in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[8]

Search Strategy

We queried PubMed and ISI Web of Science, without time and language restriction up to 31 January 2018 using the following search string:

((([survey*] OR [questionnaire*] OR [diary] OR [diaries]) AND ([social contact*] OR [mixing behavio*] OR [mixing pattern*] OR [contact pattern*] OR [contact network*] OR [contact survey*] OR [contact data])

EndNote X7 was used to eliminate duplicates and manage the search results.[9]

Inclusion Criteria

We considered studies eligible if they fulfilled all of the following criteria: (1) primary focus on face-to-face contacts of humans, implying the physical presence of at least two persons during contact; (2) contacts relevant for the transmission of close-contact infections; (3) contacts recorded using a diary-type system on paper or in electronic format; (4) full-text version available.

Exclusion Criteria

We excluded studies that involved at least one of the following: (1) primary focus on human–animal or animal–animal contacts; (2) recording contacts exclusively relevant for sexually transmitted, food-, vector-, or water-borne diseases; (3) using exclusively proximity sensor devices or observational methods to collect contact data; (4) including contacts without physical presence (e.g., phone, internet/social media contacts) or without the possibility to distinguish them; (5) recording the frequency or regularity but not the number of contacts over a given time period; (6) meeting abstracts, books, theses, or unpublished journal articles.

An overview of the selection process is presented in Figure 1. Title, abstract, and full-text were screened initially by the first author and double-checked by the second author.

Data Extraction and Analysis

Three authors (T.V.H, P.C., N.H.) designed a data input form (see eTable 1; http://links.lww.com/EDE/B552). T.V.H. extracted relevant information from selected articles and inputted it in the form. P.C. performed verification to ensure data consistency and accuracy. We structured the data synthesis according to: (1) information on surveys and relevant articles: year and countries in which surveys were performed, authors and year of first publication, and relevant publications that used the same dataset; (2) information on survey's methodology: study setting, study subjects, final sample size, response rates, sampling scheme, data collection tools, collection modes, study design (prospective, retrospective or both), and reporting period; (3) information recorded on participants and contactees; (4) characteristics of contacts reported: types, definition, location, duration, and frequency of contacts; (5) analysis results: average and median number of contacts, SD, quantiles/range, and relevance of determinants for number of contacts. Data that could not be found in individual articles were given a value of "not available" for the corresponding variable.

This systematic review aggregates information of articles in the literature, so ethical review by an independent review board was unnecessary.

RESULTS

The Screening Process

Our search retrieved 1445 nonduplicate articles, with 73 suitable articles included in the review. Figure 1 shows the study selection process.

Country Settings

The 73 remaining articles covered 64 social contact surveys conducted in 24 countries spread over five continents: 12 European,[4–6,10–33] five Asian,[31,34–39] four African,[40–43] two American,[44–49] and one Oceanian[50–52] country. More details on number of social contact surveys in each of these countries are shown in the global map. Only 14 studies were

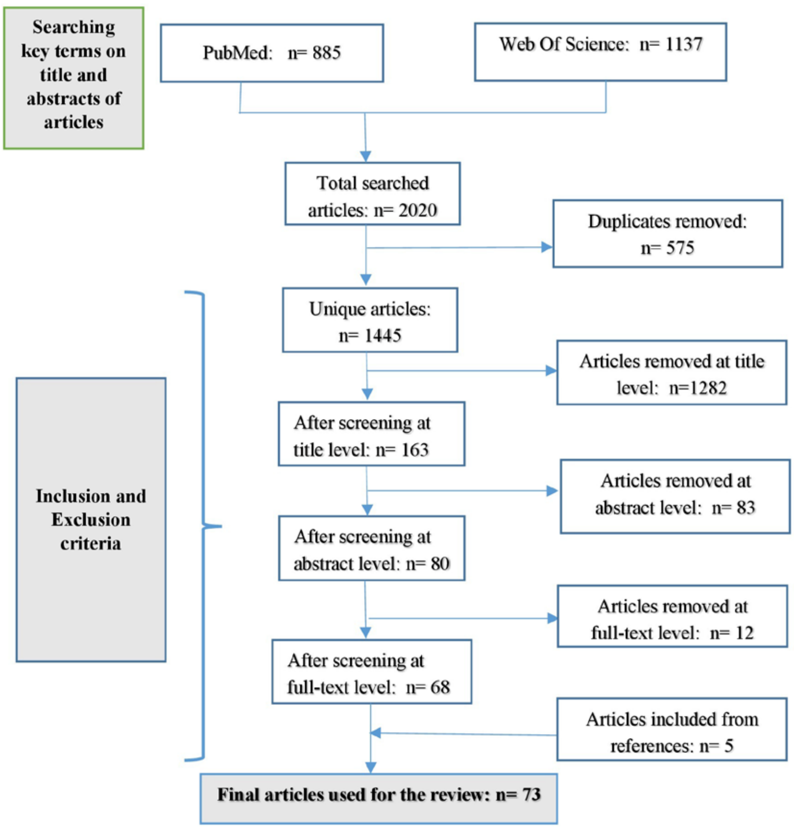

FIGURE 1. The PRISMA flowchart of the search process.

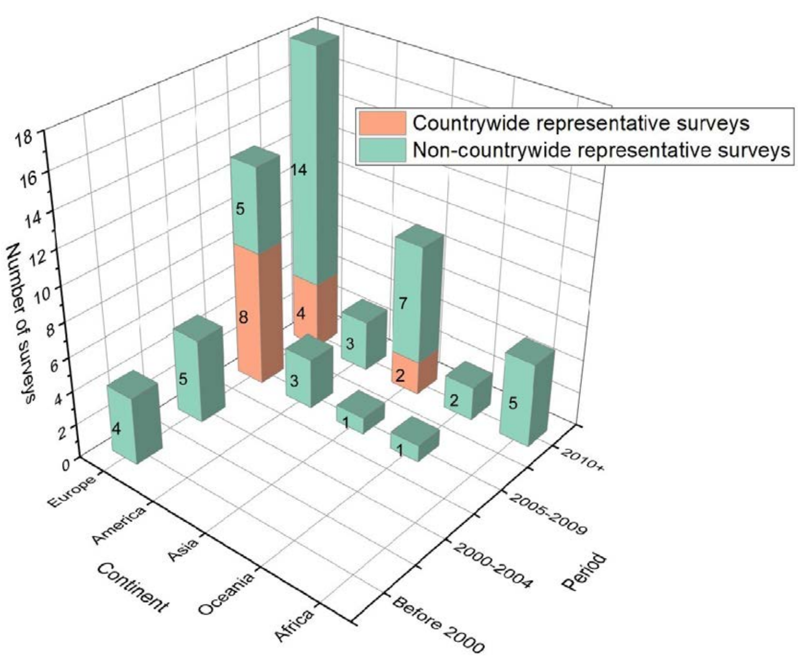

FIGURE 2. Distribution of number of representative surveys based on continents and time periods. Countrywide surveys cover the whole country in which they were conducted.

conducted at the whole-country level,[6,13,14,26,28,32,37,38] whereas remaining studies focused on a region,[12,34,35,39,41,43,47] a city or town,[5,36,40,42,51] or a specific setting (school or university, health care facility, etc.), and were therefore not representative of the entire country. Figure 2 demonstrates that 40 out of 64 of the surveys were conducted in Europe followed by Asia with 10 surveys. In contrast, only a few surveys were conducted in other regions. In this representation, we count several countries separately, even if they were included as part of a single project.[6,11,31,41]

The number of surveys greatly increased over time from only four surveys before the year 2000 up to 37 surveys after 2009, indicating that social contact surveys are increasingly conducted. In addition, no survey was conducted outside Europe before 2005. One survey did not indicate the year it was conducted. For this study, we used the publication year minus two as a proxy.[4]

Study Settings and Subjects

More than half of the social contact surveys were conducted in the community or general population (58%). Of these, there were only four household-based surveys that asked every member of each participating household to complete the survey.[14,35,39,44] The majority of surveys conducted in the general population aimed at people of all ages (65%).[6,12,14,26–28,32–34,36–40,42–44,47] In contrast, two surveys excluded infants younger than 1 year of age,[5] and one excluded children less than 2 years.[35] Four surveys focused exclusively on adults,[29,41,51] two investigated contact patterns of infants (under 11 weeks[25] and under 1 year old[52]), and one was aimed at patients with pandemic influenza AH1N1pdm09 (swine flu).[20] More specific settings of schools or universities constituted 38% of the surveys, of which 11 surveys were conducted at schools (primary schools,[16] secondary schools,[23] high schools,[48,49,53–56] or a combination of those[20,21,46]) and 13 surveys were performed in universities.[4,10,17,19,24,31,33] In addition to school or university settings, we also identified one contact survey on nurses in a health-care setting,[15] one survey at a conference,[18] and one survey in a research institute.[30]

Sample Size and Response Rate

Among social contact surveys conducted in the general population, the smallest survey only consisted of 54 participants in Switzerland,[29] and the largest survey consisted of 5388 participants in the United Kingdom.[26] The largest survey in a school/university setting contained 803 participants in Germany[17] (see eFigure 2; http://links.lww.com/EDE/B552). The response rate was reported in 36 out of 64 surveys and ranged from 4% in population-based surveys[26] up to 100% in a school-based survey.[17] Of these 34 surveys, only three considered the response rate beforehand to estimate the sample size.[15,38,47] Instead of considering the response rate, some surveys established criteria to replace those who refused or were not reached after several attempts.[39–41] Twenty surveys determined sampling weights based on demographic characteristics of the populations to reduce the effects of sample bias.[6,11–13,25,26,32,36–41,51]

Sampling Methods

Approximately half (44%) of the surveys employed convenience sampling, in which subjects were selected based on their convenient accessibility to researchers.[57] This sampling technique was also used for the sake of comparing data collection tools,[18,49,50,56,58] data collection methods,[59] or study designs.[10,17,50]

Seven surveys used random sampling.[5,26,28,34,35,51] Among these studies, only two surveys were considered representative of the entire country,[26,28] and the remaining surveys were representative of a region[34,35] or a city or town.[5,51] Three surveys employed multi-stage sampling,[15,38,39] and 10 surveys stratified sampling[12,14,25,37,40–42,47] that are easier to implement with respect to random sampling and can still remain representative. In addition, 10 surveys relied on quota sampling, which aimed to represent certain characteristics of a population (e.g., age, sex, geography, etc). Of these surveys, nine were conducted at the whole-country level,[6,13,32] and one survey focused on one specific region.[36] In addition, one survey used mixed samplings, in which a convenience sample of students was obtained in two schools, and a random sample of the general population was obtained in one province.[33] Five surveys used an online respondent-driven method, which can be considered as a snowball or chain sampling technique.[31,45,60] Only one survey did not state information on sampling techniques.[34] Finally, three surveys conducted at the general population level used a convenience sample,[29,44,50] therefore not relying on a sampling frame. More details on the distribution of sampling schemes based on time and regions are presented in eFigure 3; http://links.lww.com/EDE/B552.

Study Design

By prospective design, we mean that respondents are informed in advance of the day(s) that they are requested to record their contacts.[6,17,32] In a retrospective design, respondents recall their contacts over a past time period without prior warning or instruction that they would be requested to do so. Of 64 surveys, 29 (45%) used a retrospective design and 26 (41%) used a prospective design. Only four surveys 6%) used both designs for the purpose of comparison.[10,17,50] For five surveys (8%), it was not completely clear whether the study was prospective or retrospective.[20,24,30,48] eFigure 4; http://links.lww.com/EDE/B552 displays the trend of using study designs in social contact surveys over time, revealing that the retrospective design was more favored by researchers, except in the period 2005–2009 in which eight prospective surveys were implemented.[6,32]

Definition of Contact

Forty (63%) surveys distinguished physical and nonphysical contacts. Physical contacts were consistently defined as involving any sort of skin-to-skin touching (e.g., handshake, hug, kiss, etc). The definition of nonphysical contacts differed somewhat among surveys. Specifically, the majority of surveys using two types of contacts defined a nonphysical contact as a two-way conversation of at least three words at a distance that does not require raising one's voice.[6,10,36,39,43,50,52,54,59] In some other surveys, the definition involved close proximity (e.g., verbal communication made within 2 m) without specification of a minimum number of words to be exchanged.[13,26,30,38,41,61] Of note, that since the POLYMOD contact studies were executed,[6] its contact definition was applied in several subsequent surveys.[33,36,39,43,44,52,55] Fifteen surveys used only one type of contact, either involving a face-to-face conversation[4,5,16,23,33,45] or being in close proximity,[31,48,56] both regardless of any skin-to-skin touching[37,49] or only involving direct skin-to-skin touching.[28,40] Only one survey attempted to record casual contacts occurring in an indoor location without the requirement for a conversation or any type of touching.[42] Eight remaining surveys added kissing or intimate sexual contact as different types of contacts[17,19,29,46] or asked respondents to record contacts made in small/large groups or occasional contacts within 2 m in local transportation or crowded places separately.[17]

Reporting Time Period

Greater than half the surveys asked respondents to report contacts they made during a single day, whereas only six surveys used a reporting time period of greater than 3 days. The longest time period identified is 3 days in a prospective survey[19,50] and ten weeks in a retrospective survey.[45] Seven surveys recorded both weekdays and weekend days on the same respondents.[6,10,17,19,52–54] Finally, Eames et al[20] quantified the changes in social contact patterns experienced by individuals experiencing an episode of Influenza A(H1N1) on two randomly assigned days: one day while being ill and one day when recovered.

Characteristics of Participants and Contactees

Most surveys collected a range of demographic background characteristics of study participants (e.g., age, sex, education, household size). Some surveys also asked participants to record any influenza-like-illness symptoms they experienced on the day of surveying[31,47,53,60] or whether their day was in any way special (due to holiday, sickness, etc.).[6,12]

Among the characteristics of contactees, age and sex are considered to be the most important determinants of the mixing patterns given that they can help explain age and sex differences in the epidemiology of infectious diseases.[13] Thirty-six of 64 (56%) surveys recorded both age and sex of contactees, and 16 (25%) surveys recorded only age of contactees. In contrast, seven (11%) surveys required participants to simply report the number of different contactees without recording any of their characteristics.[17,29,45,47,49] In five surveys (8%), it was not clear what contactee characteristics participants had to report.[24,26,30,35,48] Along with age and sex, several surveys also asked participants to record health status of contactees and any symptoms they experienced, e.g., coughing, sneezing, fever, etc.[45,53–55] or whether they wore a protective mask.[53,54]

Information About Contacts

Participants were asked to record information about location, duration, and frequency of each contact in 77%, 67%, and 52% of contact surveys, respectively. All these contact characteristics were jointly recorded in 27 surveys (42%). For more details on the number of surveys considering these informations, see eFigure 5; http://links.lww.com/EDE/B552.

Mean Number of Contacts and Analysis of Determinants

Of the 64 surveys, 45 explicitly reported the average number of contacts measured without any stratification (Figure 3). To compare these survey results, we categorized them into 12 groups with different extents of representativeness (for country, region, or town or city), study designs and settings. In country-wide prospective surveys, the average number of contacts ranges from a minimum of 7.95 (95% confidence interval [CI] = 7.61, 8.29) in Germany[6] to a maximum of 26.97 (95% CI = 25.05, 28.89) in the United Kingdom.[26] In country-wide retrospective surveys, these values ranges from 12.5 (95% CI = 12.09, 12.91) in Taiwan[38] to 15.3 (95% CI = 14.4, 16.3) in Japan.[37] Six surveys conducted in the general population asked participants not to report details of professional contacts in the diary but to provide an estimate of the number and age distribution if they had more than 10 contacts (surveys in Finland,[6] Germany,[6] and the Netherlands[6]) or more than 20 professional contacts (surveys in Belgium[6,12] and France[13]). The additional professional contacts are not included in calculation of means presented in Figure 3. Among school- or university-based surveys, the highest number of contacts (70.3 [95% CI = 63.23, 77.37]) was observed in a secondary school in the United Kingdom.[23]

Figure 4 presents the number of surveys that have analyzed possible determinants for the number of contacts. For every determinant, we report the number of surveys that identified a relevant connection with the number of contacts (“yes”), the number of surveys that did not identify such a connection (“no”) and the number of surveys that did not delve into the matter. Strong evidence is identified for whether the age (34 yes vs. five no) and the household size (21 yes vs. four no) of the participant affected the number of contacts. Only five surveys identified sex as a relevant indicator for the number of social contacts, in contrast to 23 surveys that did not identify a relation. Social contacts are also affected by the daily routine (29 yes vs. six no) with a larger number of contacts during weekdays compared with the weekend (eFigure 6a; http://links.lww.com/EDE/B552 with the exception of Ref. 36). Similar results hold for term time versus holidays with all of the eight surveys that analyzed the issue identifying a larger number of contacts during term time (eFigure 6b; http://links.lww.com/EDE/B552). In addition, a self-reported healthy status is associated (five yes vs. one no) with a larger number of contacts with respect to feeling ill (eFigure 6c; http://links.lww.com/EDE/B552).

The relationship between social contacts and urbanization has been analyzed in three surveys. One survey found a larger number of contacts in peri-urban areas compared with rural areas,[43] one found the opposite,[40] and one did not find any evidence.[35] Finally, two surveys analyzed contacts during and outside flu seasons. Of these surveys, one survey[47] used a model to adjust for other factors (e.g., age and sex) and one did not.[55] However, both surveys identified no relevant effect (eFigure 6d; http://links.lww.com/EDE/B552).

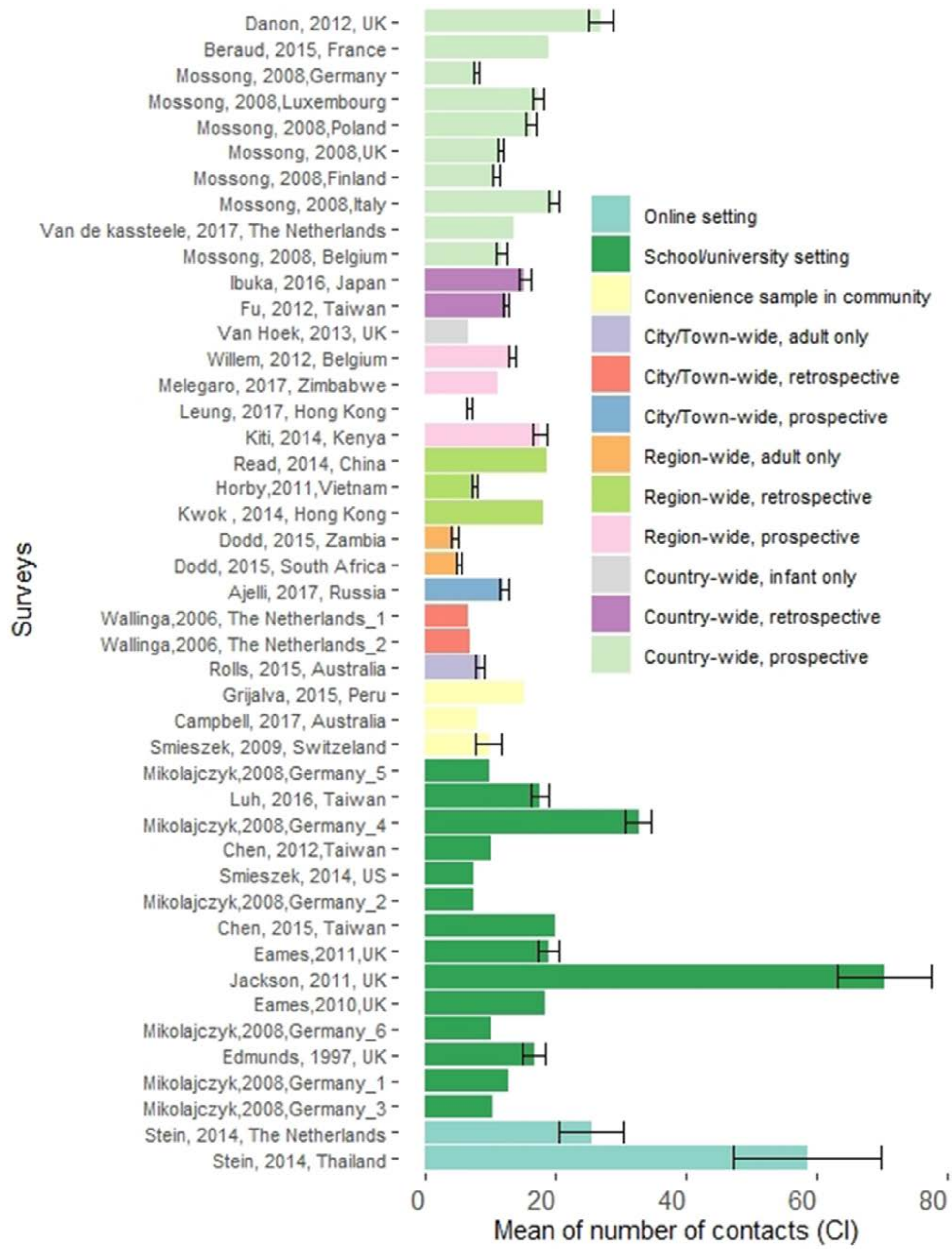

FIGURE 3. Average number of contacts measured and 95% CI. For surveys reporting mean number and SD a 95% CI for the mean was computed. Surveys are labeled according to the publication's first author, year and to the country in which the survey was performed. Ordering is performed based on increasing sample size within the specific design strata.

The Table provides summaries of all 64 social contact surveys.

DISCUSSION AND CONCLUSIONS

Social contact surveys are increasingly used to collect empirical data on human contact behavior and provide crucial inputs for mathematical models of infectious disease transmission. The POLYMOD project[6] presented the first large scale representative population surveys conducted in eight European countries. It also shared know–how both for data collection and analysis.

To date, most of these contact surveys were conducted in high- and middle-income countries, whereas low-income countries, which have a higher burden of communicable diseases, were less studied in this respect. For this reason, there is a need to continue studying contact patterns more widely and in particular in low- and middle-income settings. It is also worth noting that in low- and middle-income countries, the choice to perform a general population representative survey may be less meaningful, given the large variety of different settings (urban, rural, etc.) that are simultaneously present.

Most surveys did not clearly present sample size calculations, so we do not know to what extent important parameters, e.g., population size, confidence level, and margin of error, were taken into account.[62] Sample size estimation is even more important when one wants to compare social

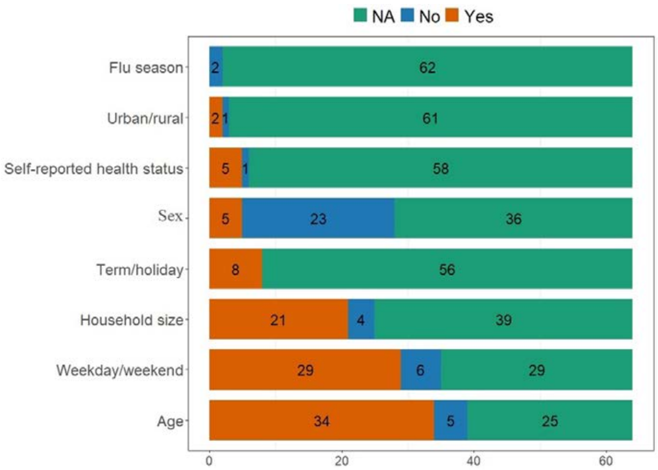

FIGURE 4. Determinants of number of social contacts. Surveys are tagged as "Yes" if a relevant connection between the number of contacts and the determinant was identified, "No" if evidence was not identified, or "NA" if the given determinant was not analyzed.

contact surveys between or among populations. Given the lack of a clear picture regarding which demographic and anthropologic factors are relevant in shaping contact patterns, inherent factors that may drive contact patterns are difficult to account for, thus making comparison among large populations, e.g., countries, even more difficult. It should be noted, however, that this is a relevant issue only when considering the contact matrix. Given the intrinsic network nature of social contact data more in general, sample size calculation becomes more involved and relies heavily on assumptions about the structural characteristic of the network and that no standard method to do so has emerged so far.[63,64] Indeed, the extensive analysis of this review has underlined a general lack of information on response rates, and a call for a better nonresponse analysis emerges as a guideline for future studies.

The prospective design is subject to less recall bias than the retrospective design. This notion can be partly explained by the fact that respondents in the former are informed in advance about which days they will be assigned for reporting their contact information. Furthermore, they are also asked to keep a diary with them and finish reporting before the surveying day is elapsed. Thus, the prospective design requires more commitment from respondents. In return, a prospective design can obtain more reported contacts compared with retrospective design.[17,50] However, large-scale studies are needed to further confirm these conclusions.

The use of self-reported diaries (paper or online) is the most commonly employed method in social contact surveys, associated with a smaller response rate with respect to e.g., face-to-face interviews. On the other hand, face-to-face interviews are more demanding in terms of fieldwork and data collection. No clear relationship in the number of contacts has been found when comparing online diaries with paper diaries,[10,36] while proximity sensors are more accurate in measuring short duration contacts, with the overall interaction pattern being similar between sensor and diaries[18,56] (for more details about data collection, see eAppendix; http://links.lww.com/EDE/B552). It should also be noted that proximity sensors usually perform a complete sampling of the interaction networks, whereas diary-based survey usually implement egocentric sampling (but complete sampling is not excluded in principle[14,35,39,44]). Although egocentric sampling does not allow estimation of several important network features, it still does not bias inference results, if properly taken into account.[65,66]

The definition of a potentially infectious contact is of crucial importance given that it will be used as a surrogate for exposure to disease.[15] Contact definitions in most surveys cannot capture potential risks from all forms of transmission modes,[30] such as fomite or indirect contact. Even for droplet transmission, using a face-to-face conversation definition as the basis for recording nonphysical contacts might lead to underreporting potentially infectious events given that susceptible individuals are likely able to contract a respiratory infection by just standing next to infected individuals who are, for example, sneezing or coughing. Furthermore, it seems even more challenging to record common touching frequency of shared material objects, such as doorknobs, water taps, etc., with which one person may be able to infect another indirectly. Indeed, the more details on potentially infectious events we attempt to collect, the greater the burden we impose on respondents. However, it seems reasonable in future studies to consider at least two contact definitions, one that involves physical contact and one that does not.

It is tempting to ask study participants to report their contacts as long as possible to gain insights into day-to-day variation. Nevertheless, the demanding task of diary–keeping may prevent many participants from recording the information for a long time in prospective studies.[67] Béraud et al[13] demonstrated that participants reported 6% (1%–10%) fewer contacts on the second day of the survey. In addition, the more contacts they reported on the first day, the larger the proportional decrease in number of contacts on the second day. For retrospective studies, a longer reporting period implies a longer recall period, with an associated larger bias. Therefore, in retrospective studies, researchers should try not to overstretch the reporting period.

This review provides information on the most relevant determinants of social contacts identified in previous studies. Asking study participants to report too many characteristics of contactees imposes a burden on participants. When designing future surveys, it is therefore important to consider which characteristics may be sufficiently relevant to include as determinants. For example, collecting age of the participants and their contacts is informative, as some studies revealed that using age-related mixing patterns helped explain observed serological and infection patterns of infectious diseases like

TABLE. Summaries of Social Contact Surveys

Year Survey Ended	Countries	Authors of First Publication	Year of First Publication	Study Settings	Study Subjects	Final Sample Size	Response Rates
1986	The Netherlands	Wallinga et al[5]	2006	General population	All excluding 0–1 year	2106	68.3%
1986	The Netherlands	Wallinga et al[5]	2006	General population	Age 1–70	1493	29.9%
1995	United Kingdom	Edmunds et al[4]	1997	University	Adult (staff, students and their families and friends)	65	70.6%
1998	United Kingdom	Read et al[24]	2008	University	Staff and students	49	NA
2003	Belgium	Beutels et al[10]	2006	University	Students and staff	73	NA
2003	Germany	Mikolajczyk et al[17]	2008	University	Students	38	50%
2004	Germany	Mikolajczyk et al[17]	2008	University	Students	196	100%
2004	Germany	Mikolajczyk et al[17]	2008	University	Students	28	100%
2004	Germany	Mikolajczyk et al[16]	2008	Primary school	Pupils	235	79.4%
2005	United Kingdom	Edmunds et al[19]	2006	University	Students	29	82.9%
2005	Germany	Mikolajczyk et al[17]	2008	University	Students	803	69%
2006	Germany	Mikolajczyk et al[17]	2008	University	Students	115	100%
2006	Belgium	Mossong et al[6]	2008	General population	People of all ages	750	NA
2006	Finland	Mossong et al[6]	2008	General population	People of all ages	1006	NA
2006	Germany	Mossong et al[6]	2008	General population	People of all ages	1341	NA
2006	Great Britain	Mossong et al[6]	2008	General population	People of all ages	1012	NA
2006	Italy	Mossong et al[6]	2008	General population	People of all ages	849	NA
2006	Luxembourg	Mossong et al[6]	2008	General population	People of all ages	1051	NA
2006	Poland	Mossong et al[6]	2008	General population	People of all ages	1012	NA
2006	The Netherlands	Mossong et al[6,a]	2008	General population	People of all ages	269	NA
2006	The Netherlands	van de Kassteele et al.[32,a]	2017	General population	People of all ages	825	NA
2007	Germany	Bernard et al[15]	2009	Hospital	Nurses	131	82%
2007	United States	Glass et al[46]	2008	School	Pupils	249	NA
2007	Vietnam	Horby et al[39]	2011	General population	People of all ages	865 /264 Hhs	NA
2008	United States	DeStefano et al[47]	2011	General population	People of all ages	4135	21.7%
2008	Australia	Mccaw et al[50]	2010	General population	University employees	65	NA
2009	United States	Potter et al[48]	2012	High schools	Pupils	246	57.9%
2009	Switzeland	Smieszek et al[29]	2009	General population	Adult	54	NA
2010	Taiwan	Chen et al[53]	2012	Junior high school	Pupils	274	67%
2010	Great Britain	Danon et al[26]	2012	General population	People of all ages	5388	3.8%
2010	United Kingdom	Eames et al[20]	2010	Schools	Pupils	119	10.8%
2010	United Kingdom	Eames et al[21]	2011	Primary and secondary schools	Pupils	135	12.3%
2010	United Kingdom	Eames et al[20]	2010	General population	Patients	317	10.6%
2010	United Kingdom	Eames et al[22]	2012	General population	Internet users	3338	NA

Sampling Scheme	Data Collection Tools	Collection Modes	Study Design	Contact Types	Recorded Information
Random sampling	Paper diary	Face-to-face interview	Retrospective	1	
Random sampling	Paper diary	Self-report	Retrospective	1	
Convenience sampling	Paper diary	Self-report	Prospective	1	Location
Convenience sampling	Paper diary	Self-report	NA	2	Location
Convenience sampling	Paper diary and online diary	Self-report on paper and on web interface	Both retrospective and Prospective	2	Location
Convenience sampling	Paper diary	Self-report	Retrospective	6	Duration, frequency
Convenience sampling	Paper diary	Self-report	Retrospective	6	Duration, frequency
Convenience sampling	Paper diary	Self-report	Retrospective	6	
Convenience sampling	Paper diary	Self-report	Retrospective	1	Duration, frequency
Convenience sampling	Paper diary	Self-report	Prospective	4	
Convenience sampling	Paper diary	Self-report	Retrospective	6	Location
Convenience sampling	Paper diary	Self-report	Both retrospective and prospective	6	Duration, frequency
Quota sampling by age, sex, geographical region, and rural/urban	Paper diary	Self-report	Prospective	2	Location, duration, frequency
Quota sampling by age and sex	Paper diary	Self-report	Prospective	2	Location, duration, frequency
Quota sampling by age	Paper diary	Self-report	Prospective	2	Location, duration, frequency
Quota sampling by age, sex, and geographical region	Paper diary	Self-report	Prospective	2	Location, duration, frequency
Quota by age, sex, geographical region, rural/urban, and day of the week	Paper diary	Self-report	Prospective	2	Location, duration, frequency
Quota sampling by age and sex	Paper diary	Self-report	Prospective	2	Location, duration, frequency
Quota sampling age, sex, geographical region, and rural/urban	Paper diary	Self-report	Prospective	2	Location, duration, frequency
Quota sampling by age and geographical region	Paper diary	Self-report	Prospective	2	Location, duration, frequency
Quota sampling based on age and geographical region	Paper diary	Self-report on paper	Prospective	2	Location, duration and frequency
Multi-stage sampling	Paper diary	Self-report	Prospective	2	Location, duration, frequency
Convenience sampling	Paper diary	Self-report	Retrospective	4	
Multi-stage sampling	Paper diary	Face-to-face interview	Retrospective	2	Location, duration, frequency
Stratified sampling	Electronic questionnaire	Computer assisted telephone interview	Retrospective	2	Location
Convenience sampling	Paper diary and A hand-held electronic diary	Self-report	Both retrospective and prospective	2	Location, duration
Convenience sampling	Paper diary	Self-report	NA	1	
Convenience sampling	Paper diary	Self-report	Retrospective	3	Duration
Convenience sampling	Paper diary	Self-report	Prospective	2	Location, duration, frequency
Random sampling	Paper diary and web-based questionnaires	Self-report	Prospective	2	Location, duration, frequency
Convenience sampling	Paper diary	Self-report	NA	2	Location, duration, frequency
Convenience sampling	Paper diary	Self-report	Prospective	2	Location, duration, frequency
Convenience sampling	Paper diary	Self-report	NA	2	Location, duration, frequency
Convenience sampling	Online diary	Self-report on web interface	Retrospective	2	Location

(*Continued*)

TABLE. (*Continued*)

Year Survey Ended	Countries	Authors of First Publication	Year of First Publication	Study Settings	Study Subjects	Final Sample Size	Response Rates
2010	Taiwan	Fu et al[38]	2012	General population	People of all ages	1943	46.2%
2010	United Kingdom	Jackson et al[23]	2011	Secondary schools	Pupils	107	83.6%
2010	South Africa	Johnstone et al[42]	2011	General population (a town)	People of all ages	571	77.4%
2010	Hong Kong	Kwok et al[34]	2014	General population	People of all ages	770	NA
2010	France	Lapidus an et al[14]	2012	General population	People of all ages	1377/601 hhs	NA
2010	China	Read et al[35]	2014	General population	Greater than 2 years old	1821/856 hhs	NA
2010	Switzerland	Smieszek et al[30]	2012	Research Institute	Staffs	50	NA
2011	Australia	Campbell et al[52]	2017	General population	Mothers with an infant	220	97.34%
2011	South Africa	Dodd et al[41]	2015	General population	Adults	1272	NA
2011	Zambia	Dodd et al[41]	2015	General population	Adults	2256	NA
2011	Peru	Grijalva et al[44]	2015	General population	People of all ages	588/114 hhs	100%
2011	Japan	Ibuka et al[37]	2016	General population	People of all ages	3146	NA
2011	Belgium	Willem et al[12]	2012	General population	People of all ages	1752	NA
2012	France	Beraud et al[13]	2015	General population	People of all ages	2033	51.1%
2012	Kenya	Kiti et al[40]	2014	General population	People of all ages	568	50.1%
2012	United States	Smieszek et al[49]	2014	High schools	Pupils, teachers and other staff	256	26.3%
2012	Sweden	Stromgren et al[28]	2017	General population	People of all ages	694	43.8%
2012	United Kingdom	Van Hoek et al[25]	2013	General population	Infant (under 11 weeks)	115	11.5%
2013	United States	Aiello et al[45]	2016	University/ residence halls	Students	590	NA
2013	Taiwan	Chen et al[54]	2015	Junior high school	Pupils	150	44%
2013	Taiwan	Luh ad et al[55]	2016	Junior high school	Pupils	373	44.37% ÷ 66.26%
2013	France	Mastrandrea et al[56]	2015	High school	Students	120	31.7%
2013	Zimbabwe	Melegaro et al[43]	2017	General population	People of all ages	1245	86.3%
2013	Australia	Rolls et al[51]	2015	General population	Adults	1307	33.5%
2013	Thailand	Stein et al[31]	2014	University	Students and their friends	220	85.6%
2013	The Netherlands	Stein et al[31]	2014	University	Students and their friends	322	89.9%
2014	Germany	Smieszek et al[18]	2016	Conference	Adult	74	24.7%
2014	The Netherlands	Stein et al[11]	2015	General population	Internet users	1451	NA
2014	Belgium	Stein et al[11]	2015	General population	Internet users	109	NA
2015	Hong Kong	Lyung et al[36]	2017	General population	People of all ages	1149	NA
2016	Russia	Ajelli et al[33]	2017	General population	People of all ages	505	NA

[a]As the multi-country survey was published while data collection was still on-going in the Netherlands, only 269 of these 825 participants were included in the study published by Mossong et al.[6] These 269 participants were also included in the study published by van de Kassteele et al.[32] Therefore, even if these studies are listed separately in this table, only one survey is counted.

Sampling Scheme	Data Collection Tools	Collection Modes	Study Design	Contact Types	Recorded Information
Three-stage systematic probability sampling	Paper diary	Face-to-face interview	Retrospective	2	Location, duration, frequency
Convenience sampling	Paper diary	Self-report	Retrospective	1	Location
Stratified sampling by age groups	Paper diary	Self-report	Prospective	3	Location, duration, frequency
Random Sampling	Paper diary	Face-to-face interview	Retrospective	2	Location, duration
Stratified sampling by geography	Paper diary	Face-to-face interview	Retrospective	2	Location, duration
Random sampling of hhs	Paper diary	Face-to-face interview	Retrospective	2	Location, duration, frequency
Convenience sampling	Paper diary	Self-report	NA	2	Duration
Convenience sampling	Paper diary	Self-report on paper	Prospective	2	Location, duration
Stratified sampling	Paper diary	Face-to-face interview	Retrospective	2	Location, duration, frequency
Stratified sampling	Paper diary	Face-to-face interview	Retrospective	2	Location, duration, frequency
Convenience sampling	Paper diary	Face-to-face interview	Retrospective	2	Location, duration, frequency
Convenience sampling	Online and paper diaries	Self-report	Retrospective	1	Location, duration
Stratified sampling by age and geography	Paper diary	Self-report	Prospective	2	Location, duration, frequency
Quota sampling for age, sex, days of week, and school holiday	Paper diary	Self-report	Prospective	2	Location, duration, frequency
Stratified random sampling	Paper diary (Text and pictorial diary)	Self-report	Prospective	1	Location, frequency
Convenience sampling	Online diary and proximity sensors	Self-report on web interface	Retrospective	1	Duration
Random sampling	Paper diary	Telephone interview	Prospective	1	Location
Random sampling	Paper diary	Self-report	Prospective	2	Location, duration, frequency
Randomized clusters of residence halls, snowball sampling	Online diary	Self-report on web interface	Retrospective	1	Location, duration
Convenience sampling	Paper diary	Self-report	Prospective	2	Location, duration and frequency
Convenience sampling	Paper diary	Self-report	Prospective	2	Location, duration and frequency
Convenience sampling	Paper diary and proximity sensors	Self-report	Prospective	1	Duration
Stratified sampling	Paper diary	Self-report on paper	Prospective	2	Location
Random sampling	Electronic questionnaire	Computer assisted telephone interview	Retrospective	2	Location, duration
Online respondent-driven sampling	Online diary	Self-report on web interface	Retrospective	1	Location
Online respondent-driven sampling	Online diary	Self-report on web interface	Retrospective	1	Location
Convenience sampling	Paper diary and proximity sensors	Self-report and wearable proximity sensor	Prospective	2	Duration
Online respondent-driven sampling	Online diary	Self-report on web interface	Retrospective	2	Location
Online respondent-driven sampling	Online diary	Self-report on web interface	Retrospective	2	Location
Quota sampling by age and sex	Paper diary and online diary	Self-report on paper and web interface	Both prospective and retrospective	2	Location, duration and frequency
Mixed samplings	Paper diary	Self-report	Prospective	1	Location

pertussis, varicella, and parvovirus-B19.[5,68,69] In addition, collecting information about location, duration, and frequency of contacts is also very essential for exploring mixing patterns and helping form effective strategies for disease prevention and control. In the case of school-aged children, a dominant number of contacts are made in school, leading to an indication that school closure can have a substantial impact on the spread of a respiratory infection.[3,20,23,70,71] Duration and frequency of contacts are important because they affect the probability of infecting another individual and if all contacts are treated equally, this may lead to wrongfully estimating the individual transmission probability.[29] Several studies found that close contacts with a duration of at least 15 minutes involving skin-to-skin touching were most predictive of the prevaccination prevalence of varicella zoster virus.[68,72] Therefore, the age of the contactee and duration of contact emerge as the most important information and should always be recorded in social contact surveys.

A comparison among all surveys based on a quantity such as the average number of contacts can be problematic. The study sample serves as the first obstacle. Given different research questions or participant availability, not all the samples studied can be considered as representative of the target population. This notion is important especially because age is a relevant determinant of social contacts, and samples in which a specific age class is over-represented regardless of study design can induce a strong bias to the number of contacts measured. For example, the study reporting the largest number of contacts (70.3[23]) was performed in a secondary school in the United Kingdom. Once these caveats are taken into account, the Table can be of value in identifying all of the surveys sharing similar features when addressing a specific research question. For example, the POLYMOD survey[6] demonstrated that the main structure of social interaction among age categories was the same among several European Union countries, although the strength of the interaction could vary between countries. On the other hand, the average number of contacts measured in sub-Saharan countries by Dodd et al[41] is considerably reduced compared with the average for high-income countries. In fact, the development level can be important in determining social interactions, e.g., due to local population density or reduced school attendance.[43,44,72] Quantifying the impact of different demographic factors on social contacts would require re-analysis of the datasets on the same basis and goes beyond the scope of this review. However, this could be performed in the future as datasets of social contact surveys will be made available from researchers in a unified format.[73]

This review used PubMed and Web of Knowledge for searching publications, possibly resulting in the omission of relevant publications. Nonetheless, the literature research step allowed us to recover more articles independently of a specific database, possibly recovering the ones we lost querying only two databases. Second, to the best of our knowledge, our search query failed to return the relevant articles of Leecaster et al[58] and Kwok et al,[74] which are eligible for this review. These articles are missing the words "survey", "questionnaire", and "diary" in the abstract and title, and therefore were missed by our searching method. The recent publication date (2016 and 2018) also prevented these articles from appearing in the references of the relevant articles. Another article that requires a specific clarification is the work of Watson et al[75] that asked participants to record whom they shared a meal with. This article was not included in our analysis, as considering a definition of social contact that is rather different from the body of this review.

Since the POLYMOD survey, there has been an increasing trend in the number of social contact surveys used to collect empirical contact data. Social contact surveys have been conducted widely in many countries, but most focused on high-income countries. These surveys used a range of different study designs with different study subjects, settings, sampling scheme to study designs, data collection tools, and data collection methods. Moreover, the definition of "contact" and its characteristics also differ, making comparison of contact patterns among surveys even more difficult. Improvements towards 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 of individuals.

This review demonstrates that contact surveys typically include of the order of a thousand participants, rely on convenience sampling, and use a retrospective design with paper diaries and self-reporting of contacts over a single day. Major determinants for this number of contactees include characteristics of the respondent (age, sex, and health status), time (weekday or weekend, and term time or vacation) and their immediate environment (household size and urban vs. rural). A typical number of different contactees reported per day is in the order of 20 for countrywide studies, a quantity that proved remarkably robust despite the many different study designs.

From the results of this review, we formulate the following recommendations for future surveys collecting social contact data relevant for the spread of respiratory pathogens.

(1) Study object: There is the need to continue the collection of social contact data, especially in low- and middle-income countries; still, in high-income countries, social contact surveys can detail interactions in epidemiologically relevant groups.
(2) Random sample: Depending on the study objectives, participants should be selected as randomly as possible, so findings can be properly extrapolated to encompass target population.
(3) Sampling: The sampling procedure should be described in detail, including response rates and information about at which stage and how participants can be excluded from the final sample, together with all the demographic factors considered when identifying the sample size.

(4) Method: Online and paper diaries both proved to be reliable to measure the overall contact matrix, but are associated with different burden for the participant, also depending on the age.
(5) Contact definition: At least two contact types should be included, one aimed to measure more casual contact and one aimed to measure physical contact.
(6) Prospective versus retrospective: Prospective design should be preferred to retrospective design, since it allows respondent to remember more contact features.
(7) Minimal contact information: Age and sex of the contacted person should definitely be included, as well as the duration, the frequency (intimate nature), and the location in which the contact took place.
(8) Scaling information: Information on the size of the possible pool of contactees (like household or school size) proves valuable for testing general assumptions on the scaling of human interactions with the size of the population.
(9) Behavioral change: Possible indicators of behavioral changes (feeling ill) should be included as well, given their large impact on disease spread.
(10) Sharing data: Finally, we want to bring to attention that several datasets referred to in this review are available in a unified format (www.socialcontactata.org[73]) that is compatible with an R package for social contact analysis (Socialmixr[76]). Complying with this standard format will improve the dissemination of future surveys' data and increase their value for the scientific community.

ACKNOWLEDGMENTS

This work received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 682540 TransMID and 283955 DECIDE).

REFERENCES

1. Anderson RM, May RM. *Infectious Diseases of Humans: Dynamics and Control*. Oxford: Oxford University Press; 1992.
2. Arregui S, Iglesias MJ, Samper S, et al. Data-driven model for the assessment of mycobacterium tuberculosis transmission in evolving demographic structures. Proceedings of the National Academy of Sciences. 2018. Available at: http://www.pnas.org/content/early/2018/03/20/1720606115.
3. De Luca G, Kerckhove KV, Coletti P, et al. The impact of regular school closure on seasonal influenza epidemics: a data-driven spatial transmission model for Belgium. *BMC Infectious Diseases*. 2018;18:29.
4. Edmunds WJ, O'callaghan C, Nokes D. Who mixes with whom? A method to determine the contact patterns of adults that may lead to the spread of airborne infections. *Proc Royal Soc Lond B*. 1997;264: 949–957.
5. Wallinga J, Teunis P, Kretzschmar M. Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am J Epidemiol.* 2006;164:936–944.
6. Mossong J, Hens N, Jit M, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* 2008;5:e74.
7. Read JM, Edmunds WJ, Riley S, Lessler J, Cummings DA. Close encounters of the infectious kind: methods to measure social mixing behaviour. *Epidemiol Infect.* 2012;140:2117–2130.
8. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med.* 2009;6:e1000097.
9. Clarivate Analytics (previously Thomson Reuters). A Smarter Way to Research. 2016. Available at: http://www.endnote.com. Accessed October 2016.
10. Beutels P, Shkedy Z, Aerts M, Van Damme P. Social mixing patterns for transmission models of close contact infections: exploring self-evaluation and diary-based data collection through a web-based interface. *Epidemiol Infect.* 2006;134:1158–1166.
11. Stein ML, van der Heijden PG, Buskens V, et al. Tracking social contact networks with online respondent-driven detection: who recruits whom? *BMC Infect Dis.* 2015;15:522.
12. Willem L, Van Kerckhove K, Chao DL, Hens N, Beutels P. A nice day for an infection? Weather conditions and social contact patterns relevant to influenza transmission. *PLoS One.* 2012;7:e48695.
13. Béraud G, Kazmercziak S, Beutels P, et al. The French connection: The first large population-based contact survey in France relevant for the spread of infectious diseases. *PLoS One.* 2015;10:e0133203.
14. Lapidus N, de Lamballerie X, Salez N, et al. 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.* 2013;8:e60127.
15. Bernard H, Fischer R, Mikolajczyk RT, Kretzschmar M, Wildner M. Nurses' contacts and potential for infectious disease transmission. *Emerg Infect Dis.* 2009;15:1438–1444.
16. Mikolajczyk RT, Akmatov MK, Rastin S, Kretzschmar M. Social contacts of school children and the transmission of respiratory-spread pathogens. *Epidemiol Infect.* 2008;136:813–822.
17. Mikolajczyk RT, Kretzschmar M. Collecting social contact data in the context of disease transmission: prospective and retrospective study designs. *Social Networks*. 2008;30:127–135.
18. Smieszek T, Castell S, Barrat A, Cattuto C, White PJ, Krause G. Contact diaries versus wearable proximity sensors in measuring contact patterns at a conference: method comparison and participants' attitudes. *BMC Infectious Diseases*. 2016;16:341.
19. Edmunds WJ, Kafatos G, Wallinga J, Mossong JR. Mixing patterns and the spread of close-contact infectious diseases. *Emerg Themes Epidemiol.* 2006;3:10.
20. Eames KT, Tilston NL, White PJ, Adams E, Edmunds WJ. The impact of illness and the impact of school closure on social contact patterns. *Health Technol Assess.* 2010;14:267–312.
21. Eames KT, Tilston NL, Edmunds WJ. The impact of school holidays on the social mixing patterns of school children. *Epidemics.* 2011;3: 103–108.
22. Eames KT, Tilston NL, Brooks-Pollock E, Edmunds WJ. Measured dynamic social contact patterns explain the spread of H1N1v influenza. *PLoS Comput Biol.* 2012;8:e1002425.
23. Jackson C, Mangtani P, Vynnycky E, et al. School closures and student contact patterns. *Emerg Infect Dis.* 2011;17:245–247.
24. Read JM, Eames KT, Edmunds WJ. Dynamic social networks and the implications for the spread of infectious disease. *J R Soc Interface.* 2008;5:1001–1007.
25. van Hoek AJ, Andrews N, Campbell H, Amirthalingam G, Edmunds WJ, Miller E. The social life of infants in the context of infectious disease transmission; social contacts and mixing patterns of the very young. *PLoS One.* 2013;8:e76180.
26. Danon L, House TA, Read JM, Keeling MJ. Social encounter networks: collective properties and disease transmission. *J R Soc Interface.* 2012;9:2826–2833.
27. Danon L, Read JM, House TA, Vernon MC, Keeling MJ. Social encounter networks: characterizing Great Britain. *Proc Biol Sci.* 2013;280:20131037.
28. Strömgren M, Holm E, Dahlström Ö, et al. Place-based social contact and mixing: a typology of generic meeting places of relevance for infectious disease transmission. *Epidemiol Infect.* 2017;145:2582–2593.
29. Smieszek T. A mechanistic model of infection: why duration and intensity of contacts should be included in models of disease spread. *Theor Biol Med Model.* 2009;6:25.

30. Smieszek T, Burri EU, Scherzinger R, Scholz RW. Collecting close-contact social mixing data with contact diaries: reporting errors and biases. *Epidemiol Infect.* 2012;140:744–752.
31. Stein ML, van Steenbergen JE, Buskens V, et al. Comparison of contact patterns relevant for transmission of respiratory pathogens in Thailand and The Netherlands using respondent-driven sampling. *PLoS One.* 2014;9:e113711.
32. van de Kassteele J, van Eijkeren J, Wallinga J, et al. Efficient estimation of age-specific social contact rates between men and women. *Ann Appl Stat*. 2017;11:320–339.
33. Ajelli M, Litvinova M. Estimating contact patterns relevant to the spread of infectious diseases in Russia. *J Theor Biol.* 2017;419:1–7.
34. Kwok KO, Cowling BJ, Wei VW, et al. Social contacts and the locations in which they occur as risk factors for influenza infection. *Proc Royal Soc Lond B*. 2014;281:20140709.
35. Read JM, Lessler J, Riley S, et al. Social mixing patterns in rural and urban areas of southern China. *Proc Royal Soc Lond B*. 2014;281:20140268.
36. Leung K, Jit M, Lau EHY, Wu JT. Social contact patterns relevant to the spread of respiratory infectious diseases in Hong Kong. *Sci Rep.* 2017;7:7974.
37. Ibuka Y, Ohkusa Y, Sugawara T, et al. Social contacts, vaccination decisions and influenza in Japan. *J Epidemiol Community Health*. 2015;70:162–167.
38. Fu YC, Wang DW, Chuang JH. Representative contact diaries for modeling the spread of infectious diseases in Taiwan. *PLoS One.* 2012;7:e45113.
39. Horby P, Pham QT, Hens N, et al. Social contact patterns in Vietnam and implications for the control of infectious diseases. *PLoS One.* 2011;6:e16965.
40. Kiti MC, Kinyanjui TM, Koech DC, Munywoki PK, Medley GF, Nokes DJ. Quantifying age-related rates of social contact using diaries in a rural coastal population of Kenya. *PLoS One.* 2014;9:e104786.
41. Dodd PJ, Looker C, Plumb ID, et al. Age- and sex-specific social contact patterns and incidence of mycobacterium tuberculosis infection. *Am J Epidemiol.* 2016;183:156–166.
42. Johnstone-Robertson SP, Mark D, Morrow C, et al. Social mixing patterns within a South African township community: implications for respiratory disease transmission and control. *Am J Epidemiol.* 2011;174:1246–1255.
43. Melegaro A, Del Fava E, Poletti P, et al. Social contact structures and time use patterns in the Manicaland province of Zimbabwe. *PLoS One.* 2017;12:e0170459.
44. Grijalva CG, Goeyvaerts N, Verastegui H, et al; RESPIRA PERU project. A household-based study of contact networks relevant for the spread of infectious diseases in the highlands of Peru. *PLoS One.* 2015;10:e0118457.
45. Aiello AE, Simanek AM, Eisenberg MC, et al. Design and methods of a social network isolation study for reducing respiratory infection transmission: the eX-FLU cluster randomized trial. *Epidemics.* 2016;15:38–55.
46. Glass LM, Glass RJ. Social contact networks for the spread of pandemic influenza in children and teenagers. *BMC Public Health.* 2008;8:61.
47. DeStefano F, Haber M, Currivan D, et al. Factors associated with social contacts in four communities during the 2007-2008 influenza season. *Epidemiol Infect.* 2011;139:1181–1190.
48. Potter GE, Handcock MS, Longini IM Jr, Halloran ME. Estimating within-school contact networks to understand influenza transmission. *Ann Appl Stat.* 2012;6:1–26.
49. Smieszek T, Barclay VC, Seeni I, et al. How should social mixing be measured: comparing web-based survey and sensor-based methods. *BMC Infect Dis.* 2014;14:136.
50. McCaw JM, Forbes K, Nathan PM, et al. Comparison of three methods for ascertainment of contact information relevant to respiratory pathogen transmission in encounter networks. *BMC Infect Dis.* 2010;10:166.
51. Rolls DA, Geard NL, Warr DJ, et al. Social encounter profiles of greater Melbourne residents, by location–a telephone survey. *BMC Infect Dis.* 2015;15:494.
52. Campbell PT, McVernon J, Shrestha N, Nathan PM, Geard N. Who's holding the baby? A prospective diary study of the contact patterns of mothers with an infant. *BMC Infect Dis.* 2017;17:634.
53. Chen SC, You SH, Ling MP, Chio CP, Liao CM. Use of seasonal influenza virus titer and respiratory symptom score to estimate effective human contact rates. *J Epidemiol.* 2012;22:353–363.
54. Chen SC, You ZS. Social contact patterns of school-age children in Taiwan: comparison of the term time and holiday periods. *Epidemiol Infect.* 2015;143:1139–1147.
55. Luh DL, You ZS, Chen SC. Comparison of the social contact patterns among school-age children in specific seasons, locations, and times. *Epidemics.* 2016;14:36–44.
56. Mastrandrea R, Fournet J, Barrat A. Contact patterns in a high school: a comparison between data collected using wearable sensors, contact diaries and friendship surveys. *PLoS One.* 2015;10:e0136497.
57. Etikan I, Musa SA, Alkassim RS. Comparison of convenience sampling and purposive sampling. *Am J Theor Appl Stat*. 2016;5:1–4.
58. Leecaster M, Toth DJ, Pettey WB, et al. Estimates of social contact in a middle school based on self-report and wireless sensor data. *PLoS One.* 2016;11:e0153690.
59. Akakzia O, Friedrichs V, Edmunds W, Mossony J. Comparison of paper diary vs computer assisted telephone interview for collecting social contact data relevant to the spread of airborne infectious diseases. In: *European Journal of Public Health*. Vol. 17. Oxford, United Kingdom: Oxford University Press; 2007:189–189.
60. Stein ML, van Steenbergen JE, Chanyasanha C, et al. Online respondent-driven sampling for studying contact patterns relevant for the spread of close-contact pathogens: a pilot study in Thailand. *PLoS One.* 2014;9:e85256.
61. Liao Q, Bai T, Zhou L, et al. Seroprevalence of antibodies to highly pathogenic avian influenza A (H5N1) virus among close contacts exposed to H5N1 cases, China, 2005-2008. *PLoS One.* 2013;8:e71765.
62. Creative Research System. Sample Size Calculator. 2016. Available at: https://www.surveysystem.com/sscalc.htm. Accessed June 2017.
63. Snijders TAB, Borgatti SP. Non-parametric standard errors and tests for network statistics. *Connections*. 1999;22:61–70.
64. Kolaczyk ED, Krivitsky PN. On the question of effective sample size in network modeling: an asymptotic inquiry. *Stat Sci*. 2015;30:184–198.
65. Handcock MS, Gile KJ. Modeling social networks from sampled data. *Ann Appl Stat.* 2010;4:5–25.
66. Krivitsky PN, Morris M. Inference for social network models from egocentrically sampled data, with application to understanding persistent racial disparities in HIV prevalence in the US. *Ann Appl Stat.* 2017;11: 427–455.
67. Fu YC. Contact diaries: building archives of actual and comprehensive personal networks. *Field methods*. 2007;19:194–217.
68. Goeyvaerts N, Hens N, Ogunjimi B, et al. Estimating infectious disease parameters from data on social contacts and serological status. *J R Stat Soc: Ser C (Applied Statistics)*. 2010;59:255–277.
69. Kretzschmar M, Teunis PF, Pebody RG. Incidence and reproduction numbers of pertussis: estimates from serological and social contact data in five European countries. *PLoS Med.* 2010;7:e1000291.
70. Hens N, Goeyvaerts N, Aerts M, Shkedy Z, Van Damme P, Beutels P. Mining social mixing patterns for infectious disease models based on a two-day population survey in Belgium. *BMC Infect Dis.* 2009;9:5.
71. Zhang T, Fu X, Kwoh CK, et al. Temporal factors in school closure policy for mitigating the spread of influenza. *J Public Health Policy.* 2011;32:180–197.
72. Melegaro A, Jit M, Gay N, Zagheni E, Edmunds WJ. What types of contacts are important for the spread of infections?: using contact survey data to explore European mixing patterns. *Epidemics.* 2011;3:143–151.
73. TransMID project and collaborators. data repository. 2017. Available at: www.socialcontactdata.org. Accessed July 2019.
74. Kwok KO, Cowling B, Wei V, Riley S, Read JM. Temporal variation of human encounters and the number of locations in which they occur: a longitudinal study of Hong Kong residents. *J Royal Soc Interface*. 2018;15:20170838.
75. Watson CH, Coriakula J, Ngoc DTT, et al. Social mixing in Fiji: who-eats-with-whom contact patterns and the implications of age and ethnic heterogeneity for disease dynamics in the Pacific Islands. *PLOS ONE*. 2017;12:1–16.
76. Sebastian Funk. Socialmixr R package. 2018. Available at: https://cran.r-project.org/web/packages/socialmixr/index.html. Accessed July 2019.