Using Time-Use Data to Parameterize Models for the Spread of Close-Contact Infectious Diseases

Emilio Zagheni, Francesco C. Billari, Piero Manfredi, Alessia Melegaro, Joel Mossong, and W. John Edmunds

Received for publication January 2, 2008; accepted for publication June 11, 2008.

Social contact patterns are a critical explanatory factor of the spread of close-contact infectious agents. Both indirect (via observed epidemiologic data) and direct (via diaries that record at-risk events) approaches to the measurement of contacts by age have been proposed in the literature. In this paper, the authors discuss the possibilities offered by time-use surveys to measure contact patterns and to explain observed seroprevalence profiles. The authors first develop a methodology to estimate time-of-exposure matrices, and then they apply it to time-use data for the United States (1987–2003). Finally, the authors estimate age-specific transmission parameters for varicella, commonly known as “chickenpox,” from age-specific time-of-exposure and seroprevalence data (United States, 1988–1994). The estimated time-of-exposure matrix reveals a strong element of assortativeness by age. In addition, there are peaks of exposure between people who were born one generation apart (i.e., parents and their children). Models based on the estimated age-specific transmission parameters fit the observed patterns of infection of endemically circulating varicella in a satisfactory way. The availability of time-use data for a large number of countries and their potential to supplement contact surveys make the methods developed extremely valuable and suitable for implementation in several different contexts.

chickenpox; communicable diseases; data interpretation, statistical; endemic diseases

Abbreviations: ATUS, American Time Use Survey; PTM, proportionate time mixing.

Several diseases, such as measles, mumps, rubella, and influenza, are transmitted by the respiratory or close-contact route. Social mixing patterns are therefore critical factors in the explanation of the transmission dynamics of a large number of infectious agents (1, 2). Knowledge of contact patterns is thus critical to devise containment strategies of a new, potentially devastating, infectious agent, such as a pandemic influenza, or to design effective control measures for endemic diseases. In particular, it is important to identify specific groups in a population that should be targeted by vaccination (3–5). In spite of the relevance of the subject, knowledge about the contact mechanisms underlying the diffusion of close-contact infectious diseases is still limited (6).

The contact structure in mathematical models of close-contact infections has usually been estimated indirectly. In populations stratified by specific individuals’ attributes, such as age, the transmission rates between age groups, which form the “who-acquires-infection-from-whom” matrix, are traditionally estimated by calibrating models to epidemiologic data, under suitable simplifying assumptions that allow reduction of the number of unknown parameters in order to make them estimable (7). Indirect approaches provide estimates of adequate contacts (8, 9), that is, the product of contacts and the corresponding risk of infection per single contact.

Recently, “direct” approaches, that is, aimed at directly filling in the elements of the contact matrix, have been increasingly used in order to overcome some limitations of “indirect” methods (10–13). The basic idea is to define an “at-risk” contact, for instance, a two-way conversation, and
to collect data from sample surveys. Usually, survey respondents are asked to complete diaries about the conversations they have throughout a randomly assigned day. Alternatively, contact matrices, together with average durations of contacts, have been estimated from secondary data sources, such as transportation surveys (14).

In this paper, we systematically explore the possibility of using a different source of data, that is, time-use data, to obtain information on contact patterns. First, we propose and discuss a general method to estimate “time-of-exposure” matrices from time-use sample surveys. Then, we apply the method to data for the United States, and we test the ability of the model based on time-use surveys to fit serologic data for varicella.

**MATERIALS AND METHODS**

**Data**

Our analysis is based on 2 different kinds of data sources, both for the United States: time-use data and seroprevalence data. Information on time use comes from the 2003 American Time Use Survey (ATUS), the 1989–1990 Activity Pattern Survey of California Children, and the 1987–1988 Activity Pattern Survey of Californians. In these surveys, respondents retrospectively record the chronologic sequence and duration of their daily activities in the form of diaries. For most activities, respondents are also asked to provide information on where the activity took place.

Seroprevalence data for varicella in the United States are obtained from the National Health and Nutrition Examination Surveys (1988–1994) and a module of the Behavioral Risk Factor Surveillance System (1991), which is a random-digit–dialed telephone survey sponsored by the Centers for Disease Control and Prevention. Details about the data sets analyzed and their consistency are provided as supplementary material to this paper. (This supplementary material is described in the text, two Web figures, and bibliography posted on the Journal’s website (http://aje.oxfordjournals.org/).

**Estimation of age-specific time of exposure**

In this section, we set up a methodology to estimate the daily amount of time that people spend together, on average, according to their age.

Several time-use surveys ask the respondent to record some information about the presence of other people during each of the activities the respondent did throughout the day. This is useful information, because it gives, for each activity, the average time spent by the respondents alone or in the presence of somebody else. In particular, the ATUS provides information about the members of the same household as the respondent who were present during the activity (e.g., their age and their relationship to the respondent). This means that we can estimate, directly from the survey, the time of exposure between members of the same households, by age. We represent the exposure by means of a matrix, $F$, whose entry $ij$ is the total time that people in the age group $i$ spend with their household members in the age group $j$ throughout an average day.

In addition to household members, people spend time with nonhousehold members during any activity considered, but the data available do not allow us to identify them. Consider, for instance, such activities as being on public transportation, at work, or in a restaurant: In these cases, people are likely to have contacts—in the sense of sharing the same room or having conversations—with a number of other people. In order to estimate the amount of time that people might spend with nonhousehold members, according to their age and during any such activity, we need to develop an indirect approach.

Our method is based on the assumption that we call “proportionate time mixing” (PTM) at the level of single activity/location and time slot. In other words, we assume that, for single activity/location and small time intervals, people allocate their time to the other participants in the activity proportionally to their relative participation in the activity. A matrix that represents the overall time of exposure to people by age, $E$, can be obtained by aggregating the matrices computed under the hypothesis of PTM at the level of single activity/location and time slot, $T$, and the time of exposure to household members’ matrix $F$.

We divide the whole day into 1,440 time slots, each of which consists of 1 minute. The number of people belonging to the age group $i$ that are in the location $h$ during the time slot $z$ is then equal to the number of minutes spent by the population in the age group $i$, in the location $h$, and during that particular time slot $z$ of the day considered. We refer to this quantity as $k_{hz}^i$, and we interpret it as a measure of person-minutes. If we let $n$ be the number of age groups in the population, then $\sum_{i=1}^n k_{hz}^i$ is the total amount of time spent by the population in the location $h$ during the time slot $z$. Conversely, $\sum_{z=1}^{1440} k_{hz}^i$ represents the total amount of time spent by people in the age group $i$ in the location $h$ throughout the entire day. When data come from a sample survey, the quantity $k_{hz}^i$ is obtained by multiplying the number of respondents in the age group $i$ that are in the location $h$ during the time slot $z$ by their respective sample weights. The assumption of PTM at the level of single activity/location and time slot implies that the time of exposure of people in the age group $i$ to people in the age group $j$, in the location $h$ and for the time slot $z$, $t_{ij}^{hz}$, can be computed as follows:

$$t_{ij}^{hz} = k_{j}^{hz} \times \frac{k_{hz}^i}{\sum_{i=1}^n k_{hz}^i}.$$

The idea is that, for each location $h$ and during the particular time slot $z$, people in the age group $i$ divide their time with people in the age group $j$ according to the relative participation of people in the age group $j$ in the activity that takes place in the location $h$ and during the time slot $z$. Consider, for example, the location “pub” and a population of 3 age groups: Assume that there are 100 people at pubs between 8:00 pm and 8:01 pm and that 20 of them belong to age group 1 (thus accounting for 20% of the population of “pub-goers” in the specific time interval), 50 belong to age group 2, and 30 belong to age group 3. By applying equation 1 to this location and time slot, we get a symmetric time-of-exposure matrix, such that people in age group 1 spend
Estimation of age-specific transmission parameters

The time-use approach relies on the assumption that the number of potentially infectious contacts between people of different age classes is proportional to their time of exposure. The proportionality factor, indicated by \( q \), is disease specific and measures the level of infectivity of the disease (11). Within the time-use framework, \( q \) represents the probability of transmission per time unit (e.g., minute) of exposure. Given an estimate of the average time of exposure of an individual in the age group \( j \) to people in the age group \( i \), and an estimate of the disease-specific infectivity parameter \( q \), we can write a time-use version of the "next generation matrix" (11, 15) as \( \mathbf{N} = (n_{ij} = (q a_{ij}) \). Within our framework, \( \mathbf{N} \) gives the age-specific amount of potential transmission time per person.

In order to estimate the infectivity parameter \( q \), we keep the elements \( a_{ij} \) fixed at their estimates based on time-use data, and we choose the value of \( q \) that maximizes the likelihood of observing the age-specific proportion of people immune to varicella in the United States prior to the introduction of the vaccine (for details on the methodology, refer to reference 11). Confidence intervals for \( q \), \( R_0 \), and the fit-to-prevalence data are derived using a bootstrap technique (16). There are 3 main sources of uncertainty associated with the estimates in the model. The first source is related to the variability of the estimates of the underlying time-of-exposure matrix. A second source is associated with the variability of the estimates of the transmission parameter, given a time-of-exposure matrix. Finally, there is uncertainty in the estimates of the seroprevalence profile by age. Our estimates of uncertainty based on a resampling technique account for the 3 sources of uncertainty.

RESULTS

Age-specific time of exposure

The estimated time of exposure reveals a strong element of assortativeness by age: People tend to spend more time with individuals of the same age. Figures 1 and 2 show contour plots of the average time-of-exposure matrices for the locations "school" and "workplace," respectively. These two matrices give a relevant contribution to the overall time-of-exposure matrix and account for most of the exposure between young people and adults, respectively. The other locations considered are "public transportation," "restaurant," "grocery store/mall," and "other public building." When the respondent is sleeping at night, no data are collected about his/her location.

Our results show a strong assortative pattern by age for children and teenagers while at school. People in the age group 5–9 years spend on average 98 minutes per day with people in the same age group at school, people in the age group 10–14 years spend on average 113 minutes per day with people in the same age group at school, and people in the age group 15–19 years spend on average 142 minutes per day with people in the same age group at school. Children divide their time at school with other children, teachers, childcare workers, and so on. We obtain estimates that represent the relative distribution of exposure to these people, and we observe a higher level of exposure to people of the same age. A slightly different pattern characterizes the location workplace (Figure 2), for which we observe that the time of exposure between adults still shows an assortative element by age but is less pronounced. In this case, the distribution of exposure is less concentrated around the diagonal elements of the time-of-exposure matrix.

Estimates of time of exposure between household members (i.e., the matrix \( \mathbf{F} \)) introduce a second important element to the average time-of-exposure matrix—the presence of the time-use matrices \( \mathbf{T} \) and \( \mathbf{F} \), we get the matrix \( \mathbf{E} \), that is, an estimate of the overall daily time of exposure between people of different age classes. By construction, at the population level, the total time of exposure of people in age group \( i \) to those in \( j \) (\( e_{ij} \)) must be equal to the total time of exposure of people in the age group \( j \) to those in \( i \) (\( e_{ji} \)). If there are \( w_i \) individuals in age group \( i \) and \( w_j \) individuals in age group \( j \), then \( e_{ij} = a_{ij} w_j = a_{ji} w_i \), where \( a_{ij} \) is the average daily time of exposure of an individual in the age group \( j \) to people in the age group \( i \). The quantity \( a_{ij} \) is a fundamental variable for our purposes and represents the time-use version of the average number of contacts by age group that has been proposed in the literature to measure mixing patterns (6, 10, 11, 13).

Estimation of age-specific transmission parameters

The total time of exposure of people in the age group \( i \) to people in the age group \( j \), for all locations and throughout an entire day, is obtained by aggregating over activities/locations and time slots as follows:

\[
t_{ij} = \frac{H}{k_i} \sum_{k=1}^{1440} \left( k_{ij}^k \times \frac{1}{n} \sum_{l=1}^{n} k_{lj}^k \right),
\]

where \( H \) is the number of activities/locations considered. The elements \( t_{ij} \) form the matrix \( \mathbf{T} \).

This method is based on the assumption of random allocation of time to people in different age groups within each time slot and activity/location: Assortativeness by age emerges from the fact that people of similar ages tend to do the same activities and tend to schedule those activities during the same time slots.

For the location "school," we have extra information related to the educational system of the country. We know, for instance, that students may attend classes in physically separated buildings according to the grade they are enrolled in and whether it is elementary, middle, or high school. We take this extra information into account by segregating the population into subpopulations of people who are "eligible to have contacts at school." In addition, the fact that respondents are supposed to record the physical location where they engage in activities (e.g., "school") and not only the activity they supposedly supposed to record the physical location where they engage in activities (e.g., "teaching") ensures that we take into account mixing between students and workers at school.

By summing up the matrices \( \mathbf{T} \) and \( \mathbf{F} \), we get the matrix \( \mathbf{E} \), that is, an estimate of the overall daily time of exposure between people of different age classes. By construction, at the population level, the total time of exposure of people in age group \( i \) to those in \( j \) (\( e_{ij} \)) must be equal to the total time of exposure of people in the age group \( j \) to those in \( i \) (\( e_{ji} \)). If there are \( w_i \) individuals in age group \( i \) and \( w_j \) individuals in age group \( j \), then \( e_{ij} = a_{ij} w_j = a_{ji} w_i \), where \( a_{ij} \) is the average daily time of exposure of an individual in the age group \( j \) to people in the age group \( i \). The quantity \( a_{ij} \) is a fundamental variable for our purposes and represents the time-use version of the average number of contacts by age group that has been proposed in the literature to measure mixing patterns (6, 10, 11, 13).

Estimation of age-specific transmission parameters

The time-use approach relies on the assumption that the number of potentially infectious contacts between people of different age classes is proportional to their time of exposure. The proportionality factor, indicated by \( q \), is disease specific and measures the level of infectivity of the disease (11). Within the time-use framework, \( q \) represents the probability of transmission per time unit (e.g., minute) of exposure. Given an estimate of the average time of exposure of an individual in the age group \( j \) to people in the age group \( i \),
Figure 1. Mean daily time of exposure (in minutes) between people in age group \( j \) (x-axis) and people in age group \( i \) (y-axis) while at school, \( a_{ij}^{\text{school}} \), United States, 1987–2003.

Figure 2. Mean daily time of exposure (in minutes) between people in age group \( j \) (x-axis) and people in age group \( i \) (y-axis) while at work, \( a_{ij}^{\text{workplace}} \), United States, 1987–2003.
of peaks on those elements that give the exposure between people who were born 1 generation apart (Figure 3). These peaks are related to the fact that people within a household tend to spend time with either those members who belong to the same generation (e.g., siblings, wife, husband) or those who are in a generation ahead or behind (e.g., parents and their children). As a result, the time-of-exposure matrix for household members, \( F \), shows peaks on both the main diagonal and on 1 of the subdiagonals, according to the time lapse between generations.

A graphical representation of the overall average time of exposure of an individual in the age group \( j \) to people in the age group \( i \), \( a_{ij} \), is given in Figure 4. We observe the highest level of assortativeness among people in the age range 10–19 years. According to our estimates, on average, they daily expose themselves to each other for a little less than 4 hours. We obtain similar numbers for the average time of exposure between children who are less than 10 years of age and people who are in the age range 25–39 years. We observe a high level of exposure, also, between people belonging to the same age group who are 65 or more years of age. Table 1 reports the estimated values (in minutes) of exposure between age groups.

Our time-of-exposure matrices are qualitatively similar to contact matrices obtained from contact surveys (11, 13). The two approaches reveal both assortativeness and peaks in exposure between people who are 1 generation apart (i.e., household contacts). Quantitatively, contact matrices tend to be more assortative than time-of-exposure matrices and, in that sense, they may be associated with larger values of \( R_0 \). On the other hand, they give little importance to household contacts, which are counted just as any other contact, whereas time-of-exposure matrices emphasize them.

### Fitting time-of-exposure matrices to the US serologic profile of varicella-zoster virus

The infectivity parameter \( q \) for varicella, which maximizes the likelihood of observing the proportions immune to the infection, given the average time-of-exposure matrix, is 0.00133 (95% confidence interval: 0.00106, 0.00160). The parameter \( q \) is disease specific and indicates how contagious the infection is. If we multiply the parameter \( q \) by the average time-of-exposure matrix given in Table 1, we obtain a time-use version of a next generation matrix for varicella in the United States. We estimate a value for \( R_0 \) equal to 7.32 (95% confidence interval: 5.31, 9.32). This value is consistent with other estimates of \( R_0 \) for varicella obtained from different models (17).
Figure 5 shows the fit of the model based on time-use data to age-specific immunity to varicella in the United States before the introduction of vaccination. The fit of the model to seroprevalence data is good overall. However, we observe a delayed rise in predicted seroprevalence from the model for the older-child age groups. This could be explained by the fact that the time-of-exposure matrix underestimates the level of assortativeness in contacts for older children and teenagers. For instance, when we compare the estimated time-of-exposure matrix with contact matrices estimated from contact surveys in European countries (13), we observe similar levels of assortativeness for the adult population. However, contact matrices show a stronger element of assortativeness for children and teenagers. An alternative explanation is that the choice of a parsimonious model based on only 1 age-independent transmission parameter \( q \) hides the fact that the transmissibility of the infection between age groups may be age dependent.

DISCUSSION

The measurement of human contact patterns has been pursued, in the literature, by applying both indirect (2) and direct (6, 13) methods. In both cases, the relevant epidemiologic variable is the number of contacts by age. Our paper considers the age-specific time of exposure as a relevant proxy for modeling close contacts by age and provides methods to estimate the time-of-exposure and mixing matrices from data collected through time-use sample surveys.

The assumption that time of exposure matters for transmission is crucial to our methodology. Because we do not know with accuracy what the routes of transmission are for varicella and other diseases, we are basically assuming that the time of exposure correlates with potentially infectious contacts or “at-risk” events. In other words, we assume that being in the same setting at the same time is an appropriate social proxy for the routes of transmission. In some cases, the assumption is certainly appropriate. For low-transmissibility infections, it is likely that the transmission of the pathogen is related to the repetition of close contacts that occur throughout an interval of exposure. This would explain the effectiveness of intrahousehold transmission, for instance. When the disease is highly transmissible, the accuracy of the time-use approach may be questioned. Nonetheless, work in progress for Italy, for which we have serologic data and both time-use and contact data, revealed that the model based on time-use data performed relatively well and very similarly to the model based on contact surveys.

In this paper, we first develop a methodology to estimate time-of-exposure matrices, and then we apply it to time-use
<table>
<thead>
<tr>
<th>Age Group</th>
<th>Mean Daily Exposure (minutes)</th>
<th>95% Confidence Interval</th>
<th>Age Group</th>
<th>Mean Daily Exposure (minutes)</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–4 years</td>
<td>154</td>
<td>93, 215</td>
<td>5–9 years</td>
<td>112</td>
<td>70, 154</td>
</tr>
<tr>
<td></td>
<td>10–14 years</td>
<td>103, 63, 144</td>
<td></td>
<td>102</td>
<td>31, 173</td>
</tr>
<tr>
<td></td>
<td>15–19 years</td>
<td>50, 0, 103</td>
<td></td>
<td>74</td>
<td>18, 130</td>
</tr>
<tr>
<td></td>
<td>20–24 years</td>
<td>135, 69, 202</td>
<td></td>
<td>47</td>
<td>0, 93</td>
</tr>
<tr>
<td></td>
<td>25–29 years</td>
<td>242, 175, 308</td>
<td></td>
<td>127</td>
<td>69, 185</td>
</tr>
<tr>
<td></td>
<td>30–34 years</td>
<td>254, 183, 325</td>
<td></td>
<td>175</td>
<td>121, 229</td>
</tr>
<tr>
<td></td>
<td>35–39 years</td>
<td>170, 109, 231</td>
<td></td>
<td>197</td>
<td>139, 254</td>
</tr>
<tr>
<td></td>
<td>40–44 years</td>
<td>76, 22, 131</td>
<td></td>
<td>137</td>
<td>85, 190</td>
</tr>
<tr>
<td></td>
<td>45–49 years</td>
<td>38, 0, 89</td>
<td></td>
<td>63</td>
<td>12, 114</td>
</tr>
<tr>
<td></td>
<td>50–54 years</td>
<td>20, 0, 77</td>
<td></td>
<td>33</td>
<td>0, 83</td>
</tr>
<tr>
<td></td>
<td>55–59 years</td>
<td>12, 0, 64</td>
<td></td>
<td>23</td>
<td>0, 68</td>
</tr>
<tr>
<td></td>
<td>60–64 years</td>
<td>15, 0, 82</td>
<td></td>
<td>26</td>
<td>0, 85</td>
</tr>
<tr>
<td></td>
<td>65–69 years</td>
<td>12, 0, 80</td>
<td></td>
<td>13</td>
<td>0, 70</td>
</tr>
<tr>
<td></td>
<td>70–74 years</td>
<td>18, 0, 89</td>
<td></td>
<td>17</td>
<td>0, 79</td>
</tr>
<tr>
<td></td>
<td>75–79 years</td>
<td>3, 0, 72</td>
<td></td>
<td>4</td>
<td>0, 61</td>
</tr>
<tr>
<td>40–44 years</td>
<td>77, 17, 138</td>
<td>38, 0, 88</td>
<td>5–9 years</td>
<td>139</td>
<td>82, 197</td>
</tr>
<tr>
<td></td>
<td>10–14 years</td>
<td>157, 103, 211</td>
<td></td>
<td>102</td>
<td>54, 150</td>
</tr>
<tr>
<td></td>
<td>15–19 years</td>
<td>78, 61, 96</td>
<td></td>
<td>80</td>
<td>60, 100</td>
</tr>
<tr>
<td></td>
<td>20–24 years</td>
<td>60, 45, 75</td>
<td></td>
<td>71</td>
<td>51, 90</td>
</tr>
<tr>
<td></td>
<td>25–29 years</td>
<td>55, 39, 70</td>
<td></td>
<td>50</td>
<td>40, 60</td>
</tr>
<tr>
<td></td>
<td>30–34 years</td>
<td>66, 54, 77</td>
<td></td>
<td>50</td>
<td>40, 59</td>
</tr>
<tr>
<td></td>
<td>35–39 years</td>
<td>90, 75, 105</td>
<td></td>
<td>58</td>
<td>48, 69</td>
</tr>
<tr>
<td></td>
<td>40–44 years</td>
<td>146, 126, 167</td>
<td></td>
<td>94</td>
<td>76, 111</td>
</tr>
<tr>
<td></td>
<td>45–49 years</td>
<td>92, 77, 106</td>
<td></td>
<td>153</td>
<td>129, 177</td>
</tr>
<tr>
<td></td>
<td>50–54 years</td>
<td>64, 50, 79</td>
<td></td>
<td>94</td>
<td>76, 112</td>
</tr>
<tr>
<td></td>
<td>55–59 years</td>
<td>52, 42, 62</td>
<td></td>
<td>73</td>
<td>55, 91</td>
</tr>
<tr>
<td></td>
<td>60–64 years</td>
<td>60, 30, 50</td>
<td></td>
<td>42</td>
<td>31, 53</td>
</tr>
<tr>
<td></td>
<td>65–69 years</td>
<td>33, 23, 42</td>
<td></td>
<td>32</td>
<td>20, 43</td>
</tr>
<tr>
<td></td>
<td>70–74 years</td>
<td>30, 15, 45</td>
<td></td>
<td>44</td>
<td>23, 65</td>
</tr>
<tr>
<td></td>
<td>75–79 years</td>
<td>22, 10, 34</td>
<td></td>
<td>26</td>
<td>19, 33</td>
</tr>
</tbody>
</table>

† The 95% confidence intervals were obtained using a bootstrap technique.
data for the United States. We document a strong heterogeneity in exposure patterns by age, in addition to a relevant element of assortativeness by age and considerable exposure between people who were born 1 generation apart (mainly explained by intrahousehold exposure). From a theoretical viewpoint, our approach has analogies with the structured mixing scheme (18). We observe that age-specific exposure is not random but that it is structured from the set of socioeconomic and institutional factors that determine the allocation of an individual’s time to life’s activities.

In the second part of this paper, we fit a model based on age-specific time of exposure to US seroprevalence data for varicella. The fit is quite good, yet we chose a highly parsimonious model with only 1 transmission parameter (11). More complex models that account for the fact that some individuals may be more infectious than others (e.g., age-dependent transmission rates) or that different activities may lead to different chances of infection (e.g., activity-specific transmission rates) could be used. However, it is noteworthy that a simple approach, based on 1 transmission parameter, given a time-of-exposure matrix, performs so well on real data. This is encouraging because there is a huge amount of information on time use that is immediately available for a large group of countries. In addition, time-use data sets are in most cases free or inexpensive and have been harmonized to allow for comparative studies (e.g., the Multinational Time-Use Study).

Our method can be applied also in those situations for which data on time use are not as detailed as they are for the United States. For several time-use surveys, for instance, data on the exact time of beginning and ending of activities are not available. We know only the total time spent by the respondent in each location throughout a day. In those cases, we can still compute the overall time of exposure between people of different age groups but at a lower level of disaggregation. We expect the time-of-exposure matrix obtained from these data to be less assortative because we cannot account for the fact that people of the same age groups tend to schedule their activities during the same time slots, but we can capture only the fact that people of the same age groups tend to do the same activities throughout an average day. We can use data from ATUS to compute a mixing matrix based on aggregated values by location. The deviation of this matrix from the homogeneous mixing one accounts for 98% of the deviation, in terms of the Frobenius norm, of the respective mixing matrix computed under PTM at the level of both single activity/location and time slot from the homogeneous mixing matrix. One important consequence of this result is that reliable matrices, from a statistical point of view, can be computed from data sets that are less detailed than ATUS. A second relevant fact to notice is that enlarging the time slot from the shortest (1 minute) to the longest (1 day) does not change the mixing pattern substantially. This evidence supports the reliability of the estimates based on disaggregation to single-minute time slots, despite the fact that, for some locations, during some time slots, these estimates could be highly stochastic. A third observation is that the time-of-exposure matrix that we estimate based on the PTM at the level of both single activity/location and time slot may still underestimate the level of assortativeness in the population. As a matter of fact, it ultimately depends on the assumption of PTM, and it may not account for several kinds of assortative mixing. For example, if one pub appealed to teenagers, whereas another one appealed to older people, there would be assortative mixing that is not reflected in the time-of-exposure matrix obtained with our method.

Time-use data are particularly interesting because they potentially supplement contact surveys. As a matter of fact, a method based on time-use data allows for measuring those contacts that standard definitions of “at-risk” events, such as a 2-way conversation, do not capture (e.g., being on public transportation, in a mall, and so on). However, the method that we propose does not account for differences in densities in various locations. By scaling the exposure between age groups in a location by the overall amount of person-minutes spent in the location, our method implements frequency-dependent mixing (19), consistent with a behavioral model in which individuals have a constant rate of contacting others, regardless of their density.

Our methodology does not take into account the different nature of contacts in different settings. For instance, contacts in a household may be more intimate and involve physical proximity, whereas contacts in a public building may be more distant. This limitation could be overcome by looking at time-of-exposure matrices disaggregated by activity/location. We foresee further research to be pursued to investigate time-of-exposure patterns in specific settings. This may improve our understanding of where transmission

**Figure 5.** Age-specific immunity to varicella, commonly known as chickenpox, before the introduction of vaccination, Third National Health and Nutrition Examination Survey (1988–1994) and Behavioral Risk Factor Surveillance System (1991). Markers indicate observations; dashed lines indicate fit according to the approach based on time-use data under the hypothesis of proportionate time mixing at the level of single activity/location and time slot; and vertical lines give the 95% bootstrap confidence intervals for the fit.
occurs and what locations should be targeted by specific interventions (e.g., school closure). We also think that more research is needed to supplement time-use data with contact surveys and epidemiologic data that can inform on transmission characteristics in different settings and for different types of individuals.

ACKNOWLEDGMENTS

Author affiliations: Department of Demography, University of California, Berkeley, California (Emilio Zagheni); Università Bocconi, Milano, Italia (Francesco C. Billari); Università di Pisa, Pisa, Italia (Piero Manfredi); Health Protection Agency, London, United Kingdom (Alessia Melegaro); Laboratoire National de Santé, Luxembourg, Luxembourg (Joel Mossong); and Health Protection Agency, London, United Kingdom (W. John Edmunds).

This study is part of POLYMOD, a European Commission project funded within the Sixth Framework Programme (contract SSP22-CT-2004502084).

This work has benefited from comments by those associated with the POLYMOD project and those attending the brown-bag seminar at the Department of Demography, University of California, Berkeley.

Conflict of interest: none declared.

REFERENCES


