







Behavioral factors and SARS-CoV-2 transmission heterogeneity within a household cohort in Costa Rica

Kaiyuan Sun¹, Viviana Loria², Amada Aparicio³, Carolina Porras², Juan Carlos Vanegas², Michael Zúñiga², Melvin Morera³, Carlos Avila², Arturo Abdelnour³, Mitchell H. Gail⁴, Ruth Pfeiffer⁴, Jeffrey I. Cohen⁵, Peter D. Burbelo⁶, Mehdi A. Abed⁶, Cécile Viboud¹, Allan Hildesheim⁴, Rolando Herrero², D. Rebecca Prevots⁷ & for the RESPIRA Study Group*

Abstract

Introduction Variability in household secondary attack rates and transmission risks factors of SARS-CoV-2 remain poorly understood.

Methods We conducted a household transmission study of SARS-CoV-2 in Costa Rica, with SARS-CoV-2 index cases selected from a larger prospective cohort study and their household contacts were enrolled. A total of 719 household contacts of 304 household index cases were enrolled from November 21, 2020, through July 31, 2021. Blood specimens were collected from contacts within 30–60 days of index case diagnosis; and serum was tested for presence of spike and nucleocapsid SARS-CoV-2 IgG antibodies. Evidence of SARS-CoV-2 prior infections among household contacts was defined based on the presence of both spike and nucleocapsid antibodies. We fitted a chain binomial model to the serologic data, to account for exogenous community infection risk and potential multi-generational transmissions within the household.

Results Overall seroprevalence was 53% (95% confidence interval (CI) 48–58%) among household contacts. The estimated household secondary attack rate is 34% (95% CI 5–75%). Mask wearing by the index case is associated with the household transmission risk reduction by 67% (adjusted odds ratio = 0.33 with 95% CI: 0.09–0.75) and not sharing bedroom with the index case is associated with the risk reduction of household transmission by 78% (adjusted odds ratio = 0.22 with 95% CI 0.10–0.41). The estimated distribution of household secondary attack rates is highly heterogeneous across index cases, with 30% of index cases being the source for 80% of secondary cases.

Conclusions Modeling analysis suggests that behavioral factors are important drivers of the observed SARS-CoV-2 transmission heterogeneity within the household.

Plain Language Summary

When living in the same house with known SARS-CoV-2 cases, household members may change their behavior and adopt preventive measures to reduce the spread of SARS-CoV-2. To understand how behavioral factors affect SARS-CoV-2 spreading in household settings, we focused on household members of individuals with laboratory-confirmed SARS-CoV-2 infections and followed the way SARS-CoV-2 spread within the household, by looking at who had antibodies against the virus, which means they were infected. We also asked participants detailed questions about their behavior and applied mathematical modeling to evaluate its impact on SARS-CoV-2 transmission. We found that mask-wearing by the SARS-CoV-2 cases, and avoiding sharing a bedroom with the infected individuals, reduces SARS-CoV-2 transmission. However, caring for SARS-CoV-2 cases, and prolonged interaction with infected individuals facilitate SARS-CoV-2 spreading. Our study helps inform what behaviors can help reduce SARS-CoV-2 transmission within a household.

¹ Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health (NIH), Bethesda, MD, USA.

² Agencia Costarricense de Investigaciones Biomédicas (ACIB) – Fundación INCIENSA (FUNIN), San José, Costa Rica. ³ Caja Costarricense de Seguro Social, San José, Costa Rica. ⁴ Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Bethesda, MD, USA. ⁵ Laboratory of Infectious Diseases, National Institute of Allergy and Infectious Diseases (NIAID), NIH, Bethesda, MD, USA. ⁶ National Institute of Dental and Craniofacial Research, NIH, Bethesda, MD, USA. ⁷ Epidemiology and Population Studies Unit, Laboratory of Clinical Immunology and Microbiology, Division of Intramural Research, NIAID, NIH, Bethesda, MD, USA. *A list of authors and their affiliations appears at the end of the paper. ✉email: kaiyuan.sun@nih.gov; rprevots@niaid.nih.gov

The household has been recognized as one of the main settings for SARS-CoV-2 transmission¹ with high secondary attack rates reported among household contacts² across multiple countries and in different phases of the pandemic^{3–6}. Even after the initial acute phase of the pandemic, public health agencies in many countries recommended home-based isolation for people with confirmed SARS-CoV-2 infections to reduce overall community transmission⁷. However, for vulnerable individuals, having a household contact with confirmed SARS-CoV-2 infection greatly increases the risk of infection, which could lead to hospitalization or even death. While vaccination became available in 2021 in many countries with high effectiveness against symptomatic infections, the emergence of high-transmissibility and immune-escape variants, such as Omicron, along with waning immunity, have rekindled the importance of non-pharmaceutical interventions. Public health agencies have provided guidelines to reduce transmission within a household setting, including mask wearing and living in separate bedrooms⁷, the effectiveness of such guidelines remain largely untested with real-world data.

Several household transmission studies have been conducted in high income countries, however, data from low- and middle-income countries are limited. Costa Rica has a universal health care system with a good infrastructure and robust surveillance system, which is ideal for conducting population-based transmission studies. Health care is centralized under the Costa Rican Social Security (Caja Costarricense de Seguro Social- CCSS) and most patients with COVID-19 are treated and followed at one of its health facilities with detailed records kept. The first case of COVID-19 in Costa Rica was detected on March 6, 2020, and soon after the CCSS Ministry of Health implemented population-level intervention measures including school closings and isolation at home for positive cases⁸.

To better estimate the secondary attack rates and understand the behavioral determinants of SARS-CoV-2 household transmission, we conduct a household serologic study nested within a larger prospective population-based study of the SARS-CoV-2 immunologic response in Costa Rica. We fit the serologic data to a chain-binomial household transmission model to account for the non-linear transmission dynamics as well as the time-varying community infection risk. Moreover, this model is able to incorporate detailed demographic, clinical and behavior risk factors of the index and household contacts. We estimate the overall household secondary attack rate, the cumulative community infection risk, and assess sources of transmission heterogeneities among household members. We find that mask wearing by the index case and avoid sharing bedroom with the index case reduce SARS-CoV-2 transmission risk, while caring for index case and prolonged interaction with the index case increase SARS-CoV-2 transmission risk. Our study demonstrates behavioral factors and preventive measures are significant drivers of SARS-CoV-2 transmission in the household setting.

Methods

Study population. For the larger prospective study, 1000 cases were recruited from three geographic areas: Puntarenas Province, Greater San Jose Metropolitan Area—(Gran Area Metropolitana), and the province of Guanacaste, and four age strata (0–19, 20–39, 40–59, 60+) using national surveillance lists provided by the CCSS and Health Ministry. Only PCR confirmed cases were included in the list from the surveillance system, because they met the case definition used for surveillance. The geographic areas were selected based on logistic considerations and represented 58% percent of the Costa Rican population. Cases were sampled randomly within each geographic area and age stratum.

Approximately 30% of cases were approached for consent to participate in the nested household study; these cases were termed “index” cases.

A household was defined as two or more people living together who shared a kitchen. To be eligible for inclusion, a contact must have spent at least one night per week in the living area since the diagnosis of the index case. After consent and enrollment, index cases and their household contacts were administered a questionnaire to ascertain demographic, clinical, and behavioral risk and preventive factors. For household contacts, symptoms related to SARS-CoV-2 were ascertained for the time period two weeks before or two weeks after the sample collection date for the index case (referred to hereafter as “date of diagnosis”). If a household contact reported a prior diagnosis of COVID, symptoms were ascertained in relation to that diagnosis. Blood samples were collected from household contacts 30 to 60 days after the date of collection of the PCR-confirmed positive sample of the index case, and serum samples were tested to ascertain the presence of SARS-CoV-2 antibodies (against both SARS-CoV-2 nucleocapsid and spike protein), as a marker of past SARS-CoV-2 infection.

Household index cases and their contacts were enrolled from December 1, 2020, through July 31, 2021. This period coincided with the middle of the first wave and the end of the second wave in Costa Rica (Fig. S1). The study was conducted immediately prior to the widespread availability of SARS-CoV-2 vaccines in Costa Rica⁹. Once study recruitment had been completed, national vaccination registries were searched to ascertain vaccination status and dates for any participants who had been vaccinated.

The RESPIRA study protocol was approved by the Central Institutional Review Board of the CCSS. (Protocol R020-SABI-000261). Informed, signed consent was obtained from all study participants or their parents (for participants younger than 18).

Serologic methods. Serum samples were tested for the presence of SARS-CoV-2 spike and nucleocapsid anti-IgG antibodies using a previously validated quantitative immunoprecipitation assay in a microtiter plate format. Technical details of the immunoprecipitation assay can be found in paper by Burbelo et al.¹⁰. At >14 days after SARS-CoV-2 symptom onset, the assay detected antibodies against SARS-CoV-2 nucleocapsid protein with 100% sensitivity and 100% specificity by the assay, whereas antibodies to spike protein were detected with 91% sensitivity and 100% specificity by the assay¹⁰. We defined seropositivity as positive to both spike and nucleocapsid antigens and considered it evidence of past SARS-CoV-2 infection. The serum samples were collected between 30 and 60 days after the index case PCR positive sample collection to allow time for seroconversion. Approximately 7.5% of the samples were incorporated into the plates in a blinded fashion to evaluate within and between plate variability. The one-way Intraclass correlation coefficient (ICC) for nucleocapsid within-plate duplicate was 0.94 with 95% CI 0.87–0.97; the ICC for spike within-plate duplicate was 0.95 with 95% CI 0.89–0.98; the ICC for nucleocapsid across-plate duplicate was 0.71 with 95% CI 0.44–0.87; the ICC for spike within-plate duplicate was 0.87 with 95% CI 0.72 – 0.94. In addition 25 pre-pandemic samples from a population study in Costa Rica¹¹ were tested as negative controls to ensure assay validity; all were classified as seronegative, as expected.

Chain binomial household transmission model. Here we consider a multi-variable chain-binomial household transmission model for SARS-CoV-2, as an extension of prior household models developed to study influenza transmission^{12,13}. The model was

fitted to the cumulative outbreak size at the end of the household outbreak, i.e., the total number of people infected, rather than the precise sequence and timeline of infections. We do not assume that all seropositive household members acquired infections from the index case and allow for community-acquired infections (prior to blood sample collection) for household members and multi-generational transmission within the household.

Specifically, let h denote a household, i an individual, with i_h^- an individual i who is serologic negative in household h and i_h^+ an individual i who is serologic positive in household h . The risk of acquiring infection from the community varies over time due to changing incidence and is written as $P_c * f(t)$, where $f(t)$ is the cumulative incidence rate from the start of the pandemic until time t in Costa Rica, and P_c is the baseline community infection risk to be estimated by the model. If we denote t_s^i the time of serology sample collection for household member i , then the likelihood of an individual i escaping infection from the community is given by:

$$l_c^i = 1 - P_c * f(t_s^i) \tag{1}$$

To model the risk of transmission between the index case and household members, we denote P_h^{ic-j} as the risk of index case ic infecting household member j in household h . We can express P_h^{ic-j} as:

$$P_h^{ic-j} = P_{index} \exp\left(\sum_k \alpha_k a_k\right) \tag{2}$$

where P_{index} denotes the baseline risk of SARS-CoV-2 transmission within the household and a_k represents risk factor k that could potentially influence household transmission risk. We include risk factors affecting transmissibility and susceptibility such as age, sex, and obesity status (self-reported) of the index case ic and of household contacts j . We also include household size, and behavioral factors such as whether the contact shared a bedroom with the index case, spent time with the index case outside the bedroom, cared for the index case, and wore a mask following diagnosis in the index case (index case and household member). We examine a sequence of models including various risk factors as shown in Supplemental Table S1. We can then express the likelihood of household member i escaping infection from the index case as:

$$l_{index}^i = \left(1 - P_h^{ic-j}\right) \tag{3}$$

To model the risk of transmission between household contacts (in i.e., chains of transmission that do not involve the index case), we denote P_h^{ij} as the risk of seropositive household contact i infecting household contact j in household h . We can express P_h^{ij} as:

$$P_h^{ij} = P_{hh} \exp\left(\sum_k \beta_k b_k\right) \tag{4}$$

where P_{hh} denotes the baseline risk of SARS-CoV-2 transmission between an infected household contact and another uninfected household contact and b_k represents risk factor k that could potentially influence household transmission risk. We consider a different baseline transmission risk from an average household member (P_{hh}) vs the index case (P_{index}) because we have slightly different behavioral variables collected from household members and index cases, especially regarding all the pairwise interactions between index case and household members. This does not assume an inherently different biologic risk from index cases and household members, but merely measurement differences. Here we include risk factors including age, sex, and obesity status of the seropositive infector i (modulating transmissibility) as well as household contacts j (modulating susceptibility). We also include

Table 1 Characteristics of the studied population stratified by index case and household members.

Characteristics	Index cases	Household contacts
Age		
0-12 years	30/304 (10%)	148/719 (21%)
13-24 years	53/304 (17%)	125/719 (17%)
25-39 years	75/304 (25%)	163/719 (23%)
40-59 years	80/304 (26%)	171/719 (24%)
≥ 60 years	66/304 (22%)	112/719 (16%)
Sex		
Male	141/304 (46%)	315/719 (44%)
Female	163/304 (54%)	404/719 (56%)
Obesity		
Yes	36/304 (12%)	41/719 (6%)
No	268/304 (88%)	678/719 (94%)
Care for index case		
Yes	—	274/719 (38%)
No	—	359/719 (50%)
NA ^a	—	86/719 (12%)
Sharing bedroom with index case (2 weeks post index diagnosis)		
Yes	—	191/719 (27%)
No	—	525/719 (73%)
NA ^a	—	3/719 (3%)
Hours (per day) with index case outside bedroom (2 weeks post index diagnosis)		
<1 h	—	236/719 (33%)
>1 h	—	473/719 (66%)
NA ^a	—	10/719 (1%)
Index case mask wearing frequency ^b (2 weeks post index diagnosis)		
>half of the time	—	189/719 (26%)
≤half of the time	—	322/719 (45%)
NA ^a	—	208/719 (29%)
Household contact mask wearing frequency ^b (2 weeks post index diagnosis)		
> half of the time	—	184/719 (25%)
≤ half of the time	—	349/719 (49%)
NA ^a	—	186/719 (26%)
Vaccination status		
Unvaccinated	296/304 (97%)	649/719 (90%)
1 st dose ^c	5/304 (2%)	29/719 (4%)
2 nd dose ^c	3/304 (1%)	41/719 (6%)

All information except vaccination status was self-reported based on questionnaires.
^aNA denotes missing data.
^bHere we looked at the mask wearing frequency of index case and household contact when they were interacting with each other under the household setting.
^cSARS-CoV-2 vaccine administered in Costa Rica was Comirnaty (BNT162b2).

household size, whether the recipient wears a mask, and the symptom status of the infector.

For a given risk factor a_k or b_k , missing information (Table 1) is treated as an independent category within the risk factor and included in the transmission model.

We can then express the probability of household contact i escaping infection from all positive household contacts as:

$$l_{hh}^i = \prod_{\{j \neq i\}} \left(1 - P_h^{ij}\right) \tag{5}$$

where $\{j \neq i\}$ represents all positive household contacts. Thus, within household h , the likelihood of household contact i being seronegative at the end of the household outbreak is given by:

$$l_-^i = l_c^i l_{index}^i l_{hh}^i \tag{6}$$

And the likelihood of an individual i being seropositive is given by:

$$l_+^i = 1 - l_-^i \tag{7}$$

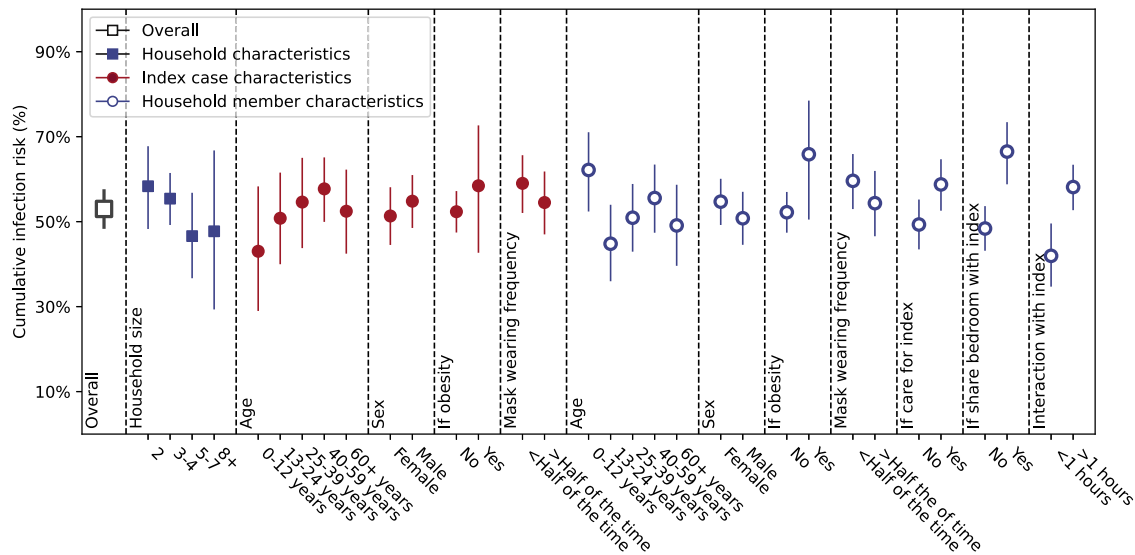


Fig. 1 The overall cumulative infection risk among household members and cumulative infection risk by different risk factors (unadjusted). The overall cumulative infection risk is calculated as fraction of seropositive among all 719 household contacts. The cumulative infection risk within a given stratum is calculated as the fraction of seropositive individuals among the household contacts within the stratum. We stratified the 719 household contacts by household level characteristics of household size; index case characteristics including index cases’ age, sex, obesity or not, mask wearing frequency; household member properties including household contacts’ age, sex, obesity or not, mask wearing frequency, if cared for index case, shared bedroom with index case or interaction frequency with index case after the diagnosis of the index case. Confidence intervals are based on a generalized estimating equation analysis applied to each risk factor one at a time that takes within household correlations into account.

For household h , the loglikelihood of observing the infection status of all household contacts is given by:

$$\log(l^h) = \sum_{\{i^h_+\}} \log(l^h_{i^h_+}) + \sum_{\{i^h_-\}} \log(l^h_{i^h_-}) \quad (8)$$

The overall likelihood of the observations across all households is given by:

$$\log(L) = \sum_h \log(l^h) \quad (9)$$

We fit the model to serology observations and used maximum likelihood method to infer parameters: estimates on P_c, P_{hh}, P_{index} reported in Fig. 2b while $\{\alpha_k\}$ and $\{\beta_k\}$ estimates were reported in Fig. 2a. 95% confidence intervals were determined by likelihood ratio test. To address potential household clustering effect, we bootstrapped over 304 households, controlling for the household size distribution as well as the distribution of age category, sex and diagnostic month of the index cases (i.e., for each bootstrap samples, the joint distribution of household size, index case’s age category (see Table 1 for the age strata), index case’s sex, and index case’s diagnostic month is the same as the original data. We used 100 times repeated bootstrapped estimates to construct bootstrapping confidence intervals as a sensitivity analysis.

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

Results

Study overview. From December 1, 2020, through July 30, 2021, a total of 986 household contacts were approached of whom 719 (73%) consented to enroll in the household study. These contacts were distributed in 304 households. This study period covered the first wave and the beginning of the second wave in Costa Rica (S1). The total household size ranged from 2 to 9 with an average household size of 3.3. Among 304 index cases, the median age was 38 (range: 0–101); 163 (54%) index cases were female. Among 719

household contacts, the median age was 32 (range: 0–93); 404 (56%) household contacts were female. Less than 10% of the study participants were vaccinated with at least one vaccine dose. Detailed demographic, clinical, and behavioral factors for both the index case and the household contacts are summarized in Table 1.

Seroprevalence among household contacts. To evaluate the burden of SARS-CoV-2 among 719 household contacts, we estimated the seroprevalence both overall and within strata, defined by: household size; age, sex, and obesity status of the index case and household contacts; and behavioral factors such as mask usage and interactions of the contact with the index case (Fig. 1). The seroprevalence along with 95% CI were estimated using univariate generalized estimating equations with household clustering. The overall seroprevalence was 53% (95% CI 48–58%) among household contacts, but seroprevalence varied substantially across different strata. In particular, for behavioral risk factors, the seroprevalence was significantly higher if the contact shared a bedroom with the index case (67% with 95% CI 59–73%) vs 48% (95% CI 43–54%, $p < 0.001$) or had interactions with the index case outside the bedroom (58% with 95% CI 53–63% for >1 h vs 41% with 95% CI 34–50% for <1 h, trend test $p < 0.001$).

Age-specific mixing patterns between the index case and household contacts are shown in Fig. S2a. The household mixing patterns in Costa Rica resemble those observed in other countries^{14,15}, with a distinct “three-band” feature. The diagonal band represents mixing with contacts of approximately the same age, while the two off-diagonal bands representing inter-generational mixing (parents living with young kids/adults living with elder parents). The mixing pattern between index and seropositive household contacts is distinctly different from that between index and seronegative contacts (Fig. S2b, c), suggesting age may be a significant risk factor associated with SARS-CoV-2 transmission. We thus further explored age as a risk factor by including age variables modulating infectivity and susceptibility in the chain-binomial household transmission models.

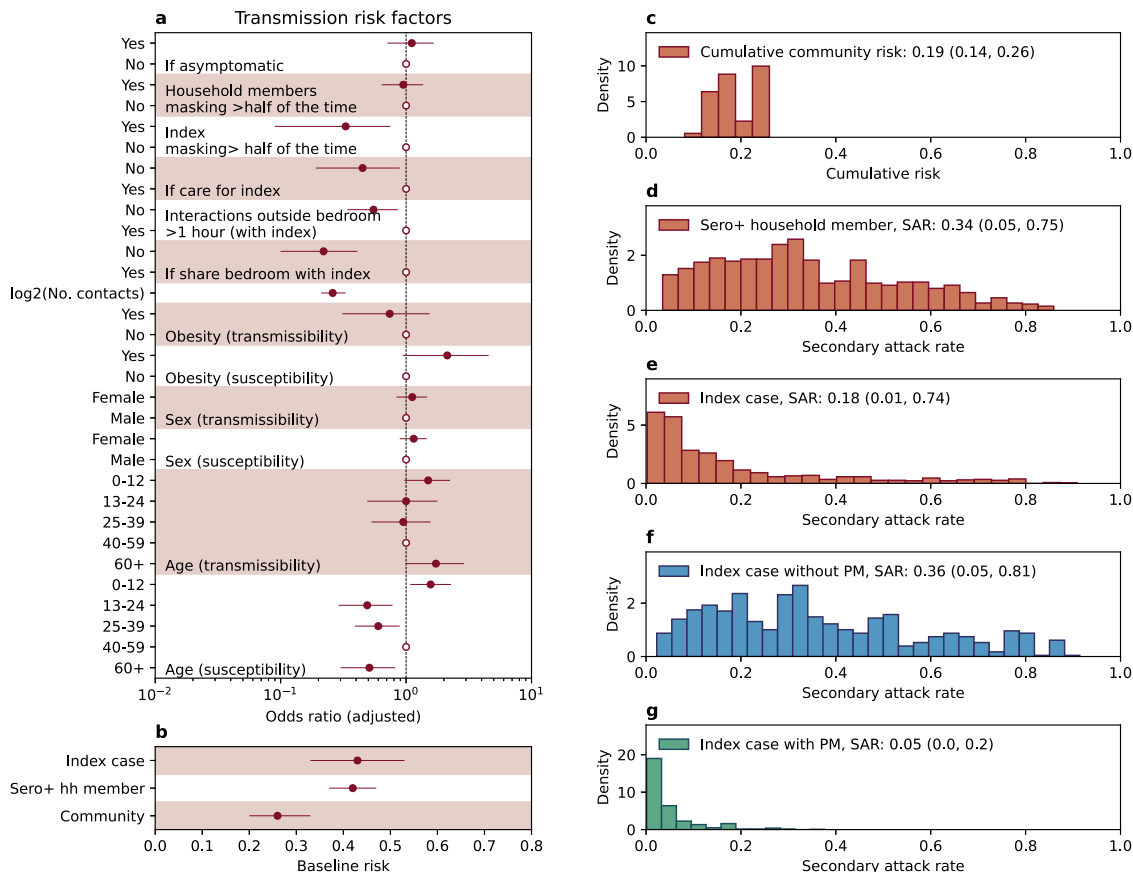


Fig. 2 Estimates from chain-binomial household transmission model. **a** Estimated odds ratios (adjusted) of the transmission risk factors fitted to 304 index cases and their 719 contacts. Solid dots and horizontal lines represent point estimates and 95% confidence intervals. Circles represent the reference class. **b** Baseline transmission risks from the index case and seropositive household members as well as baseline risks of acquiring infection from the community. **c-f** Distribution (histogram) of model projected community infection risk and household secondary attack rate across the study participants. **c** Distribution of cumulative community infection risks* **d** Distribution of the secondary attack rate attributable to seropositive household members who are not the index cases **e** Distribution of the secondary attack rate attributable to the index case. **f** Distribution of the secondary attack rate by the index case in a counterfactual scenario where no preventive measures (PM) were taken after diagnosis of the index case. **g** Distribution of the secondary attack rate by the index case in a counterfactual scenario where all preventive measures (PM) were taken after diagnosis of the index case. (*All results are from model with best fit to the data: model 15, Table S1).

Fitting chain-binomial household transmission model to the SARS-CoV-2 serologic data. The chain-binomial household transmission model fitted to serologic data suggested multiple risk factors associated with household transmission of SARS-CoV-2. We found that incorporating cumulative incidence rate in Costa Rica as a coefficient for community infection risk improved the fit of the model (Table S1 Model 1 vs Model 0), suggesting community infection risk correlates with SARS-CoV-2 circulation intensity outside the household. We found that asymptomatic index individuals were as likely to transmit SARS-CoV-2 as symptomatic index cases, indicating the significant contribution of asymptomatic transmission in the spreading of SARS-CoV-2 (Fig. 2a)¹⁶. Importantly, we found that behavioral factors were significant drivers of household transmission: sharing a bedroom with the index case (adjusted odds ratio not sharing vs. sharing: 0.22 with 95% CI (0.10–0.41), or caring for the index case (adjusted odds ratio not caring vs caring: 0.45 with 95% CI 0.19–0.89) were risk factors for transmission, while the index case wearing a mask (more than half of the time during the two weeks post diagnosis, adjusted odds ratio 0.33 with 95% CI 0.09–0.75) was protective. Avoiding interaction with the index case (<1 h) within two weeks of his/her diagnosis would reduce the risk by 45% (adjusted odds ratio vs >1 h: 0.55 with 95% CI: 0.34–0.86). Interestingly, whether household members wore a mask or not when interacting with the index case did

not significantly affect the risk of acquiring infection. Our model suggests that the number of household contacts had a strong negative association with the per-contact risk of SARS-CoV-2 transmission: doubling the number of contact numbers decreases the per-contact risk of transmission by 74% (95% CI: 67–79%). In addition, gender was neither significantly associated with SARS-CoV-2 susceptibility nor infectivity. We did not observe a significant association between age of the index case and SARS-CoV-2 infectivity but found a significant association between age of the household member and SARS-CoV-2 susceptibility: children under the age of 12 were significantly more likely to be infected when compared to age group 40–59 (OR 1.57, 95% CI: 1.08–2.28), while all other age groups were significantly less susceptible (Fig. 2a).

Utilizing the best-fitting estimate of the chain-binomial model, we projected the distribution of the community infection risk as well as the household secondary attack rate across all cohort participants. We estimated that the average cumulative community infection risk during the study period was at 19% (95% CI: 14–26%, Fig. 2c), lower than the household secondary attack rate attributable to seropositive household members (34%; 95% CI 5–75%, Fig. 2d). Interestingly, the average projected secondary attack rate by the index case was 18% (95% CI 1–74%), less than half of the secondary attack rate attributable to seropositive household members (Fig. 2e). This finding is explained by the fact

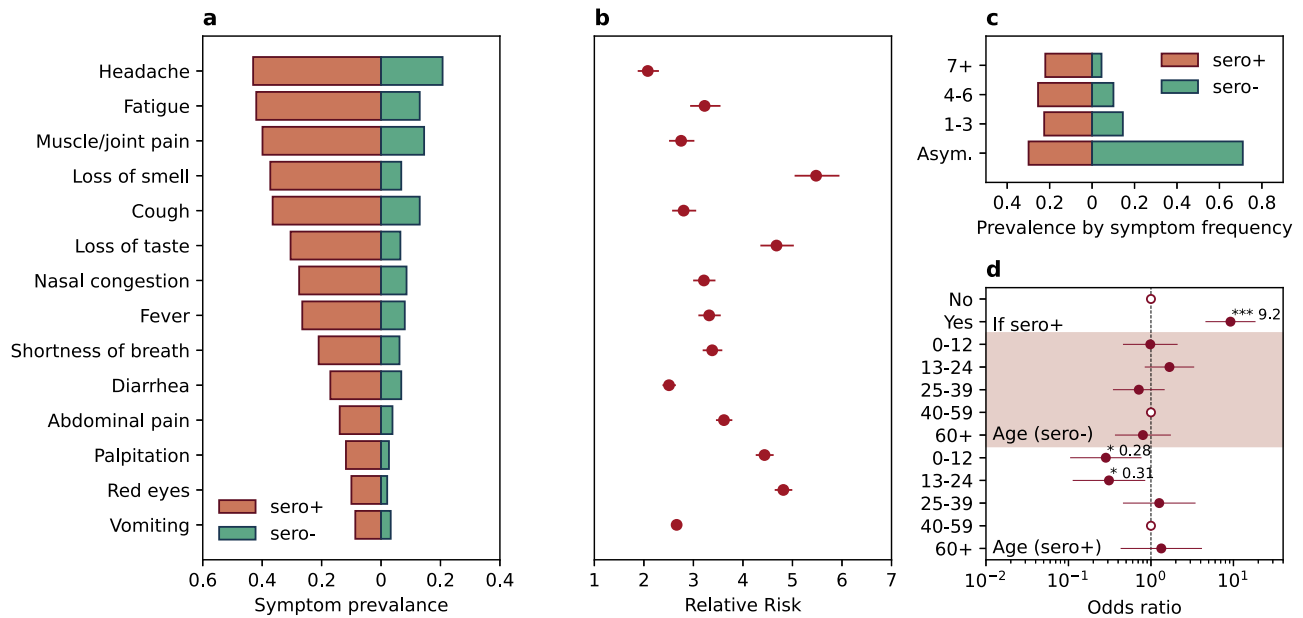


Fig. 3 COVID-19 related symptom presentations among both seropositive and seronegative individuals. **a** Prevalence of symptoms among both seropositive (red bar) and seronegative individuals (green bar). **b** The relative risk of symptom presentation in seropositive vs seronegative individuals. Panel **b** share the same y axis as panel **a**. Dots and horizontal lines represent point estimate and 95% confidence interval, based on the symptom presentations of 719 household contacts. **c** The prevalence of symptoms by symptom frequency (Asym. denotes asymptomatic individuals, 1-3 denotes individuals having 1-3 of all symptoms listed in **a**, 4-6 denotes having 4-6 of all symptoms listed in **a** and 7+ denotes having more than 7 symptoms listed in **a**. **d** Regression analysis on the risk of being symptomatic (having a least 1 symptom in **a**) by serologic status and age. “If sero +” denotes if the individual is seropositive; “Age (sero-)” denotes the age dependency of being symptomatic among seronegative individuals; Age (sero +) denotes the age dependency of being symptomatic among seropositive individuals. Dots and horizontal lines represent point estimate and 95% confidence interval, based on the symptom presentations of 719 household contacts.

that a significant fraction of the cohort population took protective measures after diagnosis of the index case, shown to be effective at reducing transmission, including avoiding sharing a bedroom, reducing interactions outside the bedroom, and wearing masks (Table 1). We also found that 30% of index cases were the source for 80% of all secondary cases’ onward transmission, indicating high transmission heterogeneity (Fig. 2e).

We further projected a hypothetical scenario in which the cohort population did not adopt preventive behavioral measures (all household members shared a bedroom with the index case, interacted with the index for more than 10 h outside the bedroom, took care of the index case who did not wear a mask most of the time). The projected secondary attack rate by the index case was 36% (95% CI 5–81%), comparable to the secondary attack rate attributable to seropositive household members (Fig. 2f). In this case, transmission heterogeneity would be much reduced (Fig. 2g), with 58% of the index cases being the source for 80% of onward transmission, suggesting variation in the adoption of preventive behavioral measures were major sources of index case’s observed transmission heterogeneity. If, on the contrary, all behavioral risk factors were avoided and preventive measures adopted, the secondary attack rate by the index case could be reduced to 5% (95% CI: 0–20%). We further conducted a sensitivity analysis of bootstrapping estimates at the households’ level, controlling for the joint distribution of household size and age category (Table 1), sex, and diagnostic month of the index cases (detailed in, to address potential household clustering effect (Fig. S4). The bootstrapping confidence intervals (Fig. S4) are wider than likelihood-ratio based confidence intervals (Fig. 2a), with the effects of index case mask wearing and duration of interaction outside the bedroom (with index case) become statistically non-significant. We also conducted a sensitivity analysis that excluded household with either index case or household members who had received vaccination prior or during the study, as shown in

Fig. S5. The results remain similar to the main results shown in Fig. 2a, with the effects of the duration of interaction outside the bedroom (with index case) become statistically non-significant.

Symptoms associated with SARS-CoV-2 seropositivity. We surveyed the presence of fourteen COVID-19 related symptoms independent of serostatus. The prevalence of each symptom is presented in Fig. 3a, along with the relative risk comparing seropositive to seronegative contacts in Fig. 3b. Overall, the prevalence was significantly higher for all 14 symptoms for seropositive individuals than for those who were seronegative (Relative Risk (RR) > 2, $p < 0.01$ for all 14 symptoms). For symptoms with a higher than 20% prevalence among seropositive individuals, loss of smell (RR = 5.5, 95% CI 5.0–6.0) and loss of taste (RR = 4.7, 95% CI 4.4–5.0) were the most predictive of SARS-CoV-2 infection. Seventy percent of seropositive individuals had at least one symptom, while only 29% of seronegative individuals reported at least one symptom (Fig. 3c). Logistic regression (Fig. 3d) of having at least one symptom against an indicator of seropositivity yielded an adjusted odds ratio of 9.2 (95% CI 4.6–18.5, $p < 0.001$). However, among seropositive individuals, the prevalence of symptom presentation differed significantly by age: persons aged 0–12 and 13–24 years were 72% and 69% less likely to be symptomatic (OR 0.28 with 95% CI 0.1–0.77 and 0.31 with 95% CI 0.11–0.85 respectively, $p < 0.05$ for both) compared with persons aged 40–59 years.

Discussion

Through fitting transmission models to household serologic data, we obtained estimates for both secondary attack rates and community transmission and identified significant behavioral measures for preventing household transmission in the pre-vaccination and pre-immune escape variant era. Although seroprevalence and household studies have been conducted in Latin

America^{17,18} our study is the first to estimate both household secondary attack rate and community infection rates, and to identify specific actionable preventive measures. This work adds to our knowledge of SARS-CoV-2 transmission in middle-income countries in Latin America, and more broadly expands our understanding of transmission in a variety of settings.

A highlight of our study is that it provides real-world evidence that preventive measures within the household, such as sleeping arrangements and reducing contacts outside the bedroom, as well as household members and infected individuals wearing masks, was significantly associated with reducing the risk of SARS-CoV-2 transmission within the household. Interestingly, our finding suggested that masks wearing by the index case is effective as “source control”. A recent household study conducted during the Omicron wave in four jurisdictions in the United States similarly found that attack rates were significantly lower among index cases who isolated or wore a mask¹. Our study emphasizes the importance of non-pharmaceutical interventions in reducing infection risk and disease burden in the household setting, especially when vaccines are not widely available or ineffective in preventing transmission.

Our study suggested that children aged <12 years were more likely to become infected. Age as a risk factor for susceptibility and transmissibility has been studied in numerous settings and a variety of designs; the effect of age is highly dependent on age-specific contact rates and is therefore difficult to disentangle from biologic effects¹⁹. In fact, several studies have observed a reduced susceptibility for SARS-CoV-2 among children, distinct from our finding^{2,20–22}. The increased susceptibility of the <12 age group in our study may be a function of behavioral factors, particularly time spent at home, as children in this age group are more likely to remain home under adult supervision and therefore have a higher risk of exposure to and in-home transmission from adult contacts.

Our study is unique in that both seropositive and seronegative individuals were asked about their symptom presentation around the time of the diagnosis of the index case, prior to the serum sample collection. The seronegative individuals served as a control group to assess symptom prevalence in non-infected populations as many of the COVID-19 related symptoms are nonspecific. Our results confirm the high rate of asymptomatic infections in the younger population and identified loss-of-taste and loss-of-smell as highly specific to SARS-CoV-2 in the pre-Omicron era.

To further explore the importance of household size and contacts, we tried models assuming logarithmic and linear relationships for the number of households contacts and found that the logarithmic model performed best (Table S1, Model 7 vs. Model 8). This suggests a power law relationship between household secondary attack rate and the number of household contacts n , i.e., $SAR \propto n^{-\gamma}$ where $\gamma = 1.7$ was estimated by our model. We found that the household secondary attack rate decreased when household size increased and that a power law relationship linked household secondary attack rate with the number of household contacts n , where

$$SAR \propto n^{-\gamma}$$

and $\gamma = 1.7$. This could be due to a dilution of household interaction intensity per household contact, whereby an individual in a large household has more household members for interactions than in a small household, and hence less propensity to interact with the index case.

The chain-binomial model suggested that the distribution of secondary attack rate by the index case is highly heterogeneous, with 30% of index cases being the source for 80% of all secondary cases' onward transmission (Fig. 2e). This heterogeneity was mainly driven by the partial adoption of the preventive measures. In the hypothetical scenario without any preventive measures (Fig. 2f), the transmission heterogeneity would be much reduced, with 6% of the index cases being the source for 80% of onward transmission. This suggests

that variations in the adoption of preventive measures contribute to the observed heterogeneities in SARS-CoV-2 transmission chains^{2,23}. Comparison of secondary attack rates across studies is limited by differences in study design, including infection ascertainment as well as follow-up and approaches for SARS-CoV-2 antigen or antibody testing. However, our secondary attack rate is somewhat higher than the SAR of 23.9% found in a large household cohort observed in South Africa from July 2020 to August 2021⁴, and is lower than that found in household studies in the United States, with SARs of 61% for Alpha variants and 55% for non-Alpha variants³.

Our study has several limitations. First, questionnaires related to behavioral factors (sharing a bedroom, interaction outside the bedroom, and caring for the index case) were only directed towards the interaction between the index individual and each household member. We could not evaluate how the interactions between (non-index) household members impact transmission. Second, we could not assess how variations in viral shedding duration and intensity across infected individuals could potentially affect transmission, as we did not collect respiratory samples from the participants. In particular, a recent study from South Africa has shown the importance of viral load and kinetics on SARS-CoV-2 household transmission²⁴. In addition, we only had the SARS-CoV-2 diagnostic date of the index case, thus the exact timing of the index case's infection remained uncertain. We also did not assess household ventilation parameters which could impact SARS-CoV-2 transmission risk within confined space²⁵. Another limitation is that our study relied on self-reported data, thus the data were subject to recall and/or social desirability biases. Finally, these estimates were from the first wave, and may not be generalizable to later epidemic waves with more transmissible or immune escape variants. However, these estimates serve as a baseline for future studies, and our findings regarding household prevention here are comparable to those found in the U.S. during the Omicron wave, suggesting the generalizability of the findings. In summary, our study from a middle-income country in Latin America points to relatively simple preventive measures to limit household transmission and suggests that simple behavioral mechanisms can explain the pervasive transmission heterogeneity reported in SARS-CoV-2.

Data availability

The investigators welcome inquiries about possible collaborations and requests for access to the dataset. Anonymized datasets as required can be shared after approval of a proposal and a signed data use agreement. All proposals for use of the data and related data use agreements will be approved by a data access committee including principal investigators from the National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH), the Costa Rican Agency for Biomedical Research (Agencia Costarricense de Investigaciones Biomédicas (ACIB)—Fundación INCIENSA (FUNIN)), and the Caja Costarricense de Seguridad Social (CCSS). Investigators interested in more details about this study, including protocols, informed consent forms, and data use agreement should contact the principal investigator and the corresponding author, D Rebecca Prevots (rprevots@niaid.nih.gov). Source data to reproduce the figures are available at²⁶ <https://doi.org/10.5281/zenodo.7793795>.

Code availability

Code to reproduce the figures are available at²⁶ <https://doi.org/10.5281/zenodo.7793795>.

Received: 20 September 2022; Accepted: 21 June 2023;

Published online: 22 July 2023

References

1. Baker, J. M. et al. SARS-CoV-2 B.1.1.529 (Omicron) variant transmission within households—four U.S. Jurisdictions, November 2021–February 2022. *MMWR Morb. Mortal Wkly. Rep.* 71, 341–346 (2022).

2. Sun, K. et al. Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. *Science* **371**. <https://doi.org/10.1126/science.abe2424> (2021).
3. Donnelly, M. A. P. et al. Household transmission of SARS-CoV-2 Alpha variant—United States, 2021. *Clin. Infect. Dis.* published online Feb. <https://doi.org/10.1093/cid/ciac125> (2022).
4. Cohen, C. et al. SARS-CoV-2 incidence, transmission, and reinfection in a rural and an urban setting: results of the PHIRST-C cohort study, South Africa, 2020–2021. *Lancet Infect. Dis.* published online March. [https://doi.org/10.1016/S1473-3099\(22\)00069-X](https://doi.org/10.1016/S1473-3099(22)00069-X) (2022).
5. Bi, Q. et al. Insights into household transmission of SARS-CoV-2 from a population-based serological survey. *Nat. Commun.* **12**, 3643 (2021).
6. Madewell, Z. J., Yang, Y., Longini, I. M. Jr, Halloran, M. E. & Dean, N. E. Household secondary attack rates of SARS-CoV-2 by variant and vaccination status: an updated systematic review and meta-analysis. *JAMA Netw. Open* **5**, e229317 (2022).
7. CDC. Quarantine and Isolation. Centers for Disease Control and Prevention. 2022; published online May. <https://www.cdc.gov/coronavirus/2019-ncov/your-health/quarantine-isolation.html> (accessed May 11, 2022).
8. Lineamientos Nacionales para la Vigilancia de la infección por Coronavirus (2019-nCoV). Ministerio de Salud Costa Rica. <https://www.ministeriodesalud.go.cr/index.php/vigilancia-de-la-salud/41-lineamientos-coronavirus/134-lineamientos-nacionales-para-la-vigilancia-de-la-infeccion-por-coronavirus-2019-ncov> (accessed June 20, 2022).
9. Fantin, R. et al. Estimating vaccine effectiveness against SARS-CoV-2 infection, hospitalization and death from ecologic data in Costa Rica. *BMC Infect. Dis.* **22**, 767 (2022).
10. Burbelo, P. D. et al. Sensitivity in detection of antibodies to nucleocapsid and spike proteins of severe acute respiratory syndrome Coronavirus 2 in patients with coronavirus disease 2019. *J. Infect. Dis.* **222**, 206–213 (2020).
11. Farhat, Z. et al. Reproducibility, temporal variability, and concordance of serum and fecal bile acids and short chain fatty acids in a population-based study. *Cancer Epidemiol. Biomarkers Prev.* **30**, 1875–1883 (2021).
12. Longini, I. M. Jr, Koopman, J. S., Monto, A. S. & Fox, J. P. Estimating household and community transmission parameters for influenza. *Am. J. Epidemiol.* **115**, 736–751 (1982).
13. Gordon, A. et al. Influenza transmission dynamics in urban households, Managua, Nicaragua, 2012–2014. *Emerg. Infect. Dis.* **24**, 1882–1888 (2018).
14. Mossong, J. et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med.* **5**, e74 (2008).
15. Mistry, D. et al. Inferring high-resolution human mixing patterns for disease modeling. *Nat. Commun.* **12**, 323 (2021).
16. Buitrago-Garcia, D. et al. Occurrence and transmission potential of asymptomatic and presymptomatic SARS-CoV-2 infections: a living systematic review and meta-analysis. *PLoS Med.* **17**, e1003346 (2020).
17. González, F. et al. Seroepidemiology of SARS-CoV-2 infections in an urban population-based cohort in León, Nicaragua. *Epidemiol. Infect.* **149**, e247 (2021).
18. Maier, H. E. et al. Clinical spectrum of SARS-CoV-2 infection and protection from symptomatic re-infection. *Clin. Infect. Dis.* published online Aug. <https://doi.org/10.1093/cid/ciab717> (2021).
19. Wang, X., Han, Q. & Kong, J. D. Studying the mixed transmission in a community with age heterogeneity: COVID-19 as a case study. *Infect. Dis. Model.* **7**, 250–260 (2022).
20. Davies, N. G. et al. Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat. Med.* **26**, 1205–1211 (2020).
21. Dattner, I. et al. The role of children in the spread of COVID-19: Using household data from Bnei Brak, Israel, to estimate the relative susceptibility and infectivity of children. *PLoS Comput. Biol.* **17**, e1008559 (2021).
22. Toh, Z. Q. et al. Comparison of seroconversion in children and adults with mild COVID-19. *JAMA Netw. Open* **5**, e221313 (2022).
23. Endo, A. Centre for the Mathematical Modelling of Infectious Diseases COVID-19 Working Group, Abbott, S., Kucharski, A.J., Funk, S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Res.* **5**, 67 (2020).
24. Sun, K. et al. SARS-CoV-2 transmission, persistence of immunity, and estimates of Omicron’s impact in South African population cohorts. *Sci. Transl. Med.* **14**, eabo7081 (2022).
25. Wang, C. C. et al. Airborne transmission of respiratory viruses. *Science* **373**. <https://doi.org/10.1126/science.abd9149> (2021).
26. Sun, K. Data and Code for “Behavioral factors and SARS-CoV-2 transmission heterogeneity within a household cohort in Costa Rica.” <https://doi.org/10.5281/zenodo.7793795> (2023).

Acknowledgements

This work was supported by the Intramural Research Programs of NIAID, NIDCR, NCI, and by the FIC. We wish to thank all participants in this study. We also thank Dr Aimee Kreimer for her support in implementing this study. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the NIH.

Author contributions

D.R.P.: Study conception, study design, data analysis, manuscript writing. K.S.: data analysis, manuscript writing, verified the underlying data. V.L., A.A., C.P., J.C.V., M.Z., M.M., C.A., A.A., A.H., R.H.—study design, data acquisition, data analysis, manuscript revision. V.L., J.C.V.—verified underlying data. M.H.G., R.P., C.V.—study design, data analysis, manuscript revision. J.I.C., P.D.B., M.A.A.—serologic testing, data analysis, manuscript revision.

Funding

Open Access funding provided by the National Institutes of Health (NIH).

Competing interests

The authors declare no competing interests.

Additional information


Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43856-023-00325-6>.

Correspondence and requests for materials should be addressed to Kaiyuan Sun or D. Rebecca Prevots.

Peer review information *Communications Medicine* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

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

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2023

for the RESPIRA Study Group

Amada Aparicio³, Karla Moreno³, Roy Wong³, Melvin Morera³, Arturo Abdelnour³, Alejandro Calderón³, Kattia Camacho³, Gabriela Ivankovich³, Adriana Yock³, Roberto Castro⁸, Rolando Herrero², Carolina Porras², Bernal Cortés², Viviana Loría², Rebecca Ocampo², Michael Zúñiga², Cristina Barboza-Solis⁹ & Romain Fantin⁹

⁸Ministerio de Salud, San José, Costa Rica. ⁹Universidad de Costa Rica (UCR), San José, Costa Rica.