

## Understanding Society User Support - Support #2117

### Household members in COVID-19 collection

06/05/2024 09:28 AM - Laura L

<b>Status:</b>	Resolved	<b>Start date:</b>	06/05/2024
<b>Priority:</b>	Normal	<b>% Done:</b>	100%
<b>Assignee:</b>			
<b>Category:</b>	COVID-19		
<b>Description</b>			
Good morning,			
I am trying to reconstruct the household members' relationships in the COVID-19 study. While the COVID-19 study documentation does not mention this explicitly and only focuses on how to identify couples, in principle I could use <i>egoalt</i> files from previous main waves of the study (e.g., wave 11, so <i>k_egoalt</i> file) and identify household members through <i>pidp</i> and <i>apidp</i> variables. Is there any issue in proceeding this way?			
Thank you in advance.			
Best, Laura			

#### History

##### #1 - 06/07/2024 03:21 PM - Understanding Society User Support Team

- Category set to COVID-19
- Status changed from New to Feedback
- Assignee deleted (Understanding Society User Support Team)
- % Done changed from 0 to 80
- Private changed from Yes to No

Dear Laura,

For waves 1-2 of the Covid-19 survey linking this from the mainstage data (for different people the nearest mainstage interview might be in different waves) is one possible strategy, and the method you're proposing sounds good. For wave 3-9 of the Covid-19 Survey you can consider using the Household relationships module (see the questionnaires <https://www.understandingsociety.ac.uk/documentation/covid-19/questionnaires/>), you can also link this back to wave 1-2 for some people.

I hope it helps.

Best wishes,  
Piotr Marzec  
UKHLS User Support

##### #2 - 06/21/2024 01:21 PM - Laura L

Dear Piotr,

Thank you very much for your feedback, this is now clear.

All the best,  
Laura

##### #3 - 07/16/2024 04:34 PM - Understanding Society User Support Team

- Status changed from Feedback to Resolved
- % Done changed from 80 to 100

##### #4 - 02/15/2025 04:01 PM - Laura L

Good afternoon,

Sorry for the follow-up question, but new doubts have emerged while reconstructing families.

I am using the COVID-19 collection, wave 6 (November 2020). I found that linking family members using the information on cohabitants provided by respondents within the COVID-19 study can lead to imperfect matching, as it is not based on identifiers but rather on respondents' reported

information about the age and sex of their cohabitants. Given this limitation, I was planning to use Calendar Data 2020, as it ensures that the data are collected as close as possible to the November 2020 COVID-19 wave.

To this end, since Calendar Data contains repeated values for some respondents, I selected the most recent record (in terms of wave) for each respondent and used it to link family members in the COVID-19 data via the *egoalt* file of Calendar Data 2020.

Do you think this procedure is viable? Or do you believe that filtering out repeated respondents in Calendar Data 2020 could introduce issues, even if it is only used for linking purposes?

Thank you in advance for your support and clarification.

Best wishes,  
Laura

**#5 - 02/21/2025 05:18 PM - Understanding Society User Support Team**

Hi Laura,

Have you tried identifying which mainstage wave is closest in time by comparing the interview dates, the interview COVID date, and the mainstage interview dates?

Best wishes,  
Piotr Marzec  
UKHLS User Support

**#6 - 02/24/2025 10:05 AM - Laura L**

Dear Piotr,

Thank you for your reply. This could be also an option. I was just wondering in principle and more generally if it is a fair procedure to pick just one respondent-wave out of the two in Calendar Data, but apparently yes. Thanks!

I have an additional concern: when using the household level identifiers from wave 10 (*j\_hidp*) provided in the COVID-19 data, I see that there are respondents in households *with more than one person*, who completed the wave 10 questionnaire (*j\_ioutcome == 11*) but are not present in either *j\_egoalt*, *j\_indall* or *j\_indresp*. Why is it the case?

Thank you very much again!

Best,  
Laura

**#7 - 02/25/2025 05:52 PM - Understanding Society User Support Team**

Hi Laura,

In principle, the Calendar Year datasets are derived from the mainstage data, so you shouldn't expect any discrepancies between the two, though you would need to verify this to be certain.

The 2019 Mainstage data, which you download alongside the Covid-19 Survey data, was released to facilitate analysis and provide users with access to the interviews from the mainstage survey, which were not yet available at the time of the first Covid-19 dataset release. However, it has not been updated since 2021, making discrepancies with the current mainstage releases more likely.

Therefore, I recommend using the most recent release of the main UKHLS survey.

Best wishes,  
Piotr Marzec  
UKHLS User Support

**#8 - 02/26/2025 10:06 AM - Laura L**

Dear Piotr,

Thank you for your reply. I see some discrepancies between the *j\_hidp* provided within the COVID-19 survey and the *j\_hidp* in the mainstage wave 10, but I will trust more *j\_hidp* in wave 10 for my analysis following your advice.

Thank you again.

Best wishes,  
Laura