PROTECT Researcher Symposium 5-6 May 2022

PROTECT
A COVID-19 National Core Study

Using WGS-enriched contact tracing data at national scale to understand SARS-CoV-2 transmission

Charlie Turner, Contact Tracing Data Team, UKHSA
WNCoV.datacontacttracing@phe.gov.uk

Source attribution from contact tracing data: questions

Key question:

- What is the source of infection for each case of COVID-19 in England?

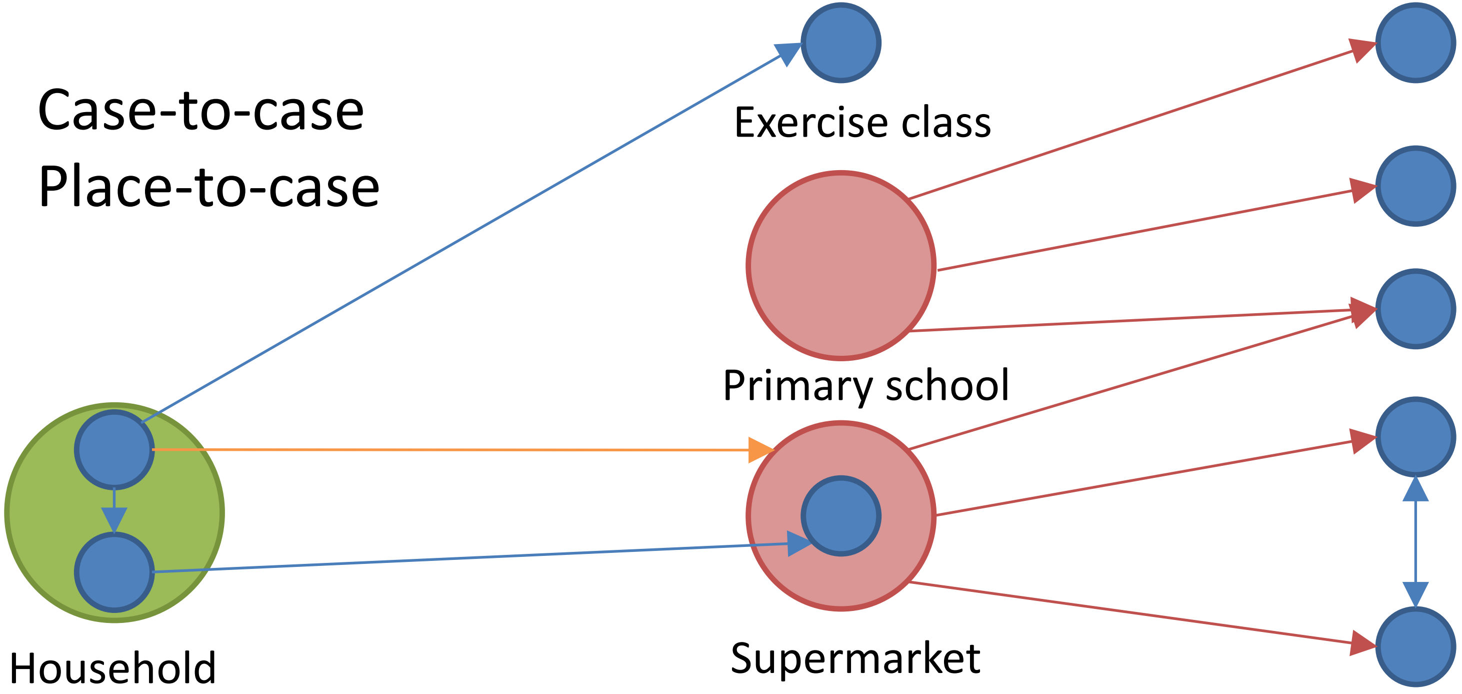
Then:

- How much transmission can be detected (… and prevented) by contact tracing?
- If contact tracing restarts, could we build real-time best-guess source attribution?
    - How early can we warn contacts that they may have been exposed?
- Which settings contribute most to transmission? What other factors are predictive of transmission?

Source attribution from contact tracing data: methods

Case-to-case

Place-to-case



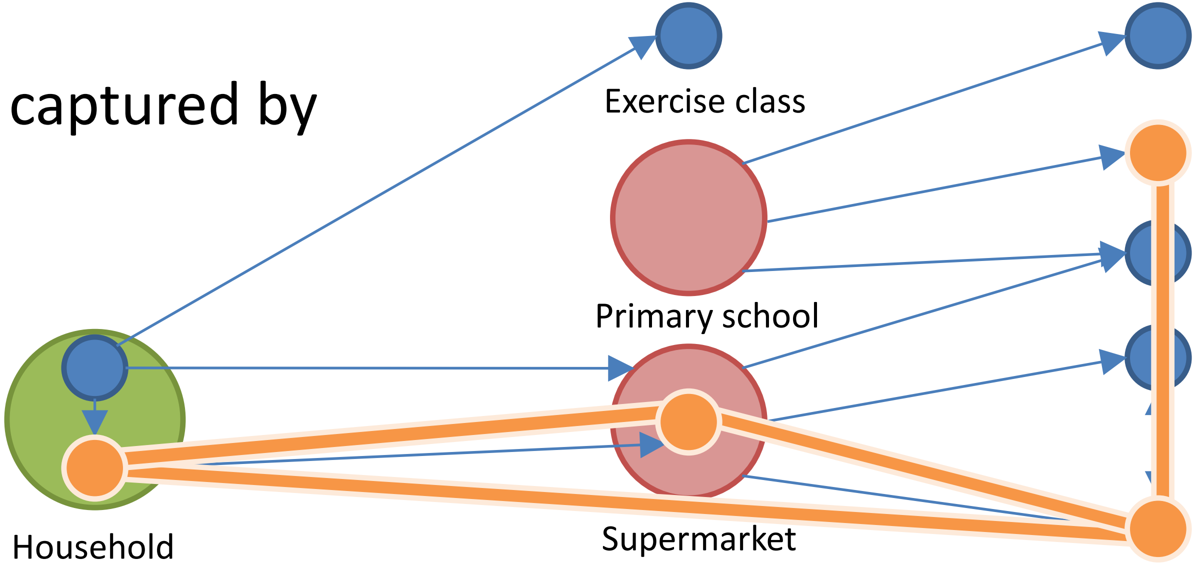
Source attribution from contact tracing data: methods

- Named contacts go on to become cases (2-14 days from exposure)
- Cases with same home address (2-14 days between cases)
- ‘Backwards’ common exposures: multiple cases reporting same location on same date, 3-7 days before symptoms or test.
- ‘Forward-Backward’ links: multiple cases reporting same location on same date,
    - at least one attended whilst infectious (from 2 days before symptoms or test onwards) and
    - at least one attended in period 3-7 days before symptoms or test.

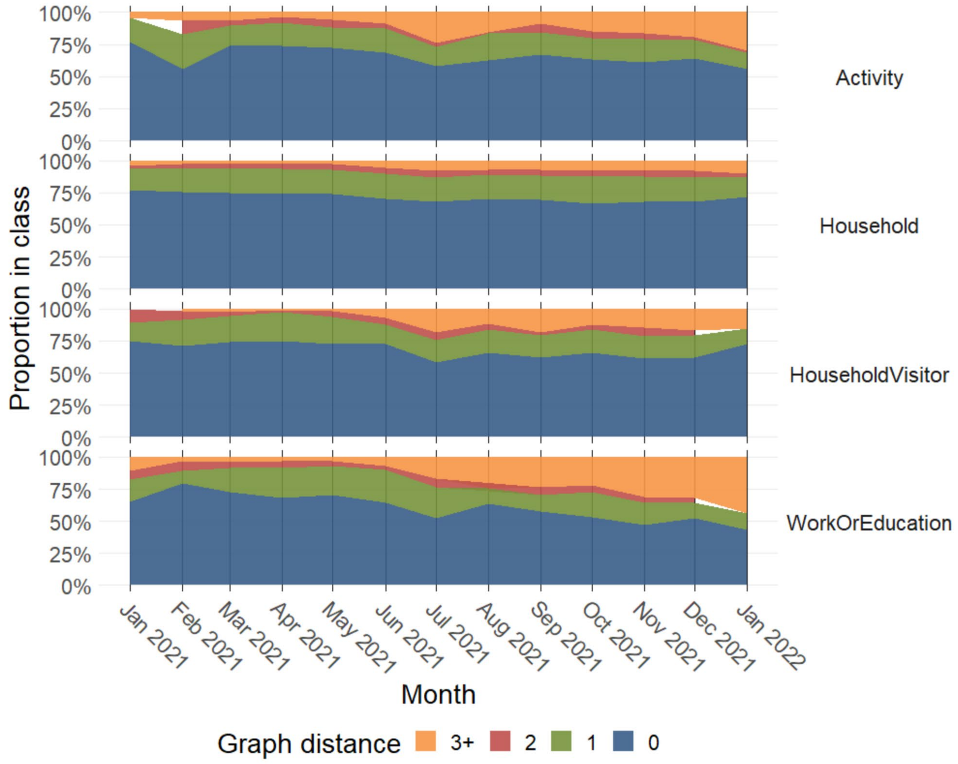
Source attribution from contact tracing data: methods

Overlay phylogenetic information for 1.4 million cases (>8.5%)

- Validate links
- Discover links not captured by contact tracing



Validation of ‘named contacts’ transmission links



- ‘Distance’ is count of Single Nucleotide Polymorphisms (SNPs) between exposing case and secondary case sequences.
- High validation of links from named Household contacts (90% with distance ≤ 1) lower for named contacts in Work/Education (82%) and Activities (84%) settings.

PROVISIONAL RESULTS

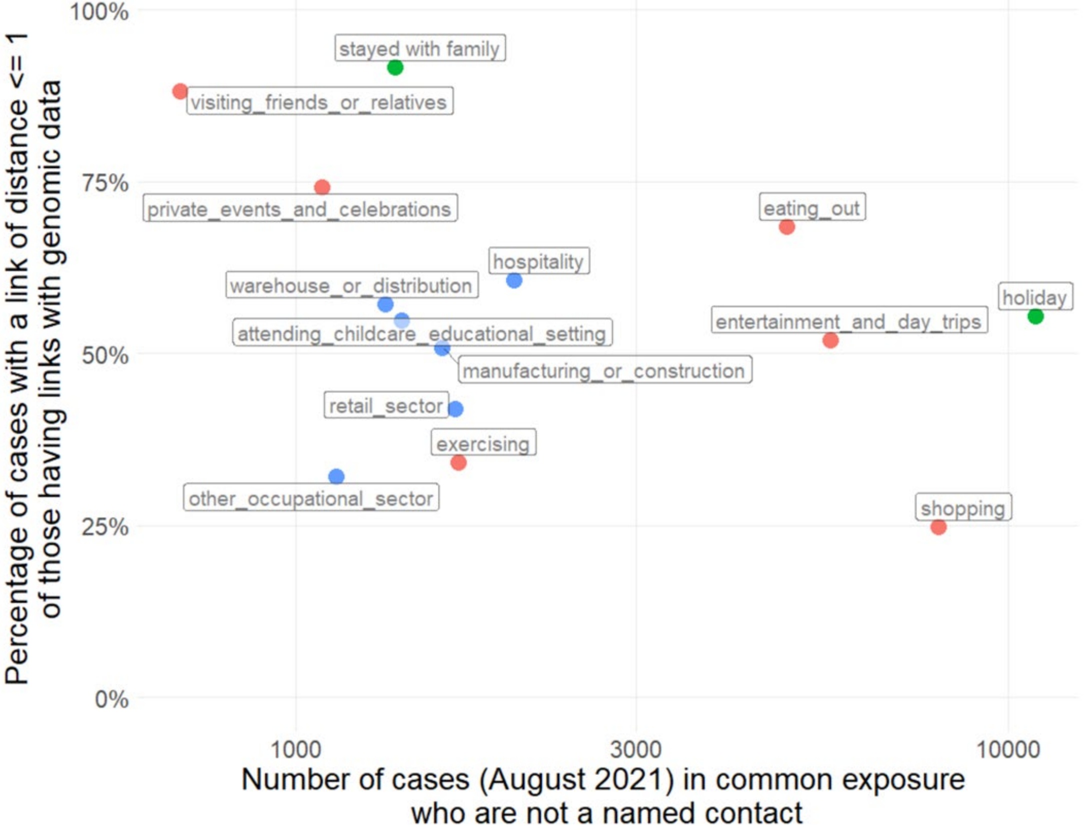
PROTECT
A COVID-19 National Core Study

Coverage and validation of transmission links by type of link

- For test period (August 2021), able to attribute a source for ~50% of cases.
- Named contacts and residential links are best validated.
- Common exposures and forwards-backwards links add more transmission links, enriching the network, but with lower confidence.

PROVISIONAL RESULTS

Coverage and validation of common exposure links by setting



- Common exposures can be used to assign place-case transmission links to many cases who were not named as contacts.
- Largest contributions of additional links from ‘holiday’ and ‘shopping’.
- Genomic validation of place-case links from common exposures varied significantly by setting.
- Home visits and private events had highest validation, shopping, exercising and “other” work settings had lowest validation.

PROVISIONAL RESULTS

Summary

- Used genomic data to validate transmission links derived from contact tracing data at an individual level and on a national scale.
- Transmission links found for ~50% of cases from contact tracing data, using combination of methods and across varied settings.
- Common exposures ‘fill in (some of) the gaps’ between households to better understand the network, and these links have some genomic validity.
- Further work on use of this data for research into settings, risk factors for transmission, surveillance and evaluation of contact tracing.

PROVISIONAL RESULTS