- Sampling bias. Who is selected for testing and which samples get sequenced may in turn result in a biased sample therefore what is highlighted might be what is most observable rather than most important. It is important to be mindful of this when comparing different studies. Care homes where richer data are available may be different to those with less data. For example, less testing might have been carried out on homes that had smaller or more drawn out outbreaks.

- It is difficult to analyse causality in the context of a rapidly changing environment. Researchers have noted in different studies that testing and intervention policy was changing during the first wave as different measures were introduced on visiting / general lockdown / enhanced PPE. Therefore, outcomes may not reflect causation or current risks.

- The quality of the data may affect the ability to deliver retrospective research studies. Issues around processing that affect the performance of viral sequencing have meant that some potentially useful data has been lost. However, a new Vivaldi study proposes to look at retrospectively combining the Vivaldi 2 cohort study data with HES data, ONS deaths, FSHC data. All studies highlighted that genomic data need to be analysed alongside detailed epidemiological data, creating a need for better linkage, multidisciplinary working and rapid overcoming of governance issues.

- In preparation for the second wave more data will be collected and linked. For example to explore in more detail the variation in sero-prevalence that has been observed. It may be that certain areas have had higher density of testing, which may therefore be more conducive to further analysis. Future studies should plan for both retrospective and prospective analyses and should be strategically aligned, prioritised with clear protocols on investigations and coordination. An assessment of seroprevalence of residents before the second wave, and the selection of care homes to track outcomes in detail based on current status would be a benefit.

4 Key findings

4.1 Although staff-to-staff transmission has been observed to have been a contributory factor in specific outbreaks, it is important not to generalise to all outbreaks and emphasise one route over another without clear evidence - studies undertaken so far indicate that multiple introductions are common. Clusters of cases have been observed but many outbreaks involve cases that are spread out over a longer period, again indicating multiple introduction routes / different lineages.

4.2 Retrospective genomic analysis and seropositive studies in care homes therefore find **evidence for multiple routes of virus ingress to care homes, but are not systematic enough to quantify the relative frequency of different routes of ingress.** Furthermore, these studies **do not definitively rule out any mode of ingress** so staff, visitors, visiting professionals, hospital discharges and new admissions and persistent infections may all contribute to the introduction of disease.

4.3 The weight of evidence is stronger in some areas than others, however. Evidence of staff to staff transmission has emerged in the genomic analysis (high confidence). Weak evidence on hospital discharge and modelling the impact of visitors (see also para 6 below) does not suggest a dominant causal link to outbreaks from these sources.

4.4 Public Health measures that reduce community incidence could be effective in reducing ingress into care homes. In the first wave, care homes appeared to show less impact from lockdown measures, with numbers continuing to rise more than cases in the community. It is thought that greater internal infection control measures now in place will make it easier to understand where infection is coming from.

4.5 Asymptomatic or atypically symptomatic presentation in residents[1] and staff mean that ingress may be hidden for a number of generations of disease.

4.6 Sequencing community tests to understand the comparator population is critical for the future.

5 Recommendations and barriers to improve virological analysis

5.1 Future research considerations identified

- Further work on the longevity and quality of antibodies. What proportion of care homes that experienced outbreaks in the first wave have evidence of sufficient resident and staff exposure to likely reduce the risk of further outbreaks or limit their size? Does the proportion of residents and staff with antibodies reduce over time due to waning of immunity or resident and staff turnover bringing in susceptible individuals? Cohort studies following particular homes in detail will be important, such as the Easter 6 study.

- How does the physical layout of a home influence transmission? Have homes that have cohorted residents onto different floors, been able to arrest the spread of the virus? How elements of ventilation & heating impact on transmission potential.

- Surveys on behavioural aspects of controlling the virus are important, for example a qualitative study interviewing staff to get further insight into how ways of working have changed.

- Exploration of severity of outcomes and what impacts this, looking at how severity varies with different co-morbidities etc.

5.2 Barriers and opportunities identified:

a. The structural set up of testing arrangements can have a major impact on the ability to analyse sequencing. The location for sample processing appears fixed without opportunity for dynamic decision making and sample redirection. **It is recommended that SAGE endorse processes to**

[1] Wide ranging symptomatology may be present in this population (See SAGE paper *Community case definitions for Covid-19*) and coupled with inability to express symptoms because of cognitive impairment or communication problesm.