Comparing temporal trends in the demographics of S+ and S- COVID cases

Christopher Overton, Julia Gog, Matt Keeling, Ellen Brooks-Pollock, Lorenzo Pellis, Leon Danon,
JUNIPER Consortium

Executive Summary

- Patterns of ethnicity over time for S+/S- varies dramatically from place to place
- In some regions (Bolton, Leicester, Birmingham, Blackburn, and Manchester), S+ cases have disproportionately affected individuals with an Asian ethnicity since early April. The ethnicity distributions are slowly converging back to that of S-.
- Other regions only observe a small discrepancy between S+ and S- ethnicity distributions (Sefton, Nottingham, Wigan, Bedford, and Central Bedfordshire) before becoming re-aligned.

Summary

- Demographic distributions of cases, and how they change over time, provide insight into epidemiological processes and events (e.g. changes in mixing patterns, importations, vaccine escape).
- We investigate temporal discrepancies between S-gene positive and S-gene negative demographic distributions. We focus on small spatial and temporal scales, due to test coverage and epidemic severity varying in space and time.
- S positive age distributions were initially perturbed due to importation of travel associated cases. After this, they have returned to the same age distribution as S negative cases.
- In many regions (e.g. Bolton, Blackburn, and Birmingham), the rise in S gene positive cases appears to have affected individuals with Asian ethnicities disproportionately, with the proportion of cases with an Asian ethnicity increasing in S positive cases from late March, which is not reflected in the S negative cases.
- This is consistent with B.1.617.2 emerging through travel related cases.
- The ethnicity distribution has not yet returned to that of S negative cases, suggesting there is a strong ethnicity bias on the current spread of B.1.617.2. However, the proportion of cases in individuals with a White ethnicity is growing, suggesting we may soon converge to the S negative distribution.
- Once these distributions converge, we should be able to obtain more reliable transmissibility estimates, since age and sex distributions have already converged.
- In other regions (e.g. Wigan, Bedford, and Sefton), there is less of a shift to individuals with an Asian ethnicity. This suggests that outbreaks in these areas may have been started by very few imported cases (either through international travel or local travel), after which community transmission has been taking place.

- Perturbations to the IMD score distribution are less obvious than those in Age and Ethnicity.

Methods

- We focus on 10 regions that have had over 200 S gene positive cases since 01/04/2021.
- In the figures, the left column considers ethnicity, the central column considers Index of Multiple Deprivation score, and the right column considers age.
- The stacked bar charts show the S negative distribution on top and S positive below.
- For IMD and ethnicity, we use the chi-squared test to estimate whether there is a significant difference between the samples. For age we use the Wasserstein distance.
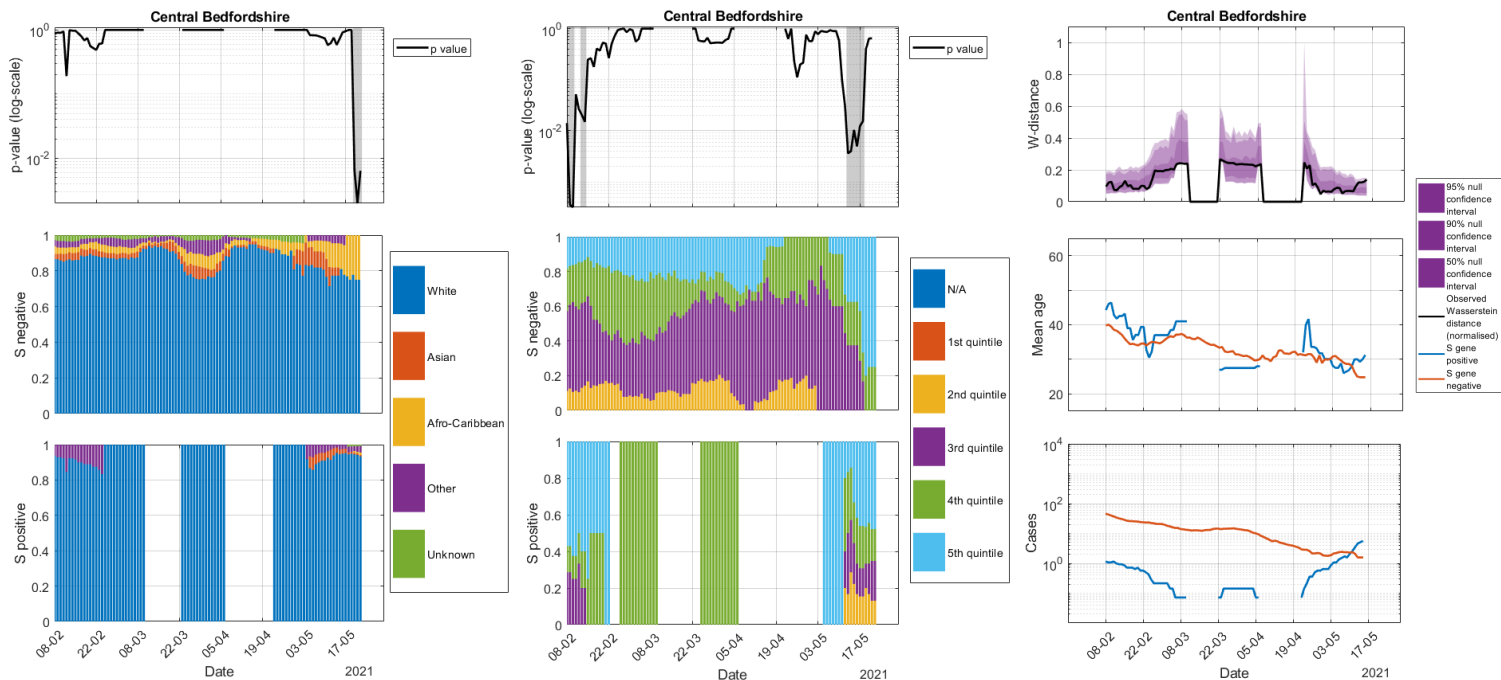
Results

Figure 1 - Central Bedfordshire: Relative age distributions are constant over time. Recent ethnicity sees S negative growing in individuals with Afro-Caribbean ethnicities and S positive cases growing in individuals with White ethnicities, with a significant difference between the two distributions. S negative cases are increasingly represented in individuals from lower deprivation areas, whereas the IMD distribution of S positive cases appears constant.
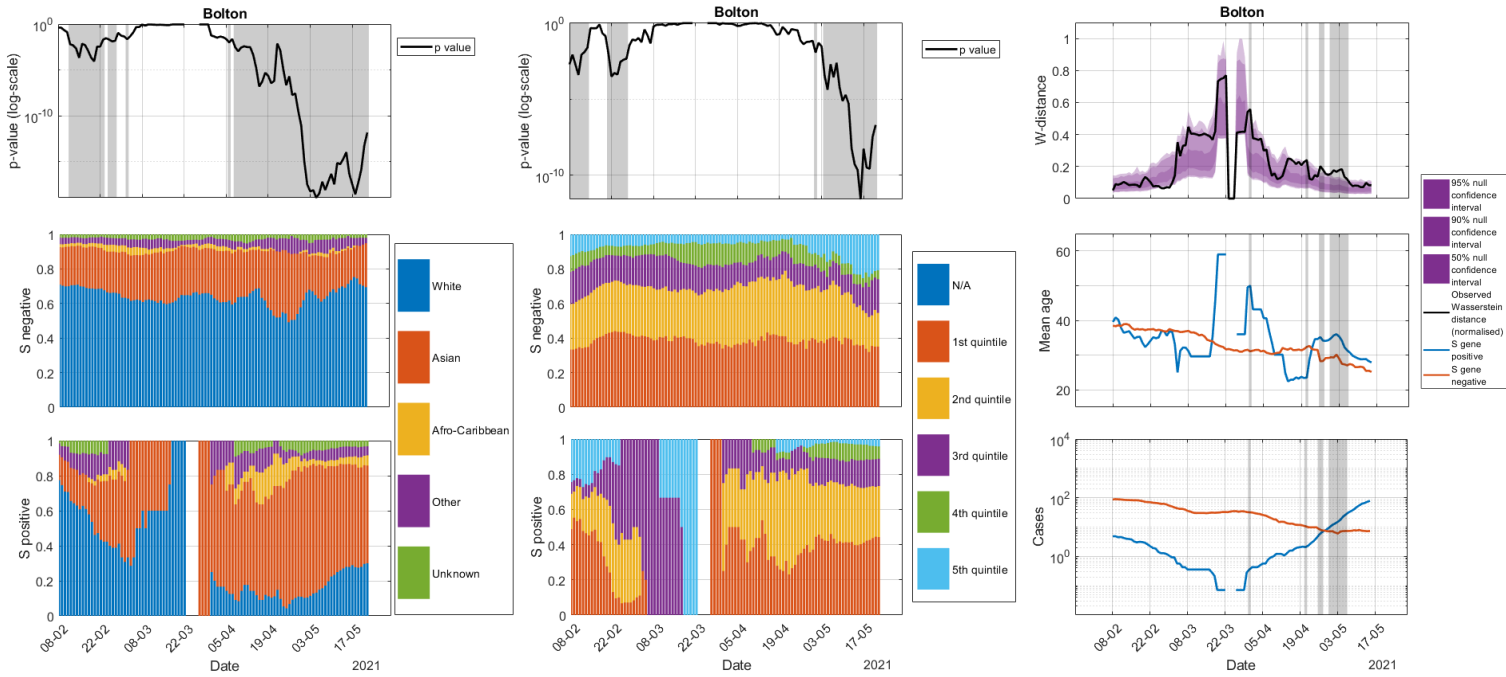
Figure 2 - Bolton: Since the end of March, the S positive ethnicity distribution is skewed towards individuals with an Asian ethnicity, whilst the S negative ethnicity distribution is skewed towards individuals with a White ethnicity. The p-value is gradually increasing, suggesting the distributions are becoming closer. The S positive IMD distributions appears relatively constant over time, whilst S negative cases are increasing arising in lower deprivation areas. Around the peak discrepancy between the ethnicity distributions (late April), there is a significant difference between the two age distributions, which was likely driven by travel associated cases.
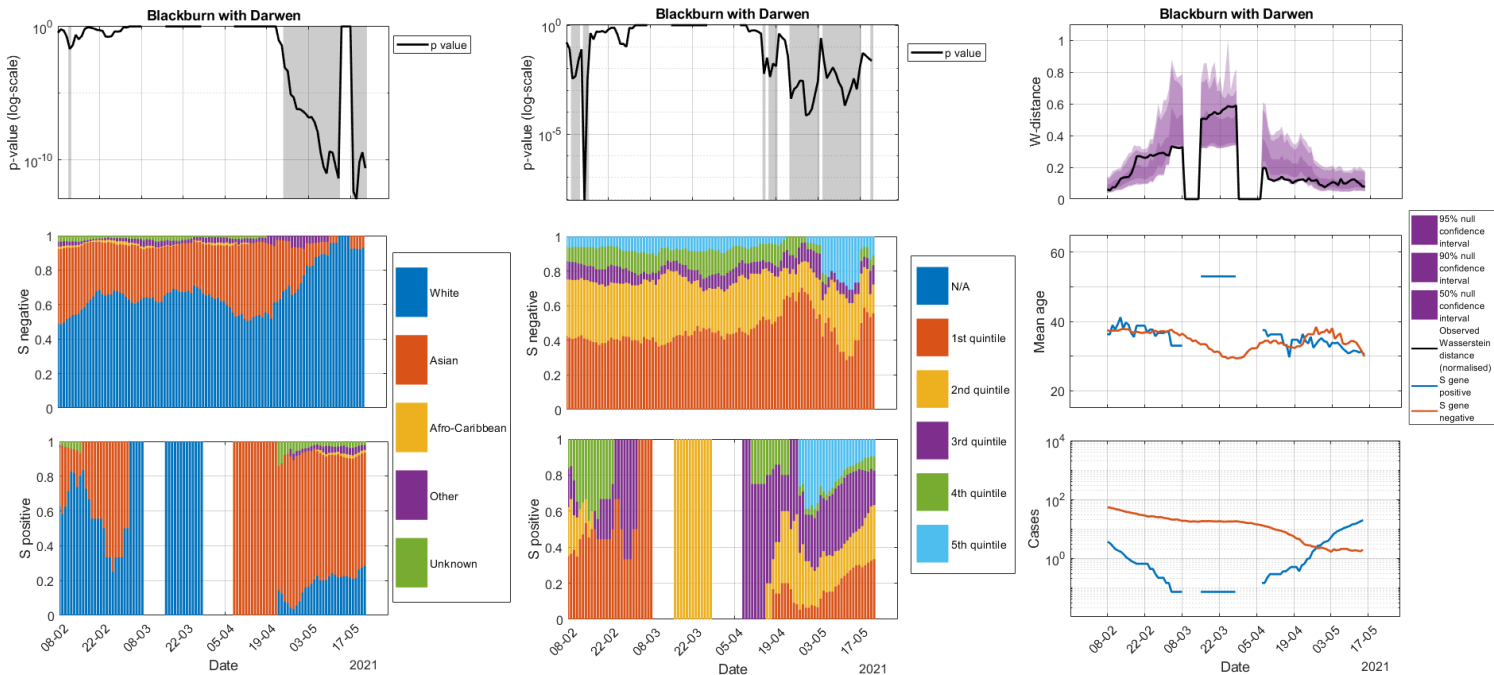

Figure 3 - Blackburn with Darwen: Since the start of April, S positive cases have mostly occurred in individuals with an Asian ethnicity, whereas S negative cases are gradually occurring more

frequently in individuals with White ethnicities. The age distributions have been consistent across the whole time series, suggesting many cases are likely to be community acquired rather than travel associated. Both S positive and S negative have observed a recent increase in the proportion of cases from lower deprivation areas. However, the timing of these peaks differs, leading to a significant difference between the distributions.
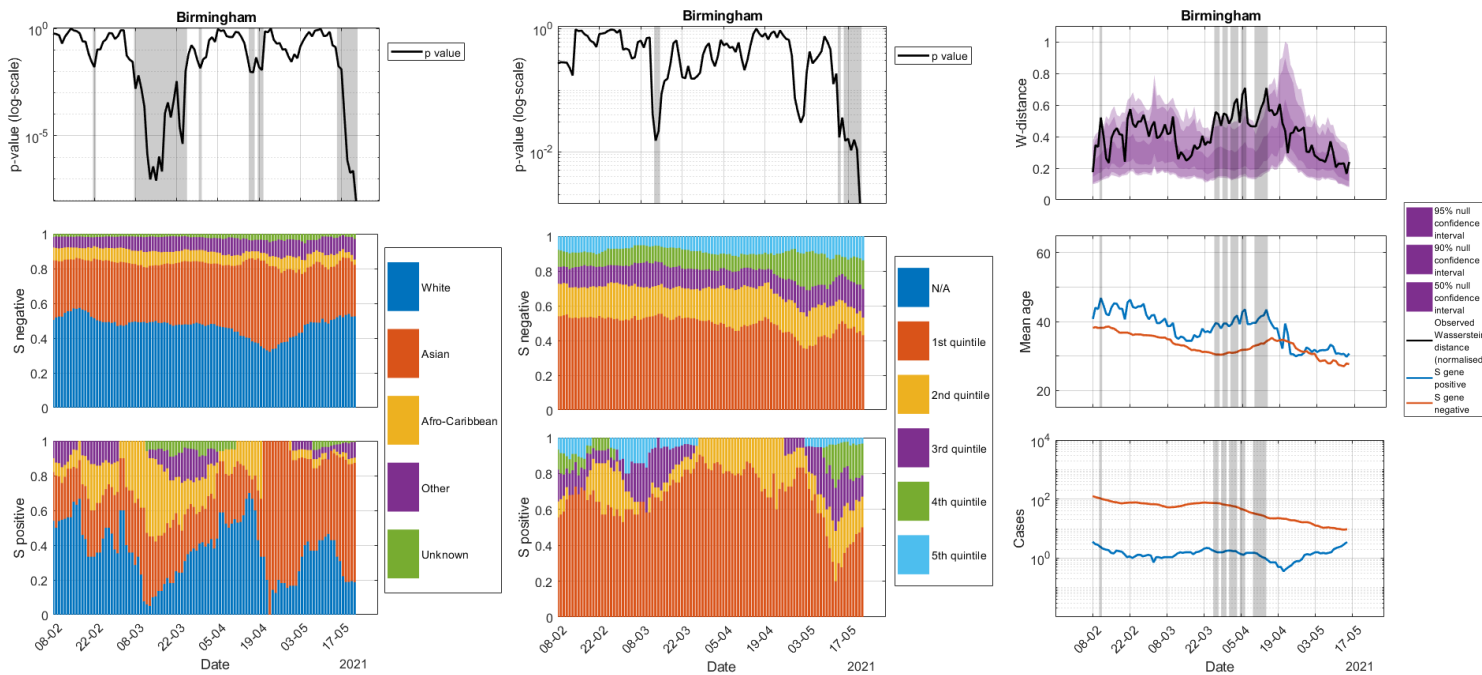

Figure 4 - Birmingham: In early March, the S positive cases observed a rise in individuals with an Afro-Caribbean ethnicity, which was not observed in S negative cases. This could be driven by B.1.351, which was spreading at the time prior to the importation of B.1.617.2. From late March, the proportion of Afro-Caribbean ethnicities in the S positive cases started to decline. From mid April, the proportion of cases from Asian ethnicities started to rise, leading to a significant difference between the current distributions. The IMD distributions have observed a recent increase in the proportion of cases from lower deprivation areas in the S gene positive data, with the S gene negative distribution remaining constant. In the age distribution, a significant difference was observed during early April.
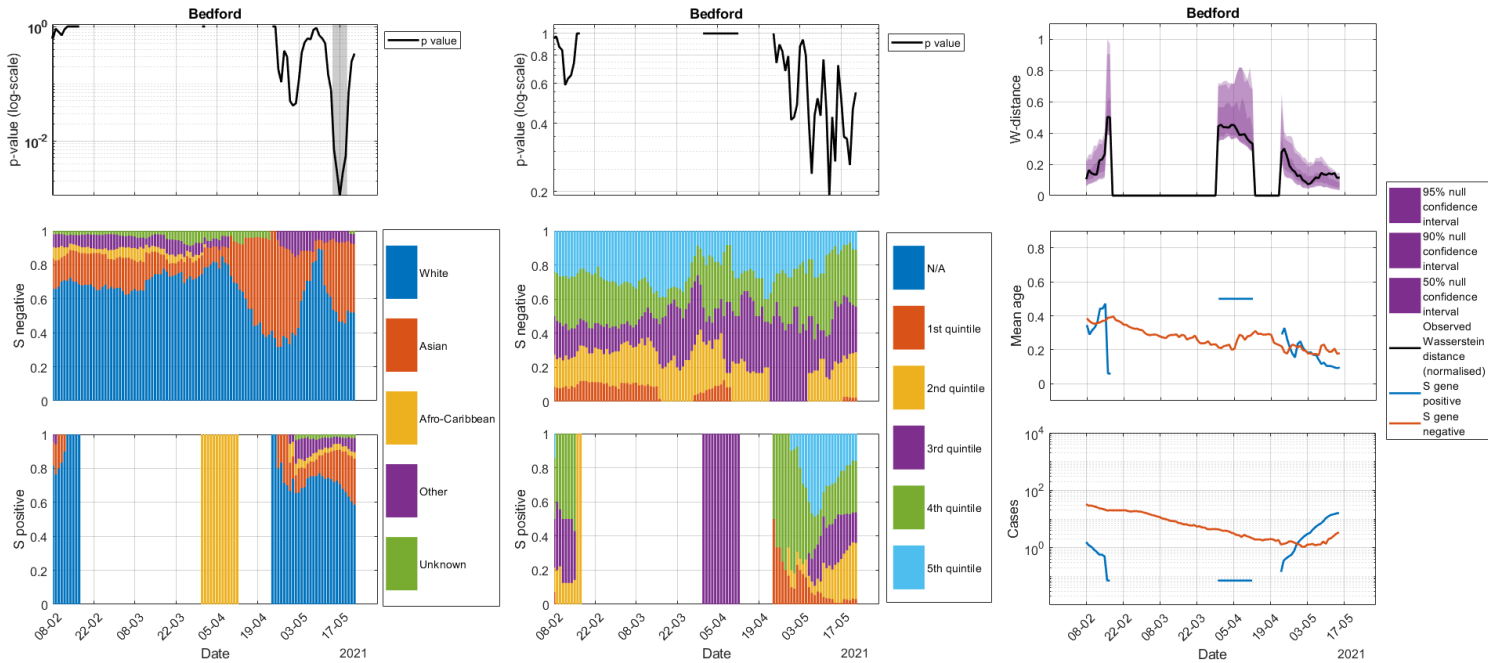
Figure 5 - Bedford: There is a short-lived difference between the ethnicity distributions. S negative cases observed a spike in individuals with an Asian ethnicity during April, prior to the arrival of B.1.617.2. With the low case numbers, there is no significant difference in the IMD distribution. The age distributions are consistent over time.
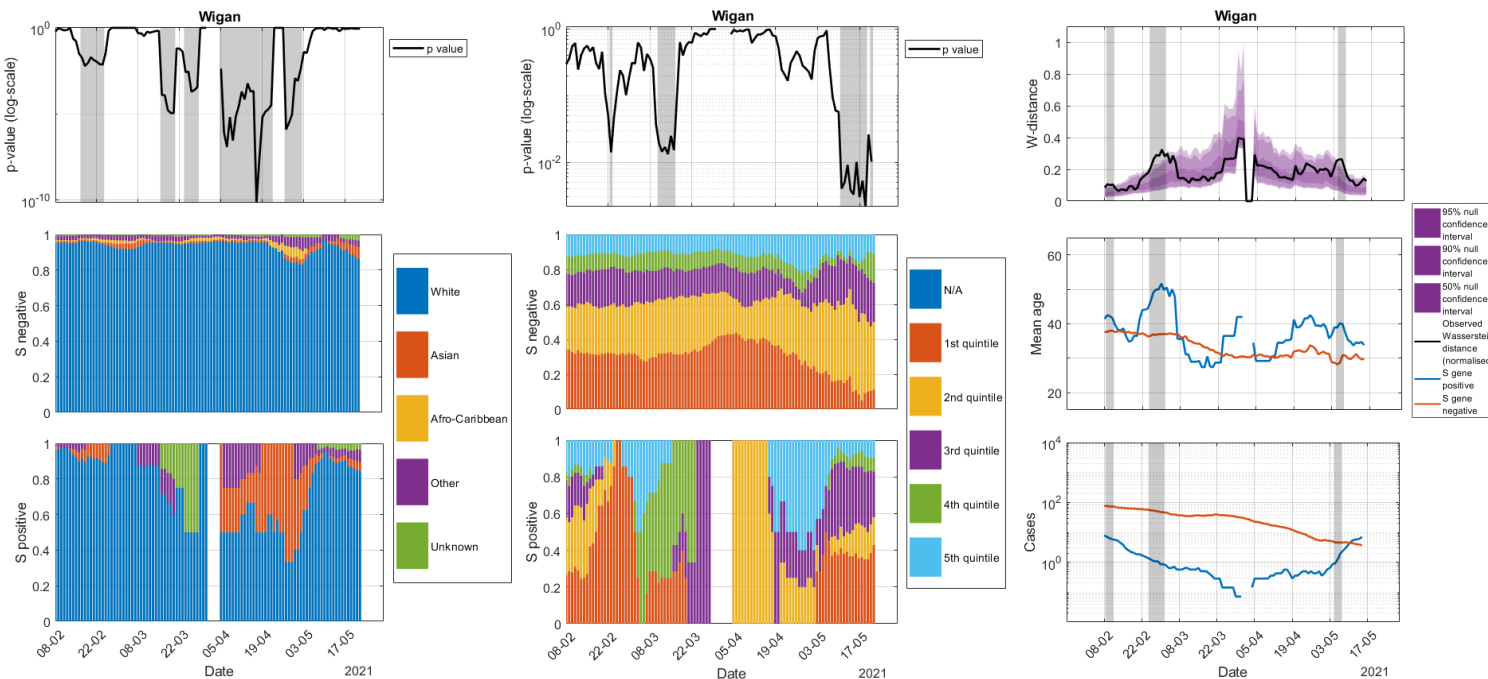

Figure 6 - Wigan: During April there is a significant difference in the ethnicity distributions, with S gene positive cases having a higher proportion of individuals with an Asian ethnicity. Since there is no significant perturbation to the age distribution, this suggests cases may be spreading from Bolton rather than importation through travel. The ethnicity distributions are now closely aligned, suggesting that transmission has quickly spread through the different ethnicities in the community. S gene

positive cases recently have a higher proportion of individuals from higher deprivation areas than S gene negative cases.
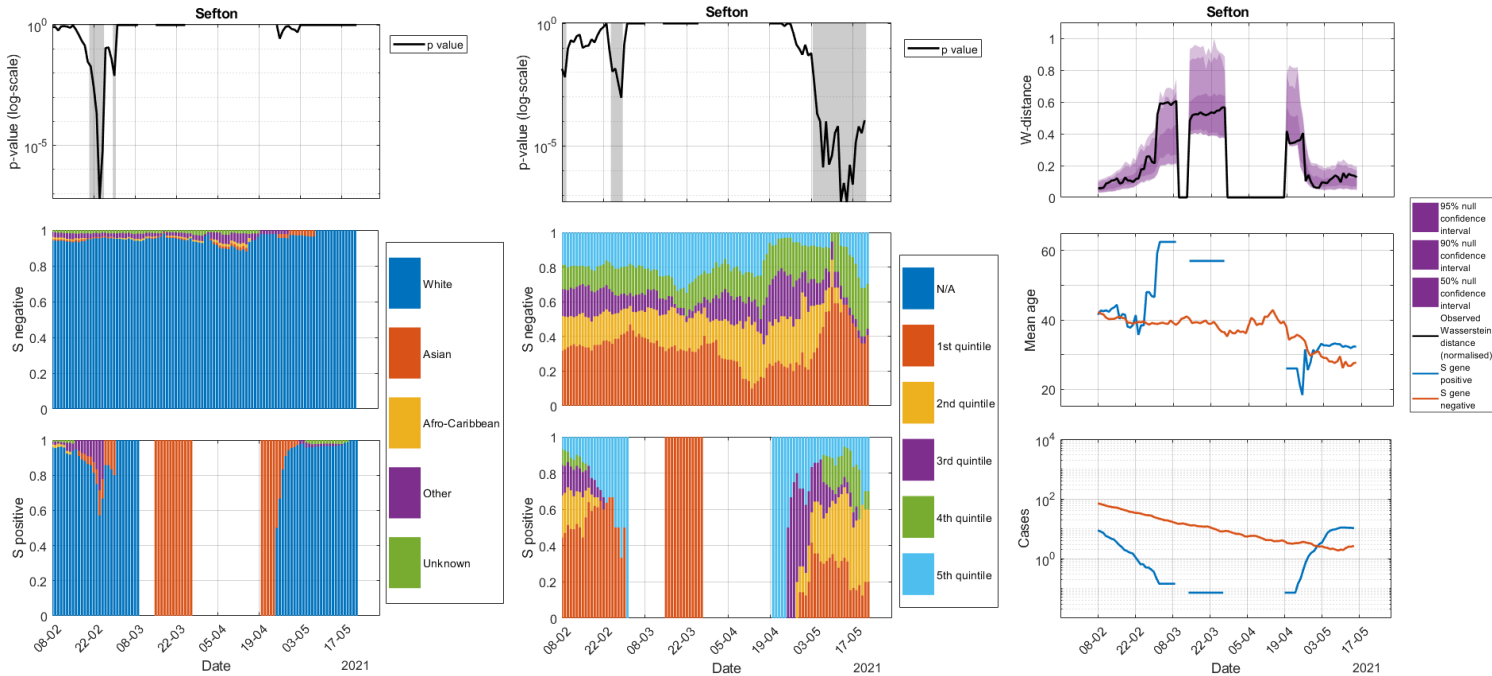

Figure 7 - Sefton: There is a short perturbation to the ethnicity distributions in mid April, but with a very small sample size it is not significant. After this, the distributions are very consistent, suggesting very few cases were imported to the area through travel, and most of the transmission occurred in the community. There is a significant difference between the IMD score distributions. However, the three last deprived quintiles seem equally represented, with the only discrepancy occurring between the two most deprived quintiles, so in practice this difference may not be significant.
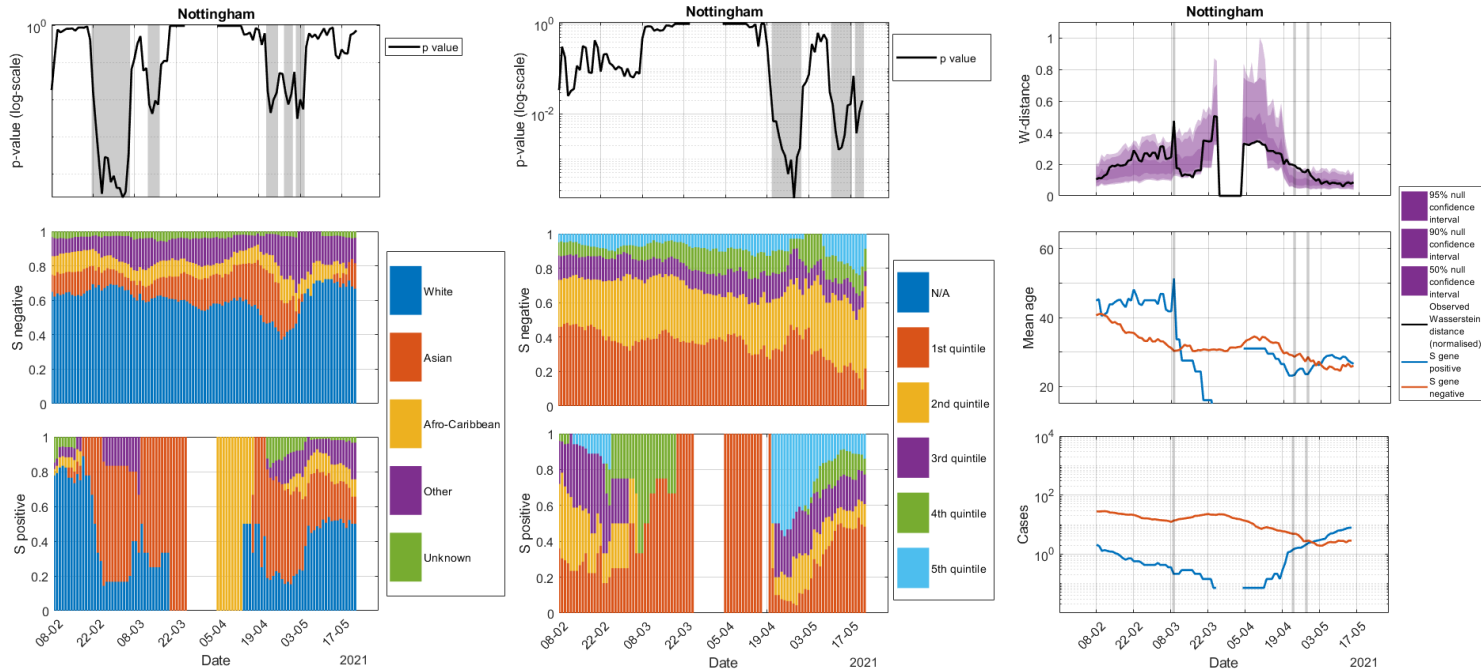
Figure 8 - Nottingham: In late April, gene positive cases appeared more skewed towards individuals with an Asian ethnicity, though the distributions have now converged. The IMD score of S gene negative cases is gradually becoming skewed to lower deprivation areas, whereas the S gene positive cases are becoming skewed to higher deprivation areas. The age distributions are consistent across most of the time series.
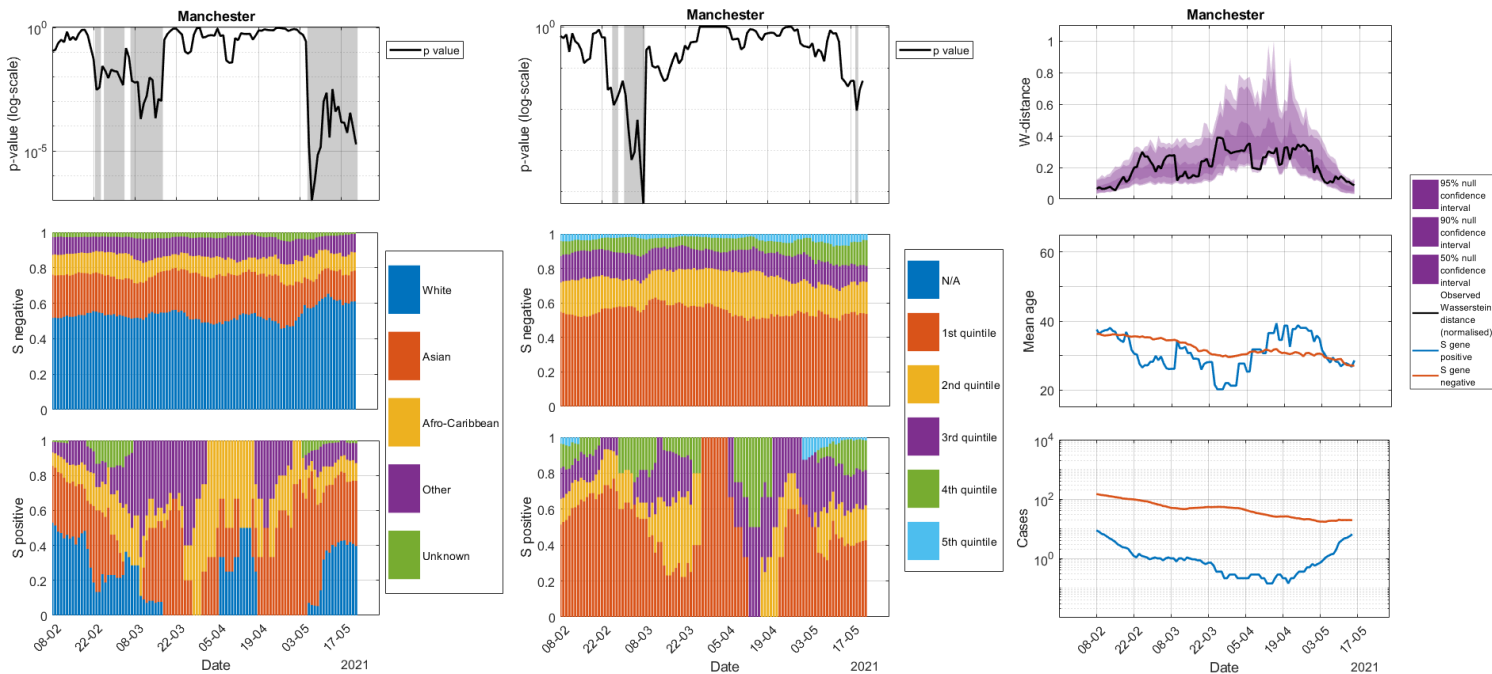

Figure 9 - Manchester: During February and March, the ethnicity distribution of S gene positive cases was skewed towards "Other". During April, the S gene positive cases became skewed towards individuals with an Asian ethnicity, and appears to be slowly returning to a similar distribution as S gene negative cases. IMD scores are relatively constant, except during late February when S gene positive cases became skewed towards lower deprivation areas.
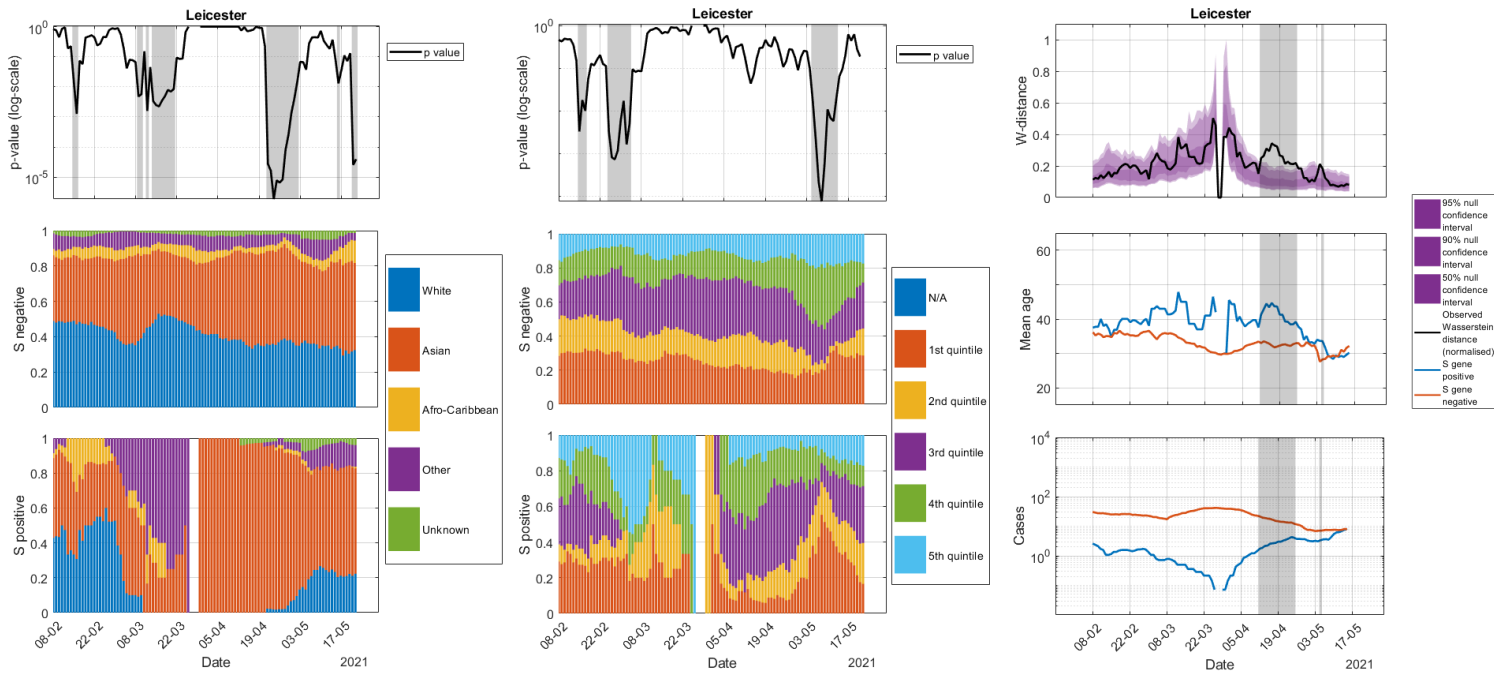
Figure 10 - Leicester: During late March and early April, the ethnicity distribution of S gene positive cases became strongly skewed towards individuals with an Asian ethnicity. This has since declined slightly, with the distribution slowly converging to the S gene negative distribution. IMD scores have been relatively constant, with the exception of early May, during which time the S gene positive cases saw a large rise in higher deprivation areas, since when the distributions have become aligned again. The age distribution of S gene positive cases saw a large perturbation in April, corresponding to the timing of multiple imported cases of B.1.617.2. Since this perturbation, the age distributions have become aligned.