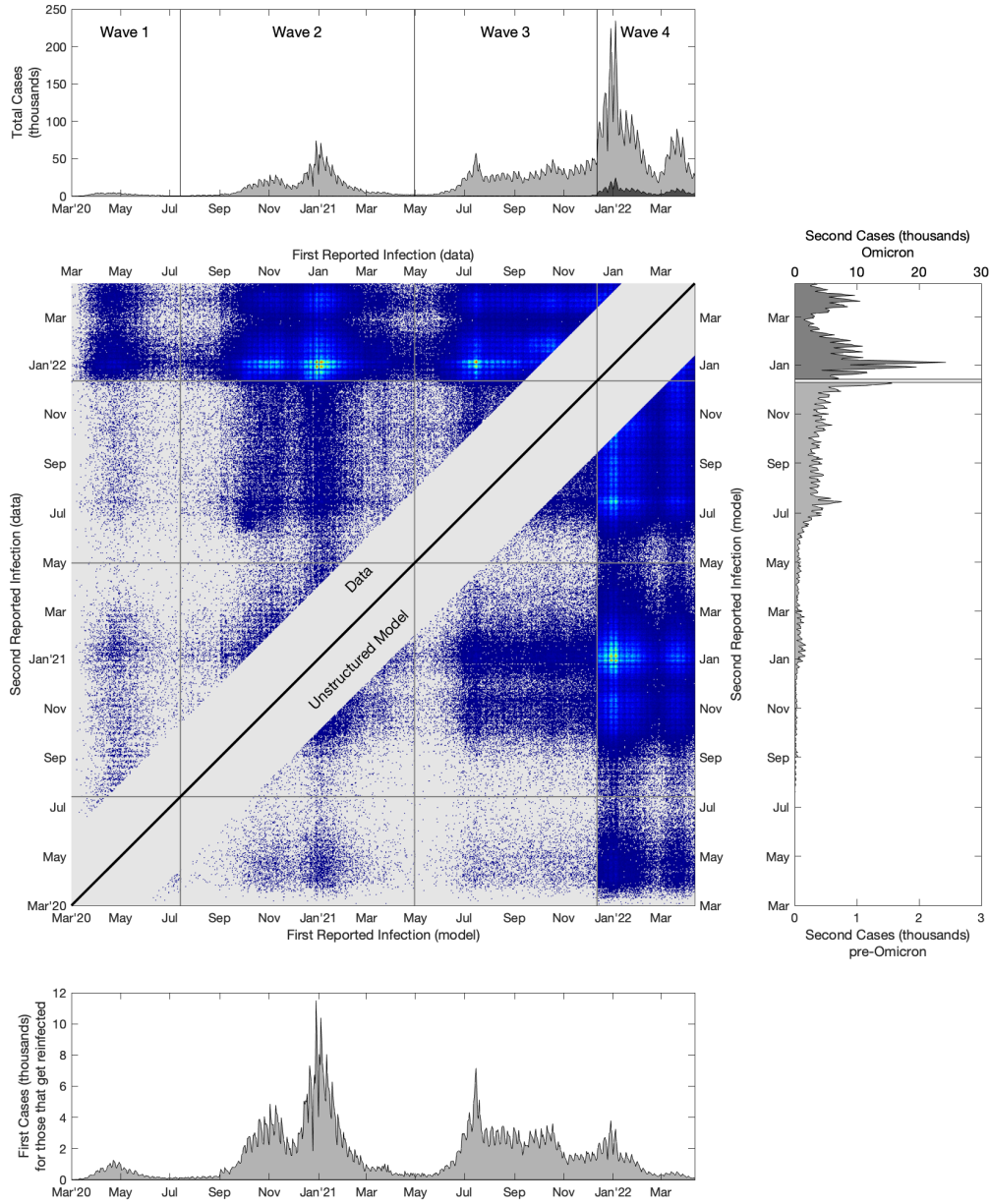## 5. Supplementary Figures

Figure S1. **A comparison of reported first and second infections in England to the model where *k* is wave specific.** The central panel shows reinfection data ($C_{t,T}$ upper-left triangle) and the corresponding model fit (in particular we show a Poisson sample of $M_{t,T}$ in the lower-right triangle, to allow a better visual comparison when $M_{t,T}$ is low), brighter colours correspond to more reported reinfections, while grey is zero reinfections in the data or model. The upper panel shows the total number of reported cases (pale grey) and the number of second reported cases (dark grey); the lower panel shows the number of reported first cases for those that report twice, while the right-hand panel shows the second reports - note that second reported cases during the Omicron wave (after 12th December 2021) are plotted on a different scale for clarity.
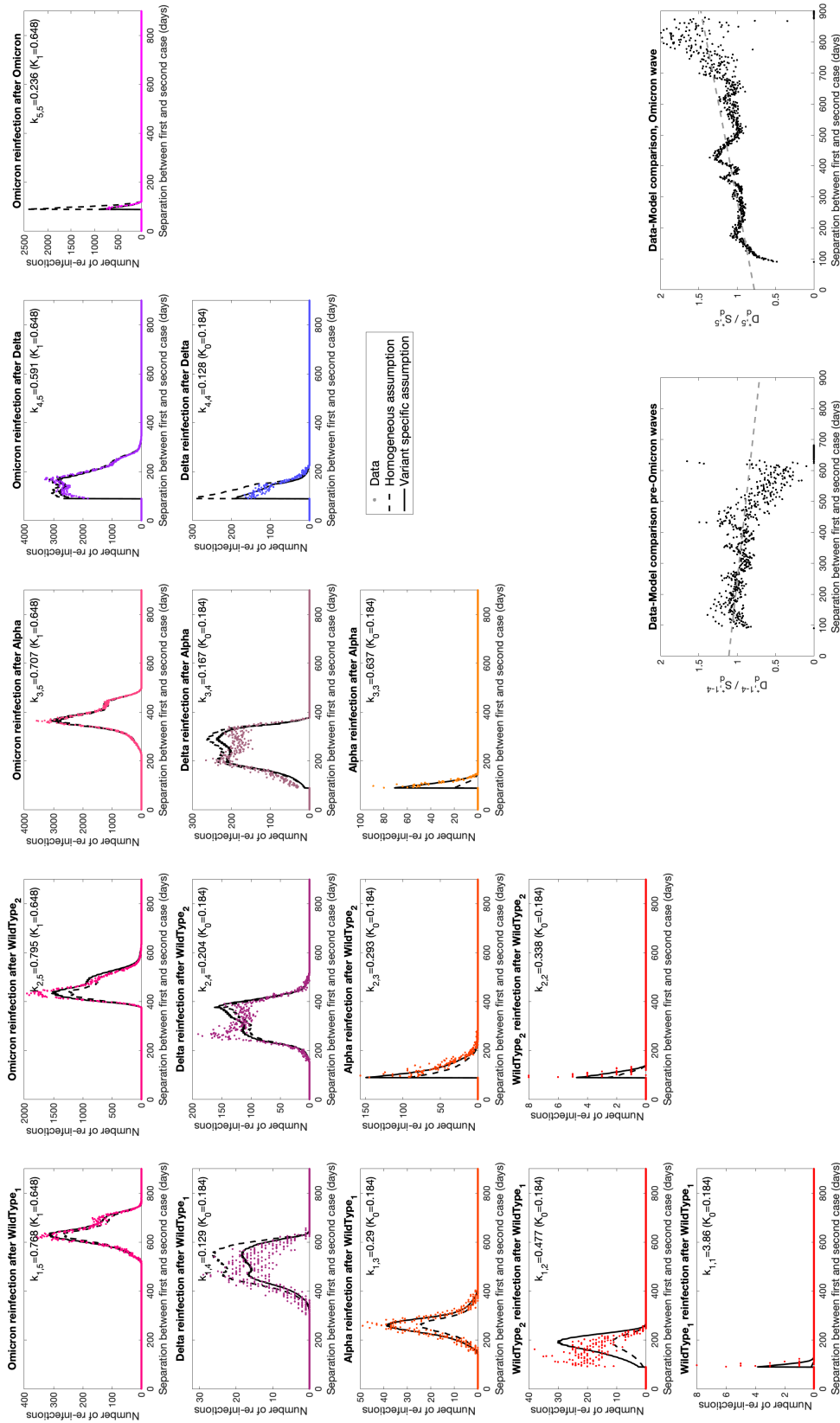
Figure S2. **Variant-variant interactions for reported reinfections, by splitting wave 2 (July 2020 to May 2021) into a Wildtype and Alpha variant component.** Each panel in the upper-left triangle shows the number of reported cases with a given delay between the first and second report ($D^{V,V}_d$, coloured dots), separated by the associated variants. For each, we also show the same numbers based on the simple model with both a variant-specific (solid lines) and homogeneous (dashed lines) assumptions for k. Unfortunately, determining the exact variant would require genotyping of all samples; while the ability to detect the S-gene is ready used as a marker of the switch between variants (with Wildtype, Delta and Omicron BA.2 being S-gene positive, and Alpha and Omicron BA.1 being S-gene negative) this only applies to individuals that are PCR tested through the TachPath system. The two panels in the bottom right perform a comparison of data and model (using the homogeneous assumption) when the second cases are from the pre-Omicron and Omicron waves respectively, enabling us to detect temporal trends - which are comparable to those in fig. 2.
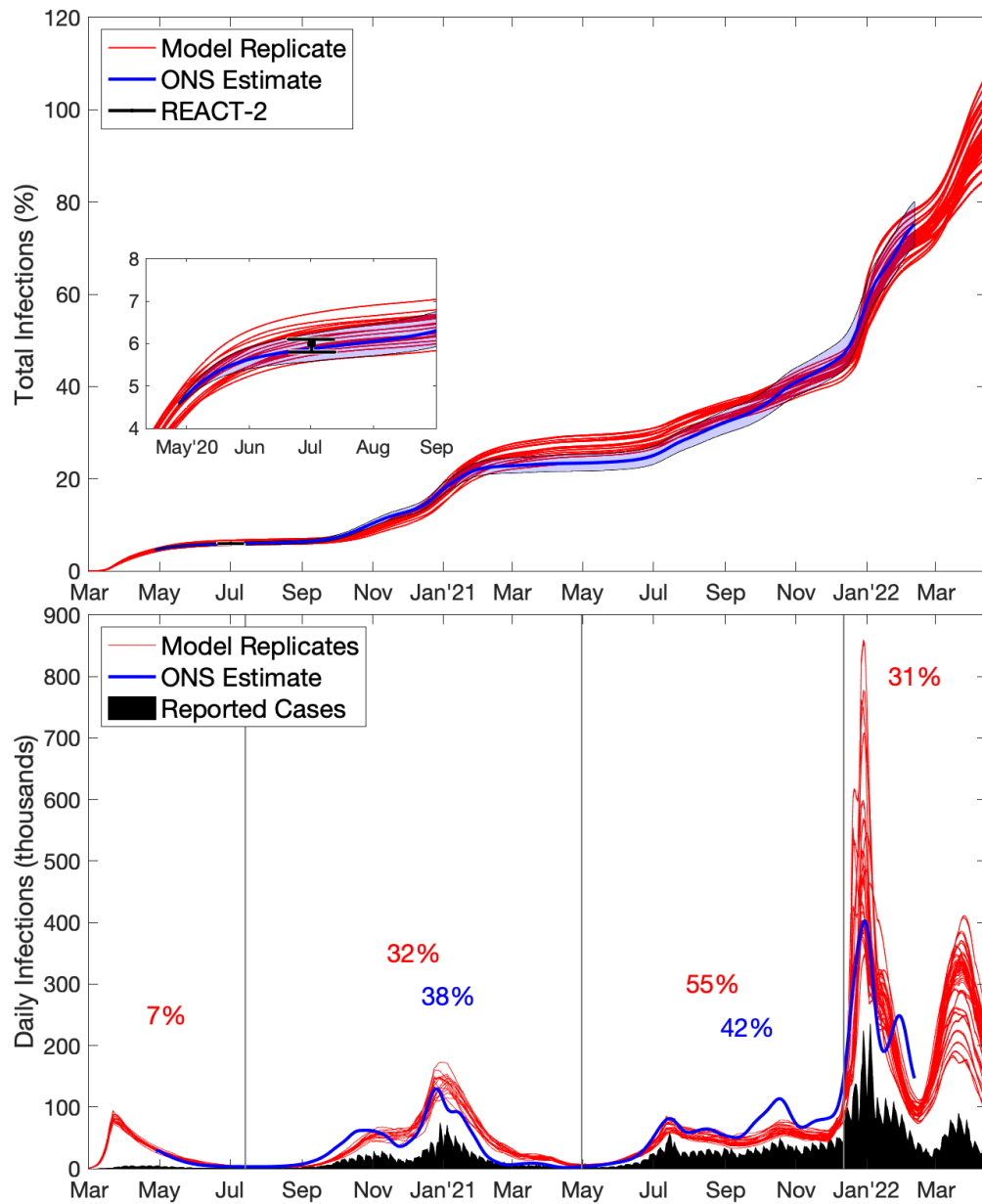
Figure S3. **Comparison between different methods of capturing the number of infections** Top: The total cumulative percentage of the population infected over time, showing three different results: (1) total number of infections from the Warwick model [20] which has been continually fitted to epidemiological data since the beginning of the epidemic and accounts for regional differences, age-structure and vaccination (red showing, 100 trajectories); the proportion of the population that has been infected from the ONS survey [14] (with the first point on 27th April 2020 increased to match model replicates, blue with 95% credible intervals); and the number of antibody positive individuals in June/July 2020 from the REACT-2 survey [5] (black point and error bars). Bottom: Estimated number of daily infections, from: reported daily cases including reinfection events (black) which forms a lower bound; model replicates including reinfection events (red) and an ONS estimate computed as the rate of change in the total number ever infected from the top graph (blue). The ONS estimate should be a reliable estimate of daily infections for the first three waves when reinfection is uncommon. For each wave (where possible) we display the fraction of infections (from either the Warwick model or the ONS estimate) that are reported as cases.