logistic res  ib0.pos\_prev ib0.hscw\_esr\_nims ib0.carehome ib1.nhsregion ib0.cev\_nims i.severelyimmunosuppressed ib1.imd5 ib15.eth1 ib15.age5 ib0.risk\_nims ib0.gender  ib40.wk\_onset  ib0.vacc\_catb

In stata the prefix ib#. means the variable is a factor and the baseline level is #

res:  0=control, 1=case

pos\_prev:  previously positive at least 90 before current sample date (0=no, 1=yes)

hscw\_esr\_nims:  health care worker flag (0=no 1 = yes)

carehome:   Geography NHS region (0=no 1=yes) carehome resident flag

nhsregion:  xxxx (coded as 1 to xx – however many there were)

cev\_nims:   Clinically extremely vulnerable flag (0=no, 1=yes)

severelyimmunosuppressed:   Severely immunosuppressed flag (0=no, 1=yes)

imd5: imd quintile (1 to 5 where 1 is most deprived)

eth1: ethnicity (baseline of 15 is white)

age5:  5 year age band (baseline of 15 is age 70-74)

risk\_nims: “at risk” population flag  (0=no, 1=yes)

gender:  female, male (0=female, 1=male)

wk\_onset: week of onset in 2021  (baseline week 40)

vacc\_catb: vaccine variable (0=unvaccinated and is used a baseline for comparison to unvaccinated, other levels cover 175+ days post dose2 and all the post dose 3 intervals).

Note: this model was fitted separately for each booster vaccine and when assessing the relative VE unvaccinated were dropped the baseline was changed the appropriate comparator level for the vacc\_catb factor. When assessing hospitalisations then this model is fitted to hospitalised cases and all controls.