\*\*import data\*\*

import delimited "C:\Users\mmn9\OneDrive - CDC\Case Control\Data management and analysis\analysis\_mpx-case-control\data\data download 18 jan 23.csv", varnames(1)

//n=852//

\*\*restrict to those who were eligible for survey\*\*

keep if inelig\_stop==1 //134 dropped, n=718)

\*\*convert dates to date format\*\*

gen first\_dose\_date2=date(first\_dose\_date, "YMD")

format first\_dose\_date2 %td

drop first\_dose\_date

rename first\_dose\_date2 first\_dose\_date

gen second\_dose\_date2=date(second\_dose\_date, "YMD")

format second\_dose\_date2 %td

drop second\_dose\_date

rename second\_dose\_date2 second\_dose\_date

gen symptoms\_date2=date(symptoms\_date, "YMD")

format symptoms\_date2 %td

drop symptoms\_date

rename symptoms\_date2 symptoms\_date

gen test\_result\_date2=date(test\_result\_date, "YMD")

format test\_result\_date2 %td

drop test\_result\_date

rename test\_result\_date2 test\_result\_date

gen control\_visit\_date2=date(control\_visit\_date, "YMD")

format control\_visit\_date2 %td

drop control\_visit\_date

rename control\_visit\_date2 control\_visit\_date

gen clinic\_date2=date(clinic\_date, "YMD")

format clinic\_date2 %td

drop clinic\_date

rename clinic\_date2 clinic\_date

\*\*case/control status - participant reported\*\*

gen CaCo\_reported=""

replace CaCo\_reported="Case" if provider\_dx== 1

replace CaCo\_reported="Control" if provider\_dx==2

tab CaCo\_reported

\*\*recode 2s to 0s for negative answers\*\*

foreach var of varlist first\_dose\_yesno second\_dose\_yesno symptoms\_yesno clinic\_yesno case\_yesno dose1\_yesno dose2\_yesno {

recode `var' (2=0)

}

\*\*restrict to thsoe with minimum self-reported data elements\*\*

gen complete=1

replace complete=0 if first\_dose\_yesno==.

replace complete=0 if first\_dose\_yesno==1 & first\_dose\_date==.

replace complete=0 if first\_dose\_yesno==1 & second\_dose\_yesno==.

replace complete=0 if second\_dose\_yesno==1 & second\_dose\_date==.

replace complete=0 if symptoms\_yesno==.

replace complete=0 if symptoms\_yesno==1 & symptoms\_date==.

replace complete=0 if provider\_dx==.

replace complete=0 if provider\_dx==2 & clinic\_yesno==.

tab complete //n=503 complete//

keep if complete==1 //n=503//

\*\*\* need to add IE date variables to complete status \*\*\*

gen complete\_updated=complete

replace complete\_updated=0 if CaCo\_reported=="Case" & (symptoms\_date==. & test\_result\_date==.)

replace complete\_updated=0 if CaCo\_reported=="Control" & clinic\_date==.

/// coding vax status at time of index event ///

\*\*define date of index event, which is date of symptom onset for cases and date of clinic visit for controls//

gen indexevent\_date=symptoms\_date if CaCo\_reported=="Case"

replace indexevent\_date=test\_result\_date if CaCo\_reported=="Case" & indexevent\_date==. //for cases that are asymptomatic, IE date is test result date//

replace indexevent\_date=clinic\_date if CaCo\_reported=="Control"

format indexevent\_date %td

//missing index event date for 79 participants//

\*\*time between index event date and 1st/2nd doses\*\*

gen firstdose\_days=indexevent\_date-first\_dose\_date if first\_dose\_yesno==1

gen seconddose\_days=indexevent\_date-second\_dose\_date if second\_dose\_yesno==1

\*\*vax status\*\*

gen vaxstatusatIE=""

replace vaxstatusatIE="Unvaccinated" if first\_dose\_yesno==0 //no doses//

replace vaxstatusatIE="Unvaccinated" if first\_dose\_yesno==1 & firstdose\_days<13 //got first dose, but <13 days from IE//

replace vaxstatusatIE="Partially vaccinated" if first\_dose\_yesno==1 & firstdose\_days>=13 & second\_dose\_yesno==0 //got first dose >=13 days from IE, did not get second dose//

replace vaxstatusatIE="Partially vaccinated" if second\_dose\_yesno==1 & seconddose\_days<13 //got second dose but <13 days from IE//

replace vaxstatusatIE="Fully vaccinated" if second\_dose\_yesno==1 & seconddose\_days>=13 //got second dose >=13 days from IE//

replace vaxstatusatIE="" if indexevent\_date==. //stata reads missings as an super high/infinite number so we need to make sure those without an index event date aren't inadvertently included in the vax status categorization

#==== Vaccination Status ----

# Time (days) between index event date and 1st/2nd doses

# Site

dose1index\_days\_site = index\_date - dose1\_date,

dose2index\_days\_site = index\_date - dose2\_date,

dose\_int\_site = dose2\_date - dose1\_date,

vax\_stat\_site\_na = as.character(NA), # make sure those without an index event date aren't included in the vax status categorization

vax\_stat\_site = fct\_relevel(as\_factor(case\_when(!is.na(index\_date) &

((dose1\_yesno == 0 & # no documented vaccination

is.na(dose2\_yesno)) |

(dose1\_yesno == 1 & # 1 dose but dose1-index date interval < 14 days

dose1index\_days\_site <= 0)) ~ "Unvaccinated",

(dose1\_yesno == 1 & dose2\_yesno != 1 & # 1 dose partial (1st dose but no 2nd dose)

dose1index\_days\_site >= 14) |

(dose1\_yesno == 1 & dose2\_yesno == 1 & # 1 dose partial (2 doses but dose interval <24 days)

dose1index\_days\_site >= 14 &

dose\_int\_site <24) |

(dose2\_yesno == 1 & # 2 doses partial (dose 2-index date interval < 14 days)

dose1index\_days\_site >= 14 &

dose\_int\_site >=24 & dose2index\_days\_site < 14) ~ "Partially vaccinated",

(dose2\_yesno == 1 &

dose\_int\_site >=24 & dose2index\_days\_site >= 14) ~ "Fully vaccinated",

TRUE ~ as.character(vax\_stat\_site\_na))), # make sure those w/o an index event date aren't included vax stat

"Fully vaccinated", "Partially vaccinated", "Unvaccinated"),