

S4 Table: Expected challenges/limitations, and steps that will be taken to minimise/mitigate effect.

Challenge/Limitation	What will be done to minimise/mitigate effect
<p>Loss-to-follow-up Some infections and symptoms may be missed where participants are not surveyed at a given study round</p>	<ul style="list-style-type: none"> • Inclusion of survey testing, primary care and Second Generation Surveillance System data to increase the chances of identifying all active infections • Inclusion of antibody data to identify infections not identified through testing for active infection
<p>Surveys may not adequately capture associated morbidity Symptom list is not exhaustive, and symptoms do not have an associated definition</p>	<ul style="list-style-type: none"> • Triangulation of long-term effect of SARS-CoV-2 infection from multiple sources • Survey list for both SIS and CoMMinS includes those symptoms considered most harmful in PPI work(1)
<p>Evolving nature of the pandemic SIS and CoMMinS are historic studies that took place when the wild-type and delta-type SARS-CoV-2 variants were circulating</p>	<ul style="list-style-type: none"> • It is importance to understand the past dynamics of SARS-CoV-2 • The health effects of earlier variants continue to be felt
<p>Misclassification bias Asymptomatic infections may be less likely to be identified than symptomatic infections (outside of testing done through the surveys) as a consequence of government testing strategy, potentially leading to either an under- or overestimation of outcome risk associated with SARS-CoV-2 infection</p>	<ul style="list-style-type: none"> • Inclusion of survey testing to increase the chances of identifying all active infections • Inclusion of antibody data to identify asymptomatic infections/infections not identified through testing for active infection
<p>Selection bias Those at risk of developing long-COVID (e.g., potentially those with symptomatic acute SARS-CoV-2 infection, or those with pre-existing health conditions) may be more likely to participate in the surveys</p>	<ul style="list-style-type: none"> • Adjustment of effect estimates for covariates • Subgroup analyses by participant demographic and clinical characteristics • Inclusion of antibody data to identify asymptomatic infections/infections not identified through testing for active infection
<p>Response bias Participants with SARS-CoV-2 infection may overreport symptoms. Survey participants are not presented with symptom definitions in either CoMMinS or SIS</p>	<ul style="list-style-type: none"> • Inclusion of antibody data to identify those with an unrecognised SARS-CoV-2 infection • Triangulation of long-term effect of SARS-CoV-2 infection from multiple sources
<p>Inadequate study power Outcome categories may be too broad and/or numbers too small to detect associations</p>	<ul style="list-style-type: none"> • Use of two different cohorts to investigate associations • Combining outcomes may conversely increase the power to detect an association
<p>Chance associations At (e.g.) the 5% level (i.e., using $p < 0.05$ as a cut-off which is a less than 5% chance that the null hypothesis of no association is false), 1 in 20 associations investigated will be statistically-significant purely on the basis of chance</p>	<ul style="list-style-type: none"> • ICD-10 and BNFC chapters selection informed by clinical consultation, documented symptomatic SARS-COV-2 in children, how well they are reported, and PPI work • Evidence across multiple outcome types will help determine plausibility of associations (e.g., comparing associations across relevant related survey symptoms, diagnoses and prescriptions for a given condition)
<p>Missing data Missing data on socio-demographic and clinical covariates may bias, or reduce the power to detect, effect estimates, if missing data are not random (i.e., more likely for certain covariate categories) and confound/modify the association between exposure and outcome</p>	<ul style="list-style-type: none"> • Use of both research data and EHR data to inform covariates • Exploration of patterns in missing data and use of multiple imputation where appropriate
<p>Incomplete outcome reporting in EHRs Not all outcomes recorded due to incomplete or variable recording by healthcare professionals, and individuals not presenting to healthcare services</p>	<ul style="list-style-type: none"> • Association of SARS-CoV-2 infection with multiple different types of outcomes investigated to increase likelihood of capturing effect of infection on long-term health outcomes
<p>Use of data on recent symptoms as a proxy for persistent symptoms A symptom being reported as experienced in multiple survey rounds does not necessarily mean the symptom was persistent the whole time. SIS and CoMMinS do not define ‘recent’ symptoms in the same way</p>	<ul style="list-style-type: none"> • The assumption that repeated symptoms correlate with persistent symptoms is not necessarily an issue if long-COVID symptoms are fluctuating • Persistent symptoms will be analysed in two ways for SIS, with use of data collected specifically on persistent symptoms as well as repeated data collection on recent symptoms
<p>Outcomes due to other causes The outcomes investigated (symptoms, school absences due to illness, diagnoses of conditions, medication use and health service attendance) have multiple other potential causes/explanations</p>	<ul style="list-style-type: none"> • Use of a control/unexposed group • We are utilising all available data, and comparing hazard ratios for time since SARS-CoV-2 diagnosis, with follow up without or before infection for each outcome • Adjustment of effect estimates for covariates

	<ul style="list-style-type: none">• Sensitivity analyses
Differential clinical course in different individuals The long-term effects of SARS-CoV-2 infection may be seen in different CYP in different ways	<ul style="list-style-type: none">• Use of broad chapters of diagnoses and prescriptions to more fully capture all possible long-term effects of SARS-CoV-2 infection• Clustering and latent class analyses planned