Deviations from protocol – Long-term kidney outcomes after COVID-19: a matched cohort study using the OpenSAFELY platform

Section	Original protocol	Deviation	Justification
4.3 Study designs	We will compare people after SARS-CoV-2 infection to both historical and contemporary cohorts in order to address epidemiological challenges to be able to triangulate meaningful conclusions.	The comparison with the historical cohort was designated as a sensitivity analysis.	We had expected the analysis with contemporary comparators may not be robust due to non-ascertainment of asymptomatic/mild COVID-19 infection. However, because of the absence of effect in those with non-hospitalised COVID-19, this did not prove to be an important bias. For clarity and avoidance of repetition, we assigned our historical cohort analysis as a sensitivity analysis in our final manuscript.
4.4.2 Historical comparator groups	[We will match historical comparators] 2 years before the first recording of SARS-CoV-2 infection in the exposed case.... In our secondary analysis, we will compare individuals who remain alive 28 days after admission to hospital with COVID-19 with those who remained alive 28 days after admission to hospital for pneumonia (including influenza) from 1 February 2018 up to two years before the end of follow-up for the exposed group (unmatched)	We matched historical comparators three years before first recording of COVID-19 infection. For secondary analysis, we followed-up patients hospitalised for pneumonia from 1 February 2017.	This allowed us to follow-up COVID-19 cases up to 31 January 2023 and historical comparators up to 31 January 2020 (i.e., before the start of COVID-19 epidemic in the UK).
4.4.2 Historical comparator groups & 4.4.3 Contemporary comparator groups	...comparator groups will comprise individuals from the general population under follow-up in OpenSAFELY, each matched five to one to each individual from the recorded SARS-CoV-2 group on age (within 1 year), sex, NHS Sustainability and	Individuals with COVID-19 were matched to three comparators rather than five. Individuals were matched on age, sex, and STP only (i.e., not additionally matched on IMD).	Due to depletion of potential matches, we reduced the number of matches to three. We adjusted for IMD (rather than including as an additional matching

	Transformation Partnership (STP, geographical areas configured for regional reorganisation available in our data), IMD decile (a measure of socioeconomic deprivation based on postcode), and date…		factor) to maximise the number of potential matches.
4.5.1 Exposure	We will also be able to investigate whether outcomes vary based on infection severity (hospitalisation, critical illness, mechanical ventilation and acute kidney injury [including requiring renal replacement therapy]) as well as calendar time and vaccination status…	We did not use mechanical ventilation or renal replacement therapy as additional levels of exposure.	We decided there would be little added benefit in disaggregating mechanical ventilation from critical care admissions, especially in the context of heightened competing risk of death. We did not include acute renal replacement therapy as an additional level of exposure as it was included in our outcome definition (we defined kidney failure as incident dialysis [including any acute renal replacement therapy records recorded between initial COVID-19 record and index date 28 days later on day 1 of follow up], kidney transplantation, or estimated glomerular filtration rate <15ml/min/1.73m^2). However, we investigated the effect of including and excluding acute renal replacement therapy (recorded in the 28 days after initial COVID-19 record) from our outcome definition through sensitivity analyses. Instead of calendar time and vaccination status being investigated as additional levels of exposure, we fitted these as interaction terms in secondary analysis to be able to determine stratum-specific

			effects on individuals with and without COVID-19.
4.5.3.3 Covariates (clinical)	We will describe several clinical conditions including: diabetes mellitus, hypertension, ischaemic heart disease, heart failure, atrial fibrillation, stroke, peripheral vascular disease, non-haematological cancer, haematological cancer, chronic respiratory disease, chronic liver disease, dementia, other chronic neurological diseases, systemic lupus erythematosus, vasculitis, rheumatoid arthritis, HIV and sickle cell disease. We will also describe albuminuria and/or proteinuria. In addition, we will describe coded CKD as clinical management may vary (e.g. greater routine follow-up) between people diagnosed and subsequently coded as having CKD compared to people without recognised CKD. …we will request baseline electronic frailty index (eFI) calculations for each individual aged ≥65 years… We will describe the number of individuals who were being prescribed drugs from the following classes in primary care: insulin, other antidiabetic drugs, immunosuppressants and anticoagulants.	In addition to diabetes mellitus, hypertension and non-haematological cancer, we grouped some covariates into composite categories: 1) cardiovascular disease included ischaemic heart disease, heart failure, atrial fibrillation, stroke and peripheral vascular disease; and 2) immunosuppressive diseases included haematological cancer, HIV, rheumatoid arthritis and systemic lupus erythematosus. We did not include the remaining clinical covariates (i.e., chronic respiratory disease, chronic liver disease, dementia, other chronic neurological diseases, vasculitis, and sickle cell disease). We did not extract albuminuria or proteinuria. We did not extract coded CKD. We did not extract eFI. We did not extract prescription data.	We grouped covariates to improve model performance. We did excluded some covariates due to relatively low numbers or risk of misclassification. We did not extract albuminuria or proteinuria due to underascertainment. We did not extract coded CKD as our models would be adjusting for baseline eGFR instead.[1,2] In the OpenSAFELY platform, eFI was only extracted in December 2020. eFI would therefore have been missing for individuals entering the cohort in the earlier months of the COVID-19 pandemic (i.e., March 2020 to December 2020) and at more likely to be misclassified for those entering the cohort later in the study period (as the eFI was not updated after December 2020, so would not have captured deficits occurring between then and a subsequent cohort entry). We did not extract prescription data as we were adjusting for the conditions they were treating. Immunosuppressants will often be unascertained in primary care records due to secondary care prescribing.
4.6.1 Main analyses	Due to the competing risk of death, we will use Fine and Gray regression models (using	We used Cox models stratified by matched set to obtain hazard ratios, overall, and	Cox models quantify average cause-specific hazards over the course of follow-

	calendar time as the underlying timescale) to estimate subdistribution hazard ratios… We will run sequential regression models: 1. Minimally adjusted for age (using spline functions) and sex, 2. Additionally adjusting for ethnicity, IMD, smoking and BMI, 3. Additionally adjusting for clinical comorbidities. We will assume people with missing BMI are not obese and people without a documented smoking status are non-smokers,	stratified by time period instead of Fine and Grey models. We ran fully-adjusted models only. We undertook complete case analysis without making assumptions about individuals with missing BMI or smoking status data.	up, as well as more meaningful period-specific hazards. These are more informative for policymaking than subdistribution hazards. We ran only fully-adjusted models to be able to present the most meaningful estimates only after adjustment for potential confounding variables. We performed complete case analysis to minimize effects of misclassification.
4.6.2 Secondary analyses	*A priori,* we intend to fit interaction terms to obtain hazard ratio estimates stratified by age group, ethnicity, quintile of IMD, frailty group (for those aged 65 years and over only), diabetes, baseline eGFR, coded CKD, and calendar time of infection. Stratified SHRs will also be reported for critical care admission, mechanical ventilation, acute RRT and AKI after COVID-19. We will describe Kaplan-Meier graphs for survival after COVID-19. We will obtain *p*-values using log rank tests.	We fitted interaction terms for age, sex, ethnicity, diabetes, baseline eGFR, vaccination status and COVID-19 wave (i.e. calendar time of infection). We obtained stratum-specific hazard ratios overall, and by hospitalization status only (i.e. hospitalised and non-hospitalised). We did not produce Kaplan-Meier graphs.	We felt sex was an important potential effect modifier. We felt interaction for IMD would not provide important additional information. We did not investigate interaction by frailty or coded CKD status: 1) We were unable to capture reliable time-updated measures of frailty; and 2) evidence indicates that coded CKD status does not capture CKD status as reliably as CKD stage established using serum creatinine test results to calculate eGFR measures.[2] Investigating interactions by critical care admission or acute kidney injury would have been difficult to interpret due to the high competing risk of death with our outcomes.

			We did not have access to the data after June 30 2024 to be able to obtain standardized incidence curves as planned. We undertook assessment for proportional hazards using Schoenfeld residuals and presented period-specific hazard ratios.
4.6.3 Sensitivity analyses		Our historical cohort analysis was presented as a sensitivity analysis.	For clarity and avoidance of repetition, we assigned our historical cohort analysis as a sensitivity analysis in our final manuscript.
4.6.3.2 Sensitivity analyses (eGFR reduction)	In case a difference in 50% reduction in eGFR is not identified, we will also investigate for 40% and 30% reduction in eGFR.	We did not investigate for 40% and 30% reduction in eGFR.	We identified an important effect on 50% reduction in eGFR and so these additional sensitivity analyses were not required.
4.6.3.3 Sensitivity analyses (incident eGFR <30ml/min/1.73m^2)	As healthcare services have been impacted heavily by the pandemic, there may not have been sufficient routine testing of patients to be able to detect a reduction in eGFR. As patients who are highest risk may be prioritised, we will also investigate for incident eGFR <30ml/min/1.73m^2...	We did not investigate for incident eGFR <30ml/min/1.73m^2.	We identified an important effect on 50% reduction in eGFR and so this additional sensitivity analysis was not required.
4.6.3.4 Sensitivity analyses (eGFR equation)	We will assess how many individuals change CKD stage from the original CKD-EPI equation after switching to the 2021 equation. If there are marked differences in the number of people with CKD and within each CKD stage, we will undertake a sensitivity analysis using the 2021 equation instead for eGFR.	We did not undertake an analysis using the 2021 CKD-EPI eGFR equation.	This equation has not yet been validated for use in clinical practice in the UK.

4.6.3.5 Sensitivity analyses (Age)	As our main analyses will adjust for age in broad age bands, to ensure that we have fully accounted for age as a potential confounder, we will repeat analyses with age as the underlying timescale (allowing us to finely account for age).	We did not undertake this sensitivity analysis.	Our main analysis implicitly adjusted for exact age as our models were stratified by matched set (with age as one of the matching variables).
4.6.3.7 Sensitivity analyses (Smoking)	In our main analyses we will assume that individuals with missing smoking status were non-smokers. To test this assumption, we will repeat our analyses restricting to those with known smoking status.	In our main analyses, we undertook complete case analysis without making assumptions about individuals with missing smoking status data.	We decided to perform complete case analysis (i.e. the planned sensitivity analysis) to minimize effects of misclassification.
4.6.3.8 Sensitivity analyses (BMI)	In our main analyses we will assume that individuals with missing BMI data have normal BMI. To test this assumption, we will repeat our analyses restricting to those with known BMI.	In our main analyses, we undertook complete case analysis without making assumptions about individuals with missing BMI status data.	We decided to perform complete case analysis (i.e. the planned sensitivity analysis) to minimize effects of misclassification.