Appendix 1: STUDY PROTOCOL

Version history

V1	10/01/2024	RT drafted skeleton protocol; funding application submitted.
V2	11/04/2024	RT updated the protocol and data available for consortium registration of study.
V3	30/05/2024	RT circulates protocol to Investigator team and updated the review suggestions
V4	04/06/2024	RT uploaded protocol to Git repository

Study Title: **Beyond pandemic: Tracing the longtime cardiovascular risks and related deaths in multiethnic young adults after SARS-COV-2 infection and vaccination.**

Lay Summary

In the aftermath of COVID-19 pandemic there has been a disturbing rise in deaths of young adults due to heart and circulation issues. Emerging research had already indicated a link between COVID-19 infection and development of heart and circulatory problems that leads to cardiovascular disease (CVD) and deaths in people. The risks for CVD and deaths after COVID-19 infection were evidenced in both immediate and medium term follow up periods up to 1 year among the overall population in England and Wales. However, it's impacts beyond a year is yet to be investigated especially to understand how long the damages made in the heart and circulation system of a person after COVID-19 infection sustain and combinedly interacts with other new infections in future.

Research Question

In people who have had SARS-COV-2 infection compared with people who have not, are there higher rates (expressed as hazard ratios with time since COVID-19 disease) of outcome events (1. fatal CVD, 2.CVD risk events) before and after adjustment for potential confounders? What is the interaction effect of overlapping infections on these outcome events before and after adjustment for potential confounders?

Study Population

Young patients alive and aged 18 to 59 years (on 1st January 2020) who have a record in the primary care extract with a known sex and deprivation index.

Data Source

NHS Digital TRE for England (up to Dec2024, and Dec 2025 release)

- Primary care data (GP Data for Pandemic Planning and Research via General Practice Extraction Service, GPES) Second Generation Surveillance System (SGSS) COVID-19 infection laboratory testing data

- Hospital episode statistics Admitted Patient Care (HES APC)
- Office for National Statistics (ONS) death registration records
- Community dispensing data

NHS Wales (via SAIL Databank TRE)

- COVID C20 (all people alive and resident in Wales from 01/01/20) and C16 (counterfactual from 01/01/16 to end 2019) total population cohorts 3.2M. Censored by migration out of Wales and death.
- Patient Episode Database for Wales (PEDW).
- Consolidated mortality with 4 separate mortality data sources (Welsh Demographic Service Dataset (WDSD) – population spine weekly flows, ONS - monthly and daily flows (ADDE and ADDD), and records from the MPI (Master Patient Index) - daily flows (CDDS)).
- Pathology data COVID-19 test results (PATD) daily flows from NHS and Public Health Wales (PHW) laboratories and Lighthouse laboratories, including antigen and antibody testing - Pillar 1,2 and 3.
- Primary care (available on 80% that provide data to SAIL on a monthly basis containing all diagnostic, referral, prescribed medication (WLGP). Plus 100% GP daily flow since January 2020 for of COVID-19 specific coding and symptom respiratory codes (GPCD)).
- Community dispensing (WDDS).
- Intensive Care National Audit & Research Centre (ICCD, ICNC).
- Critical Care Data Set (CCDS).
- The ONS Census (CENW).
- Outpatient Database for Wales (OPDW) and Emergency department data set (EDDS)

Cohorts and Approach:

TITLE	**Cohort 1**	**Cohort 2**
1.COHORTS	Non vaccinated COVID-19 infections in young adults (18 – 59 years) in 2020 prior to vaccines becoming available.	Vaccinated COVID-19 infections in young adults (18 – 59 years) in 2021 after vaccines become available
1.1 Start date	01/01/2020, which is the approximate start date of the pandemic in the UK.	8th December 2020 (date of start of vaccine rollout)
1.2 End date	7th December 2020 (start of vaccine rollout on 8th Dec 2020)	31/12/2021, to facilitate a 4 year follow up of this cohort within the project implementation period, for the latest data release in year 2025. Note that cohort2 is going to be predominantly of Delta wave infections as Omicron wave of infections took over in Dec 2021(UK Health Security Agency, this report).
1.3 Inclusion criteria	On study start date: -Alive -Known age of ≥18 and <60	On study start date: -Alive -Known age of ≥18 and <60

	-Known sex -Known deprivation -Have a record in the primary care -Not known to be living outside of England /Wales.	-Known sex -Known deprivation -Have a record in the primary care -Not known to be living outside of England /Wales.
1.4 Exclusion criteria	• SARS-CoV-2 infection recorded prior to their index date. • any patients/records where a linkage is not possible.	• SARS-CoV-2 infection recorded prior to their index date [Note: these individuals are required for a sensitivity analysis and so would not be removed at the data extraction stage] • They have a record of one or more vaccination doses prior to their index date. • They received a vaccination prior to 08-12-2020 (i.e., the start of the vaccination program/study) • They received a second dose vaccination before their first dose vaccination. • those with a second vaccine dose but no first vaccine dose • They received a second dose vaccination less than three weeks after their first dose. • those with booster dose before first or second vaccine dose • those with a booster dose but no first and second vaccine dose • those with an interval between second vaccine dose and booster dose of < 3 months • those with a third vaccine dose before first or second vaccine dose • those with a third vaccine dose but no first and second vaccine dose • They received mixed vaccine products before 07-05-2021
1.5 Follow-up start	Follow-up will start at the latest of the following dates (i.e., an individual's index date): - Study start date.	- Study start date - Two weeks after their second vaccination (i.e. on fully immunised state)
1.6 Exposure	SARS-COV-2 infection	SARS-COV-2 infection
1.7 Exposure definition	Date of positive SARS-COV-2 PCR antigen test in SGSS data; and/or Date of confirmed diagnosis code in Primary care data; and/or	Date of positive SARS-COV-2 PCR antigen test in SGSS data; and/or Date of confirmed diagnosis code in Primary care data; and/or

	Start date of episode with confirmed diagnosis in any position in SUS data SUS; and/or Date of death with SARS-COV-2 infection listed as primary or underlying cause in ONS death registry	Start date of episode with confirmed diagnosis in any position in SUS data SUS; and/or Date of death with SARS-COV-2 infection listed as primary or underlying cause in ONS death registry
1.8 Follow-up end	Follow-up will end at the earliest of the following dates: • Death • Outcome event • Vaccination • Study end date (07/12/2024), Allowing for minimum 4 year follow up of the cohort.	Follow-up will end at the earliest of the following dates: • Death • Outcome event • Study end date (31/12/2025), Allowing for minimum 4 year follow up of the cohort.
2. **Hypothesis**	There is a higher risk of 'outcome' events after SARS-COV-2 infection than before or without infection.	
3. Outcomes	WP1: Fatal CVD events WP2: CVD risk outcomes such as, 1. New hypertension 2. New hypercholesterolemia 3. New Diabetes 4. New Obesity WP3: WP1 & 2 Outcomes (to investigate the role of overlapping infections)	
4.Outcome definitions	WP1: CVD events will be defined as fatal if they are followed by death of any cause within 28-days or are only recorded as fatal (i.e. reported only in death records). WP2: Each CVD risk outcome is defined as the first event of that type following the start of follow up in one of the following data sources: HES, primary care or ONS death registry. Date of onset defined as: date of start of SUS or hospital admission spell with event; OR date of General Practitioner (GP) consultation with event; OR death with event (whichever comes first). WP3: Over lapping infections will be defined as any 1. Bacterial, 2. Viral, 3. Fungal and/or mix of these categories of infection entered to EHR after their first SARS-COV-2 infection.	
5.CVD events	Individual CVD event definitions to define the fata outcomes: 1.**Arterial events** ▪ Incident myocardial infarction in primary care or HES or ONS death; ▪ Incident ischaemic stroke or stroke of unknown type or retinal infarction in primary care or SUS or ONS death; ▪ Incident non-stroke non-MI arterial embolism in primary care or SUS or ONS death. 2.**Venous events** ▪ Pulmonary embolism in HES or ONS death; ▪ Deep vein thrombosis in HES or ONS death;	

	▪ Portal vein thrombosis in HES or ONS death; ▪ Other deep vein thrombosis in HES or ONS death; ▪ Thrombosis during pregnancy (excluding intracranial venous thrombosis) in SUS or ONS death; ▪ Intracranial venous thrombosis during pregnancy in HES or ONS death; ▪ Intracranial venous thrombosis22 in HES or ONS death. 3.**Other vascular events** ▪ Intracerebral haemorrhage in primary care or HES or ONS death; ▪ Mesenteric thrombus in HES or ONS death; ▪ Spinal stroke in HES or ONS death; ▪ Heart failure in HES or ONS death; ▪ Angina in HES or ONS death. ▪ Transient ischaemic attack in HES or ONS death 4. **Composite events** *Arterial*: fata events out of first of ischaemic stroke or stroke of unknown type or retinal infarction or myocardial infarction or other arterial thrombosis. *Venous*: fatal events out of first of pulmonary embolism or lower limb deep venous thrombosis or intracranial venous thrombosis or intracranial venous thrombosis during pregnancy or portal vein thrombus or other deep vein thrombosis.
6.Potential confounders	Defined up to the start of study, with a look back in GP and hospital admission (HA) for each person: o Sex: categorical, GP; o Age in years continuous, at inception; GP; o Ethnicity 5 categorical initially, but to 19 categorical granularities when sufficient event size observed, most recent recorded prior to inception; GP and if missing from GP data then from HA; o Area-level Deprivation continuous, most recent recorded prior to inception; HA or GP; o Region: East of England, London, Midlands, NE and Yorkshire, North West, South East, South West, Scotland, Wales, most recent residence prior to inception; GP; o Consultation rate: number of primary care contacts in the year prior to inception; GP; o Medications: total number of medications by BNF chapters prescribed within three months prior to the inception date; GP; o Diabetes: yes/not recorded from start of record to inception; GP or HA or medication; o Depression: yes/not recorded from start of record to inception; GP or HA; o Obesity: yes/not recorded from start of record to inception or BMI>30; GP o Cancer yes/not recorded from start of record to inception; GP or HA; o COPD yes/not recorded from start of record to inception; GP or HA: o CKD yes/not recorded from start of record to inception; GP or HA; o Liver disease: yes/not recorded from start of record to inception; GP or HA; o Major Surgery yes/not recorded from start of record to inception; GP or HA; o Hypertension yes/not recorded from start of record to inception; GP or

	HA or medication; ○ Dementia yes/not recorded from start of record to date of linkage; GP or HA; ○ Smoking current/ex-/never/unknown most recent prior to inception; GP; ○ Antiplatelet yes/no/unknown (unknown if patient's GP practice not in prescription extract) BNF codes starting '0209' with at least one prescription within three months prior to the inception date; dispensing; ○ BP lowering yes/no/unknown (unknown if patient's GP practice not in prescription extract) using DMD list of antihypertensives ○ Lipid lowering yes/no/unknown (unknown if patient's GP practice not in prescription extract) BNF codes starting '0212' at least one prescription within three months prior to the inception date; dispensing; ○ Anticoagulant yes/no/unknown (unknown if patient's GP practice not in prescription extract) BNF codes starting '020802' (exclude: '0208020I','0208020W') with at least one prescription within three months prior to the inception date; dispensing; ○ COCP yes/no/unknown (unknown if patient's GP practice not in prescription extract) BNF codes starting '070301' with at least one prescription within three months prior to the inception date; dispensing; ○ HRT yes/no/unknown (unknown if patient's GP practice not in prescription extract) BNF codes starting '0604011' with at least one prescription within three months prior to the inception date; dispensing; ○ History of stroke yes/not recorded any stroke type (from 'outcomes above, ischaemic stroke, intracranial haemorrhage, or stroke of unknown type) recorded from start of record to inception; GP or HA; ○ History of MI: yes/not recorded any MI type (from 'outcomes above) recorded from start of record to inception; GP or HA; ○ History of angina: yes/not recorded any angina type (from 'outcomes above) recorded from start of record to inception; GP or HA; ○ History of VT: yes/not recorded and venous event (from 'outcomes above) recorded from start of record to inception; GP or HA; ○ History of thrombophilia: yes/not recorded from start of record to inception, defined as ICD-10 D68.5 D68.6; acquired or inherited; GP or HA; ○ History of thrombocytopenia (from outcomes above) yes/not recorded from start of record to inception; GP or HA; ○ History of any outcome type: (from outcomes above) yes/not recorded from start of record to inception; GP or HA;
7.**Missing data**	All analyses will be "complete case" analyses (note that many potential confounders are defined using the presence versus absence of specific codes in the EHRs, so by definition have no missing values). We will not use multiple imputation or indicators for missing data.
8. Meta-analysis	We will use inverse-variance weighted meta-analysis to pool estimates of post-COVID hazard ratios from common models across nations, examining heterogeneity using I2 statistics and reporting summary estimates and confidence intervals.

References

Knight, R., Walker, V., Ip, S., Cooper, J.A., Bolton, T., Keene, S., Denholm, R., Akbari, A., Abbasizanjani, H., Torabi, F. and Omigie, E., 2022. Association of COVID-19 with major arterial and venous thrombotic diseases: a population-wide cohort study of 48 million adults in England and Wales. Circulation, 146(12), pp.892-906.

Cezard, G.I., Denholm, R.E., Knight, R., Wei, Y., Teece, L., Toms, R., Forbes, H.J., Walker, A.J., Fisher, L., Massey, J. and Hopcroft, L.E., 2024. Impact of vaccination on the association of COVID-19 with cardiovascular diseases: An OpenSAFELY cohort study. *Nature Communications*, *15*(1), p.2173.

Ip, S., North, T.L., Torabi, F., Li, Y., Abbasizanjani, H., Akbari, A., Horne, E., Denholm, R., Keene, S., Denaxas, S. and Banerjee, A., 2024. Cohort study of cardiovascular safety of different COVID-19 vaccination doses among 46 million English adults. *medRxiv*, pp.2024-02.