**Protocol: Predicting and validating risk of pre-pandemic and excess mortality in individuals with chronic kidney disease using national Electronic Health Records**

**Lay summary**

Chronic kidney disease (CKD) is a common comorbidity associated with increased risk of severe coronavirus infection and poor clinical outcomes. The pandemic has had both direct (through infection) and indirect impact (through changes in health services, economic upheaval and behavioural factors [1], [2]). The direct impact on individuals with CKD and other underlying conditions is related to baseline risk, influenced by age, multimorbidity and other socio-demographic factors. However, previous studies of COVID-19 in CKD have been small scale (12-1099 cases), mostly focused on end-stage CKD, and ignored major comorbidities. Thus, using large-scale, population-based electronic health records, in people with incident CKD we aimed to (a) identify the most common comorbidities; (b) estimate 1-year (pre-pandemic) risk of mortality and (c) predict excess deaths related to COVID-19 and validate our findings over 1-year of pandemic based on pre-pandemic risk of mortality at different population infection rates and relative risks.

|  |
| --- |
| **Background** *Approx. 300 words2 summarising why the question(s) you are addressing matter, and how your project fits within the broad scope of* [*CVD-COVID-UK*](https://www.hdruk.ac.uk/projects/cvd-covid-uk-project/)) |

Chronic kidney disease (CKD) carries major global disease burden, both as a risk factor for morbidity and mortality, and as the end syndrome of underlying risk factors and diseases [3],[4], such as cancers[5] and CVD [6]. During the coronavirus (COVID-19) pandemic, CKD has been associated with increased risk of poor outcomes [7], [8]. Despite its clinical and public health importance, CKD research to-date in all stages, multi-morbidity, or the general population [9] has been limited.

On the other hand, few risk stratification tools are used in clinical practice for individuals with CKD or prediction of CKD, and those that include CKD, usually do not consider different CKD stages. Better characterization of baseline risk in people with CKD may inform individual and population approaches to CKD prevention and treatment and integrated management of chronic diseases. This proposal fits well into the scope of CVD-COVID-UK, specifically in terms of (i) understanding the prognostic role of COVID-19 on future CKD risk groups, and (ii) providing a way toward personalised healthcare through stratified risk estimation by key clinical and demographic confounders.

**Research Methods**

**Study design and Data sources**

We conduct a retrospective, population-based cohort study using primary and secondary care data linked to Office for National Statistics (ONS) death registration [10]. The base cohort consists of 3,862,012 adults aged 30 years or older and registered with a general practice between Jan 1, 1997, and Jan 1, 2017, with ≥1 year of follow-up data, having demographic (age, gender, ethnicity, deprivation decile) and baseline characteristics (comorbid conditions, risk factors, and medications) recorded over time.

We use NHSD TRE data (recorded events across GDPPR and HES) [11] including all adults >= 30 years by 1 March 2020 (n=54 million); the diagnoses in HES and primary care linked to ONS data since 1997 will be investigated to identify CKD patients. The data will be then linked to COVID-19 data ranging from 1 March 2020 to 1 March 2021. The inclusion criteria of a COVID-19 event will be defined as any of a positive test (polymerase chain reaction or lateral flow), a coded diagnosis in primary or secondary care or a COVID-19 diagnosis on a death certificate.

**NHS England (TRE)**

***HES: Hospital Episode Statistics***

* Admitted Patient Care
* Adult Critical Care
* Outpatients
* Accident & Emergency/ Emergency Care Data Set

***ONS:*** Deaths

***GDPPR:*** GPES Data for Pandemic Planning and Research

***COVID-19 SGSS:*** Second Generation Surveillance System (+ve results from pillars 1 and 2)

**CKD Phenotype**

We will include all individuals with incident and prevalent CKD, using International Classification of Diseases 10th edition (ICD10) and Read codes, and estimated glomerular filtration rate (eGFR) levels. Based on clinical guidelines [12], patients were mutually exclusively classified into either CKD Stages 3 (eGFR 30-60) , 4 (eGFR 15-29), 5 (eGFR<15 mL/min/1.73 m2), Stages 1 or 2 [10], using a validated CALIBER phenotype. We defined incident cases of stages of CKD by first incident of a highest stage since individuals could have multiple records and disease progression could occur over time.

**COVID phenotype**

*PRIMARY*

• A positive PCR test and date of test available in testing data.

*SECONDARY*

• Everyone with a positive PCR test or hospitalisation, or primary care record of COVID taking first positive test or first hospital or primary care record coded as COVID from any source

**Underlying Conditions**

Definitions of comorbid conditions and risk factors will be derived from Health Data Research (HDR) UK–CALIBER [13] and the latest validated COPD phenotype [14]. Phenotypes were defined using Read and ICD-10 codes from hospital and primary care information recorded in primary care [10]. IN NHSD TRE, we will employ mapped SNOWMED Concept IDs from Caliber library and we will extract relevant concepts to our CKD code list using code browser.

**Outcomes**

1. Prevalence of underlying conditions in individuals with prevalent and incident (chronic kidney disease using CPRD data.
2. Prevalence of underlying conditions in individuals with incident chronic kidney disease (CKD), by sex, ethnicity, age category and CKD stage using CPRD data
3. One-year all-cause mortality (%) in individuals with incident and prevalent chronic kidney disease (n=294381) by number of underlying conditions, age, sex, CKD stage and deprivation using CPRD data.
4. One-year (pre-pandemic) risk of mortality for moderate-risk conditions in individuals with chronic kidney disease using CPRD data.
5. Estimated one-year excess mortality by population infection rate and relative impact of the pandemic using Lancet 2020 model.
6. Absolute number and age distribution of excess deaths over 1 year of the pandemic in individuals with chronic kidney disease: (a) predicted using Lancet 2020 model1; (b) observed2; and (c) predicted using Lancet 2020 model, assuming observed infection rate and relative risk3.

**Statistical Analysis**

We will identify the most prevalent comorbidities in incidence and prevalence cohorts using CPRD GOLD (linked to HES/ONS) and estimate the impact of each condition (as well as co-existing pair of conditions at baseline date) on mortality using odds ratio (Wald method), relative risk, and proportional hazard models. We utilise Kaplan-Meier 1-year all-cause mortality estimates by age bands, sex, ethnicity, deprivation (most deprived vs least deprived), CKD stage, and underlying conditions listed by the NHS as moderate- or high-risk [15].

We estimate excess mortality using recently published [16], [17] model, incorporating relative impact of mortality (compared with pre-pandemic mortality): relative risk (RR) 2·0 and 3.0 [18] under different population infection rates (IR) of 10%, 40% and 80%. Our validation will involve analysis of the actual observed COVID-19 mortality in individuals with CKD in the NHSD TRE, and calculation of actual IR, RR and age distribution. We use R, version 3.4 in the University College London ISO27001 Safe Haven for all analyses, except for external validation in NHSD TRE.

**References**

[1] H. Yao, J.-H. Chen, and Y.-F. Xu, “Patients with mental health disorders in the COVID-19 epidemic,” *The Lancet Psychiatry*, vol. 7, no. 4, p. e21, 2020.

[2] V. Pereira-Sanchez *et al.*, “COVID-19 effect on mental health: patients and workforce,” *The Lancet Psychiatry*, vol. 7, no. 6, pp. e29–e30, 2020.

[3] P. Cockwell and L.-A. Fisher, “The global burden of chronic kidney disease,” *Lancet*, vol. 395, no. 10225, pp. 662–664, 2020.

[4] B. Bikbov *et al.*, “Global, regional, and national burden of chronic kidney disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017,” *Lancet*, vol. 395, no. 10225, pp. 709–733, 2020.

[5] B. Stengel, “Chronic kidney disease and cancer: a troubling connection,” *J. Nephrol.*, vol. 23, no. 3, p. 253, 2010.

[6] R. T. Gansevoort *et al.*, “Chronic kidney disease and cardiovascular risk: epidemiology, mechanisms, and prevention,” *Lancet*, vol. 382, no. 9889, pp. 339–352, 2013.

[7] B. M. Henry and G. Lippi, “Chronic kidney disease is associated with severe coronavirus disease 2019 (COVID-19) infection,” *Int. Urol. Nephrol.*, pp. 1–2, 2020.

[8] A. Extance, “Covid-19 and long term conditions: what if you have cancer, diabetes, or chronic kidney disease?,” *BMJ*, vol. 368, 2020.

[9] R. Saran *et al.*, “US renal data system 2017 annual data report: epidemiology of kidney disease in the United States,” *Am. J. kidney Dis. Off. J. Natl. Kidney Found.*, vol. 71, no. 3 Suppl 1, p. A7, 2018.

[10] S. Denaxas *et al.*, “UK phenomics platform for developing and validating electronic health record phenotypes: CALIBER,” *J. Am. Med. Informatics Assoc.*, vol. 26, no. 12, pp. 1545–1559, 2019.

[11] “Trusted Research Environment service for England,” *May 2021*. https://digital.nhs.uk/coronavirus/coronavirus-data-services-updates/trusted-research-environment-service-for-england.

[12] A. Forbes and H. Gallagher, “Chronic kidney disease in adults: assessment and management,” *Clin. Med. (Northfield. Il).*, vol. 20, no. 2, p. 128, 2020.

[13] V. Kuan *et al.*, “A chronological map of 308 physical and mental health conditions from 4 million individuals in the English National Health Service,” *Lancet Digit. Heal.*, vol. 1, no. 2, pp. e63–e77, 2019.

[14] J. K. Quint *et al.*, “Validation of chronic obstructive pulmonary disease recording in the Clinical Practice Research Datalink (CPRD-GOLD),” *BMJ Open*, vol. 4, no. 7, p. e005540, 2014.

[15] “Methodology, COVID-19 – high risk shielded patient list identification,” *NHS Digit. Engl.*, p. version 9, 2020.

[16] A. Banerjee *et al.*, “Estimating excess 1-year mortality associated with the COVID-19 pandemic according to underlying conditions and age: a population-based cohort study,” *Lancet*, 2020.

[17] A. G. Lai *et al.*, “Estimating excess mortality in people with cancer and multimorbidity in the COVID-19 emergency,” *medRxiv*, 2020.

[18] P. Spychalski, A. Błażyńska-Spychalska, and J. Kobiela, “Estimating case fatality rates of COVID-19,” *Lancet. Infect. Dis.*, 2020.