Predicting and validating the impact of the COVID-19 pandemic in individuals with chronic kidney disease

Muhammad Dashtban, PhD[1], Mehrdad A Mizani, PhD[1], Prof Spiros Denaxas PhD[1], Prof Dorothea Nitsch, MD[2], Prof Jennifer Quint[3], Richard Corbett, MD[4], Jil B Mamza, PhD[5], Tamsin Morris, BSc[5], Prof Mamas Mamas[6], Prof Deborah A Lawlor, PhD[7], Prof Kamlesh Khunti, FMedSci[8], CVD-COVID-UK Consortium[9], Prof Cathie Sudlow, OBE FRSE[9], Prof Harry Hemingway, FMedSci[1,10], Prof Amitava Banerjee, DPhil[1,10,11,12]

[1]Institute of Health Informatics, University College London, London, UK
[2]Department of Non-Communicable Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK
[3] Faculty of Medicine, National Heart & Lung Institute, Imperial College London, London, UK
[4]Imperial College Healthcare NHS Trust, London, UK
[5]Medical and Scientific Affairs, BioPharmaceuticals Medical, AstraZeneca, Cambridge, UK
[6]Keele Cardiovascular Research Group, Centre for Prognosis Research, Keele University, UK
[7]School of Social and Community Medicine, University of Bristol, Bristol, UK
[8]Diabetes Research Centre, University of Leicester, Leicester, UK
[9]British Heart Foundation Data Science Centre, Health Data Research UK, London, UK
[10]Health Data Research UK, University College London, London, UK
[11]Barts Health NHS Trust, London, UK
[12]University College London Hospitals NHS Trust, London, UK

Abstract

Chronic kidney disease(CKD) is associated with increased risk of baseline mortality and severe COVID-19, but analyses across CKD stages, and comorbidities are lacking. In prevalent and incident CKD, we investigated comorbidities, baseline risk, COVID-19 incidence, and predicted versus observed 1-year excess death. In national English data(NHSD-TRE; n=56 million), we conducted a retrospective cohort study in prevalent and incident CKD(March 2020 to March 2021) of prevalence of comorbidities by incident and prevalent CKD, SARS-CoV-2 infection and mortality. We assessed baseline mortality risk, incidence and outcome of infection by comorbidities, controlling for age, sex and vaccination. We compared observed versus predicted 1-year mortality at varying population infection rates(IR) and pandemic-related relative risks(RR) using our published model in pre-pandemic CKD cohorts(NHSD TRE and CPRD). Among individuals with CKD(prevalent:1,934,585, incident:144,969), comorbidities were common(73.5% and 71.2% with ≥1 condition, and 13.2% and 11.2% with ≥3 conditions, in prevalent and incident CKD), and associated with SARS-CoV-2 infection, particularly dialysis/transplantation(OR 2.08, 95% CI 2.04-2.13) and heart failure(OR 1.73, 1.71-1.76), but not cancer(OR 1.01, 1.01-1.04). One-year all-cause mortality varied by age, sex, multimorbidity and CKD stage. Compared with 34,265 observed excess deaths, in NHSD-TRE and CPRD data respectively, we predicted 28,746(83.9%) and 24,546(71.6%) deaths (IR 10% and RR=3.0), and 23754(69.3%) and 20283(59.2%) deaths (observed IR 6.7% and RR 3.7). In the largest, national-level study to-date, individuals with CKD have high burden of comorbidities and multimorbidity, high risk of pre-pandemic mortality and a high risk of pandemic mortality. Treatment of comorbidities, non-pharmaceutical measures, and vaccination are priorities in people with CKD and management of long-term conditions is important during and beyond the pandemic.

Introduction

Chronic kidney disease (CKD) carries major global disease burden, as a risk factor for morbidity and mortality, and as the end syndrome of underlying risk factors and diseases[1,2], such as cancers[3] and CVD[4]. During the coronavirus (COVID-19) pandemic, CKD has been associated with poor prognosis[5,6]. Despite clinical and public health importance, CKD research to-date in all stages, multi-morbidity, or the general population[7] using national-level data has been limited.

The pandemic has had both direct (through infection) and indirect impact (through changes in health services, economic upheaval and behavioural factors[8,9]). The direct impact in individuals with CKD and other underlying conditions is related to baseline risk, influenced by age, sex, multimorbidity and other socio-demographic factors[10]. However, previous studies of COVID-19 in CKD have been small-scale (12-1099 cases[5]), mostly focused on end-stage CKD, and have ignored major comorbidities (either most common in CKD or related to risk of COVID-19 mortality). 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.

CKD, already known to increase baseline risk of mortality, is associated with increased risk of SARS-CoV2 (severe acute respiratory syndrome coronavirus 2) infection, disease severity, hospital[11] and intensive care admission[12], and mortality. The role of other risk factors and underlying conditions in risk of COVID-19 in people with CKD requires more detailed investigation[13–15]. There are clinical practice tools for risk stratification of COVID-19 patients in the community and hospitals, but inclusion of CKD is as a binary variable, and so the spectrum of risk faced by individuals with CKD has not been fully considered. Such analyses are important in risk communication to patients, public and health professionals, as well as policies to suppress infection rate (IR) such as social distancing and physical isolation. Meanwhile, more nuanced investigation of the risk associated with CKD may inform clinical care, COVID-19 vaccination strategies as well as public health approaches to CKD after the pandemic[16,17,18,19].

Using national, population-based electronic health records (EHR), in individuals with prevalent and incident CKD, we investigated: (1) underlying conditions; (2) mortality risk; (3)incidence of SARS-CoV-2 infection, and (4)prediction and validation of pandemic-related excess deaths.

Methods

Study design and Data sources

We conducted a retrospective, population-based cohort study using NHS Digital Trusted Research Environment for England (NHSD TRE)[20]: a national database developed for pandemic-related research, linking primary care[21], Hospital Episode Statistics Admitted Patient Care (HES APC), COVID-19 trajectories [22], COVID-19 vaccination, and mortality information from the Office for National Statistics (ONS) Civil Registration of Deaths (**Figure S1**). To investigate multi-morbidity, baseline risk, incidence, and mortality, in individuals with CKD (≥18 years), we defined "prevalent CKD" as ≥6 months prior to the onset of pandemic (1st March 2020) without prior history of COVID-19, and "incident CKD" as new onset from 1st March 2020 to 1st March 2021 without history of COVID-19 prior to developing CKD. To predict 1-year COVID-19-related excess deaths based on pre-pandemic mortality risk, prevalent CKD at 1st Jan 2019 was defined using similar criteria. To show applicability of our methods to less complete, less up-to-date datasets, we also used CPRD Gold data (as in our previous research[15]) to define prevalent CKD at 6th April 2014 by either diagnosed CKD or two eGFR measures (by MDRD-4 algorithm[23]) ≥6 months prior to index date.

Having an underlying condition, for all cohorts, was defined as having ≥6 months' history of the condition: (a) prior to index date for prevalent CKD, and (b) prior to incidence date for incident CKD. Number of underlying conditions, where stated, was based on 6 conditions: chronic obstructive pulmonary disease (COPD), asthma, cardiovascular disease (CVD), cancer, diabetes, and chronic liver disease. COVID-19 mortality was defined as mortality within 28 days of a positive test. For SARS-CoV-2 incidence rate in prevalent CKD, disease-free time was estimated from earliest date prior to death or first-dose vaccination. Incident CKD was defined as SARS-CoV-2 positive ≥14 days after developing CKD. Disease-free time was measured from date of incident CKD. Crude incidence rate did not account for vaccination or other factors.

Phenotypes

Definitions of underlying conditions were derived from Health Data Research (HDR) UK–CALIBER, a comprehensive platform with validated definitions of underlying conditions[24]. Phenotyping was performed (a) in primary care (GDPPR) using SNOMED CT concepts, and in secondary care (HES APC) using ICD-10 codes. For CKD phenotyping (including CKD stages, dialysis, and transplant), we extracted SNOMED CT concepts systematically using off-line NHS Digital SNOMED CT Browser (**Table S1**). CVD was defined as a composite of stroke (non-specified, ischaemic, haemorrhagic, transient ischaemic attack, subarachnoid haemorrhagic),

heart failure, arrhythmias, acute myocardial infarction, cardiomyopathy, atrial fibrillation, deep vein thrombosis, isolated calf vein thrombosis, and pulmonary embolism[25]. Obesity was defined as body mass index (BMI)>40kg/m^2. Diabetes included all types of diabetes. Implementation of phenotypes is publicly available:https://github.com/BHFDSC/CCU003_03/tree/main/phenotypes.

Statistical analysis

Underlying conditions: We estimated prevalence of underlying conditions in prevalent and incident CKD, stratifying by age, gender, CKD stage, or dialysis/transplantation. We compared prevalence of underlying conditions in infected versus non-infected for (a) all CKD patients, and (b) "non-survival" group, using odds ratio (Wald method) and Mantel-Haenszel chi-squared test with 95% confidence intervals. *Mortality risk:* With SARS-CoV-2 infection as exposure and 1-year all-cause mortality as outcome, we estimated adjusted relative risk (RR), stratified by underlying conditions, for both prevalent and incident CKD, using generalised linear model (GLM) with Poisson distribution (log link) after adjusting for: (a) age, and (b) age and other potential cofounders by exact matching based on ≥1 vaccination dose, age groups (5-year intervals) and sex, assessing matching quality using distributional plots. To estimate overall effect of having an underlying condition, analyses were repeated with GLM for each condition, reporting respective RRs (with "SARS-CoV-2 positive" as another potential confounder in exact matching). *Incidence of SARS-CoV-2 infection:* We estimated crude incidence rate of SARS-CoV-2 infection per 10,000 person-week, stratified by underlying conditions for incident and prevalent CKD. *Predicting and validating pandemic-related excess deaths:* By Kaplan-Meier analyses, we estimated pre-pandemic baseline risk of 1-year all-cause mortality for prevalent CKD in NHSD TRE (2019) and CPRD cohorts (2014). We validated our recent model[14,15] (to predict COVID-19-related excess death) using our risk estimates and applying 1-year population IR of 10%, and overall RR of mortality (set at 3) based on previous reports[15,26]. We predicted total excess deaths by: (a) age groups and number of underlying conditions, and (b) underlying conditions; using assumed, and observed IR, and RR. The analysis was performed according to a pre-specified analysis plan published on GitHub https://github.com/BHFDSC/CCU003_01) including implementations, and phenotypes.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. MD, MAM and AB had full access to all the data in the study and AB had final responsibility for the decision to submit for publication.

Results

Overall population characteristics

We included 1,934,585 individuals with prevalent CKD (mean age:77.4±12 years, 58.0% female, 12.7% CKD stage >3, 4.4% dialysis/transplantation) and 144,169 with incident CKD(mean age:73.9±12.8 years, 51.9% female, 9.2% CKD stage >3, 2.4% dialysis/transplantation)(**Figure S1, Table S2-S3**). Among those with prevalent and incident CKD, 91.5% and 86.6% were >60 years, 48.0% and 36.1% were >80 years. In the first year of the pandemic, in those with prevalent and incident CKD, 6.7% and 7.8% were infected, 1.8% and 1.7% had died from COVID-19, and 8.9% and 7.0% had died from all causes.

Underlying conditions

Comorbidities were more common in prevalent than incident CKD, and in males, in older individuals, and at CKD stage 4 and 5, especially CVD (prevalent CKD 42.5% vs incident CKD 39.6%) and diabetes (prevalent CKD 30.5% vs incident CKD 28.8%) (**Figure S2-S5**). Looking at comorbidity pairs, the most common combinations were two CVD subtypes (e.g. 20.5% for AF and CVD), diabetes with CVD (15.3%) and cancer with CVD (11.6%) in prevalent CKD (**Figure S2**). 73.5% and 13.2% of individuals with prevalent CKD and 71.2% and 11.2% of those with incident CKD had ≥1 and ≥3 underlying conditions, respectively (**Table S2-S3**). SARS-CoV-2 infection rates were higher in incident than prevalent CKD (e.g., 39.2 vs 28.1 per 10,000 person-weeks for chronic liver disease, 37.9 vs 25.8 for stage 5, 31.7 vs 24.1 for heart failure) (**Figure 1**). Comorbidities were associated with infection, compared with non-infected individuals, particularly for dialysis/transplantation (OR 2.08, 2.04-2.13 95% CI) and heart failure (OR 1.73, 1.71-1.76), but not for cancer (OR 1.01, 1.01-1.04). Across all comorbidities, association with infection was reduced in the non-surviving group. Cancer (OR 0.80, 0.78-0.82), atrial fibrillation (OR 0.90, 0.87-0.92) and chronic liver disease (OR 0.83, 0.74-0.93) were less likely in infected people with prevalent CKD who did not survive, compared with non-infected people. In non-survivors, only diabetes, dialysis/transplantation and asthma were more common in infected than non-infected cases in both prevalent and incident CKD (**Figure S5, Table S4**).

Mortality risk

One-year all-cause mortality varied by age, sex, multimorbidity and CKD stage (e.g. 0.2% in age ≤50 years, no comorbidities and stage 3; 29.9% in age >80 years, ≥3 comorbidities and stage 5) (**Figure S6**). The relative risk of 1-year all-cause mortality associated with SARS-CoV-2 infection was comparable between incident and prevalent cases of CKD, and highest for those on dialysis/transplantation (prevalent CKD: RR 1.70, 1.67-1.73; incident CKD: RR 1.50, 1.37-1.63), or having chronic liver disease (prevalent CKD: RR 1.61, 1.55-1.66; incident CKD: RR 1.85, 1.65-

2.06), after adjusting for age, sex, COVID-19 vaccination, and positive COVID-19 test (**Table 1**) with appropriate matching (**Figure S7)**. The relative risk of 1-year all-cause mortality was highest for diabetes (prevalent CKD: RR 1.32, 1.30-1.33; incident CKD: RR 1.38, 1.30-1.45), and asthma (prevalent CKD: RR 1.27, 1.24-1.30; incident CKD: RR 1.31, 1.21-1.42), after adjusting for age, sex, and first-dose vaccination (**Table S5**). The incidence risk of mortality was significantly lower in vaccinated CKD than non-vaccinated (**Table S6**) after exact matching and adjusting based on age, sex, and being tested positive for SARS-CoV-2 infection. Vaccine efficacy seemed to be highest in CKD patients with dialysis or asthma comparing to other underlying conditions.

Incidence of SARS-CoV-2 infection

The incidence of infection was higher in incident CKD (20.5 [95% CI 20-21] per 10,000 person weeks) than prevalent CKD (15.0, 14.9-15.1), across all underlying conditions and CKD stages, even after accounting for vaccination (**Figure 1, Table S7**), Incidence of infection was highest in individuals with dialysis/transplantation (prevalent CKD: 28.1, 27.6-28.6; incident CKD: 39.2, 35.2-43.6 per 10,000 person weeks) and lowest in those with cancer (prevalent CKD: 15.7, 15.5-15.9; incident CKD: 22.2, 21.2-23.4 per 10,000 person weeks).

Predicting excess death

Observed IR (6.7%) and observed RR (3.7) were used in our prediction model (**Figure S8**) with the NHSD TRE (1st Jan 2019, n= 1,727,130, mean age: 77.0±12.0 years, 58.4% female) and CPRD (6th April 2014, n= 174,648, 77.0 ±11.9 years, 61.2% female) cohorts of individuals with prevalent chronic kidney disease (**Table S8**). Pre-pandemic 1-year all-cause mortality in the CPRD cohort (**Figure S9**) was comparable to the NHSD TRE cohort. by number of underlying conditions, age, sex and CKD stage.

Using NHSD-TRE and CPRD data, our model predicted 28,746 (83.9%) and 24,546 (71.6%) deaths with IR 10% and RR=3.0, respectively, and 23754 (69.3%) and 20283 (59.2%) deaths, respectively, with IR 6.7% and RR 3.7, compared with 34,265 observed excess deaths (**Table 2**). For NHSD TRE data, the prediction of COVID-19 deaths was significantly improved using IR 10% and RR 3.0, compared with IR 6.7% and RR 3.7 (e.g. 90.6% vs 71.2% for COPD, 94.4% vs 74.2% for heart failure, and 90.3% vs 71.0% for cancer). The model under-predicted for asthma (77.0% vs 60.5%) and diabetes (76.0% vs 59.7%) and over-predicted for dialysis/transplantation (124.3% vs 97.7%)(**Table 3**). The predicted proportions of COVID-19 deaths by age group were comparable to the observed proportions, for both NHSD TRE and CPRD data (**Figure S10**), e.g. in individuals aged >80 years (observed 75.0% and predicted 73.8% using NHSD TRE data and 76.6% using CPRD data).

Discussion

In this large, nationally representative cohort study of individuals with CKD, we had four findings. First, comorbidities and multi-morbidity were common, and associated with SARS-CoV-2 infection and severe COVID-19. Second, one-year mortality risk was high and dependent on age, underlying condition, stage of CKD, and incidence or prevalence of CKD, ranging from 0.5% to 37.2%. Third, the UK burden of COVID-19 excess deaths in individuals with CKD was over 34000 in 1 year and predictable using a simple, parsimonious model and routine EHR. Fourth, we showed that vaccination was associated with reduced mortality risk.

Diabetes and CVD are well-documented as major risk factors and comorbidities in people with CKD, whether in epidemiologic[27,28] or therapeutic research[29]. We describe, for the first time, distribution of comorbidities and multimorbidity across the whole spectrum of CKD, both prevalent and incident CKD in up-to-date national data for England. These data are important for planning services for treatment and prevention in individuals with CKD both during and after the pandemic. For example, 7% of individuals with incident or prevalent CKD have both diabetes, and cancer; >10% have CVD and cancer. Projections of direct and indirect impact of COVID-19 have not considered overlap between diseases and treatments, probably leading to underestimation. Our finding of higher infection rates in those with dialysis/transplantation may be related to detection bias due to some regular monitoring of those patients for COVID-19 symptoms, resulting in a better detection of SARS-CoV-2 infection. In this context, developing a new condition (such as incident CKD) could potentially increase the contacts with health service that could have resulted in higher detection of infection in incident CKD than prevalent CKD. Despite that, the low rates observed for cancer patients could be related to shielding strategy in clinically vulnerable patients in the UK. Our results are in line with prior studies[13] showing higher infection rates in those with CKD. Future research should also address subtypes of CKD and trajectory by comorbidity profile to guide and prioritise preventive clinical and public health interventions.

We provide detailed large-scale, population-based analyses to provide patients, health professionals and policymakers with understanding of pre- and post-COVID-19 mortality risk in people with CKD, based on age, and underlying conditions, incident versus prevalent diseases. Despite increasing clinical, societal and scientific interest in precision medicine, CKD has not been comprehensively investigated whether in terms of aetiology, prognosis or prevention research[1,2,28]. Such granular, personalised data can inform risk prediction and public health projections to translational research and conversations with patients about individual risk. Moreover, such approaches are needed to help future research in long COVID.

Excess deaths have been the main metric to measure direct and indirect COVID-19 impact, whether overall or in individuals with particular diseases[14,15]. We present the first analyses in individuals with CKD. These are projections over 1 year based on a published model[15] and consistent with current estimates of the UK's COVID-19 deaths[26,27,30,31]. The variations in pre- and post-COVID-19 mortality based on age, and underlying conditions, are consistent with observed variation in mortality rates during the pandemic[27,32]. The greater prediction accuracy of our model using assumed IR and RR values (10% and 3 respectively), compared with observed values (6.7% and 3.7) is likely to reflect underestimation of infection rate, even in near-complete, national data. Further validation of our prediction model is required across different diseases, patterns of multimorbidity and countries. Our approach highlights the *feasibility* of large-scale use of EHR for pandemic preparedness, even less contemporary, less complete data (e.g. CPRD from 2014), and *validity* of our estimates of infection and excess deaths. For example, our infection rate estimates in non-dialysis patients with prevalent CKD (14.4, 14.3-14.5 per 10,000 person weeks) were comparable with a recent meta-analysis (16, 4-33 per 10,000 person weeks[33]).

Strengths and Limitations

This is the largest study to-date of individuals with CKD in national EHR to consider a wide range of comorbidities and COVID mortality but has several limitations. Laboratory testing was not available, and phenotyping was based on SNOMED CT concepts with potential underestimation. We used validated CALIBER phenotypes[25] and methods[34], but biases are possible[35]. We only investigated impact of underlying conditions, or effect of SARS-CoV-2 infection by individual comorbidities. Further studies should investigate comorbidity clusters and progression of CKD and outcomes. We were unable to study detailed ethnic categories due to data quality in EHR. Our model rests on baseline risks. Under- or over-estimation of excess deaths is possible for some underlying conditions being differentially affected by specific health policies (e.g. shielding), or by indirect effects of the pandemic (e.g. cancelled procedures).

Implications for research and policy

There are three policy implications. First, our findings are consistent with a "syndemic", describing convergence of an infectious disease, under-treated non-communicable diseases and social determinants of health[36] requiring multi-disciplinary, rather than traditional, disease- and specialty-specific responses. Second, given high comorbidity burden, particularly CVD and cancer, it is important to mitigate against indirect effects, likely to disproportionately affect people with CKD[14]. Third, routine data can provide patients, public, professionals, and policymakers with tailored risk information since mortality is highly variable based on age, sex, multi-morbidity and disease

stage, which can inform pre-pandemic and pandemic management, such as social isolation policies and vaccination prioritization in individuals with CKD.

There are three research implications. First, clustering approaches may inform and clarify subtype classification, trajectories, and risk prediction in CKD. Second, possible mechanisms underlying observed differences in mortality by age, comorbidities, ethnicity, stage of CKD and other factors need investigation. Third, pathophysiology of CKD as a risk factor and an outcome in COVID-19 warrants further study, informing aetiology, prevention, and intervention research.

Conclusions

In conclusion, individuals with CKD have high burden of multimorbidity and high risk of pre-pandemic mortality across all stages of CKD and in prevalent and incident disease. We showed that the direct burden of pandemic could be predicted using pre-pandemic, large-scale EHR data. The combined data for multimorbidity, CKD stage, and age could help prioritise patients for vaccination and post-COVID policies, and design of stratified pathways for CKD patients.

Acknowledgments

This work is carried out with the support of the BHF Data Science Centre led by HDR UK (BHF Grant no. SP/19/3/34678). This study makes use of de-identified data held in NHS Digital's TRE for England and made available via the BHF Data Science Centre's CVD-COVID-UK/COVID-IMPACT consortium. This work uses data provided by patients and collected by the NHS as part of their care and support. We would also like to acknowledge all data providers who make anonymised data available for research.

Ethical approval

Approval for the study was granted by the Independent Scientific Advisory Committee (20_074R) of the Medicines and Healthcare products Regulatory Agency in the UK in accordance with the Declaration of Helsinki. The North East-Newcastle and North Tyneside 2 research ethics committee provided ethical approval for the CVD-COVID-UK/COVID-IMPACT research programme (REC No 20/NE/0161).

Data availability

The data used in this study are available in NHS Digital's TRE for England, but as restrictions apply, they are not publicly available (https://digital.nhs.uk/coronavirus/coronavirus-data-services-updates/trusted-research-environment-service-for-england). The CVD-COVID-UK/COVID-IMPACT programme led by the BHF Data Science Centre (https://www.hdruk.ac.uk/helping-with-health-data/bhf-data-science-centre/) received approval to access data in NHS Digital's TRE for England from the Independent Group Advising on the Release of Data (IGARD) (https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/independent-group-advising-on-the-release-of-data) via an application made in the

Data Access Request Service (DARS) Online system (ref. DARS-NIC-381078-Y9C5K) (https://digital.nhs.uk/services/data-access-request-service-dars/dars-products-and-services). The CVD-COVID-UK/COVID-IMPACT Approvals & Oversight Board (https://www.hdruk.ac.uk/projects/cvd-covid-uk-project/) subsequently granted approval to this project to access the data within the TRE for England. The de-identified data used in this study was made available to accredited researchers only.

Funding statement

This study was funded by AstraZeneca UK Ltd. The British Heart Foundation Data Science Centre (grant No SP/19/3/34678, awarded to Health Data Research (HDR) UK) funded co-development (with NHS Digital) of the trusted research environment, provision of linked datasets, data access, user software licences, computational usage, and data management and wrangling support, with additional contributions from the HDR UK data and connectivity component of the UK governments' chief scientific adviser's national core studies programme to coordinate national covid-19 priority research. Consortium partner organisations funded the time of contributing data analysts, biostatisticians, epidemiologists, and clinicians.

Contributors

AB conceived the research question. AB, JBM, and TM obtained funding. AB and MD designed the study and analysis plan. SD, CS, and the BHF Data Science Centre CVD-COVID-UK/COVID-IMPACT consortium prepared the data, including EHR phenotyping in the CALIBER open portal. CS is the Director of the BHF Data Science Centre and coordinated approvals for and access to data within the NHS Digital TRE for CVD-COVID-UK/COVID-IMPACT. MD prepared the CKD cohorts (including phenotyping of CKD stages), designed incidence study, and performed statistical analysis. MAM provided all required implementations for adding phenotypes, and vaccination data in TRE England, beside insightful comments throughout research. AB and MD drafted the initial and final versions of the manuscript. All authors critically reviewed early and final versions of the manuscript.

Declaration of interests

DN is Director of Informatics Research for the UK Kidney Association and on the steering group for two Glaxo-SmithKline-funded studies in Sub Saharan Africa, unrelated to this research. JBM and TM are employed by AstraZeneca UK Ltd, a biopharmaceutical company. DAL has received funding from Wellcome, the European Research Council (ERC Advanced grant and a Horizon 2020 grant), US National Institute of Health, Diabetes UK, Roche Diagnostics and Medtronic Ltd for research unrelated to that presented here. KK is director of the University of Leicester Centre for Black Minority Ethnic Health, trustee of the South Asian Health Foundation,

and chair of the ethnicity subgroup of the UK Scientific Advisory Group for Emergencies (SAGE), he has acted as a consultant, speaker or received grants for investigator-initiated studies for AstraZeneca, Novartis, Novo Nordisk, Sanofi-Aventis, Lilly and Merck Sharp & Dohme, Boehringer Ingelheim, Bayer, Berlin-Chemie/Menarini Group, Janssen and Napp. AB is supported by research funding from the National Institute for Health Research (NIHR), British Medical Association, AstraZeneca, and UK Research and Innovation, and Trustee of the South Asian Health Foundation. HH is a National Institute for Health Research (NIHR) Senior Investigator. HH is funded by the National Institute for Health Research University College London Hospitals Biomedical Research Centre, Health Data Research UK (grant No. LOND1, which is funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care [England], Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division [Welsh Government], Public Health Agency [Northern Ireland], British Heart Foundation and Wellcome Trust). AB and HH are part of the BigData@Heart Consortium, funded by the Innovative Medicines Initiative-2 Joint Undertaking under grant agreement No. 116074. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA; it is chaired, by DE Grobbee and SD Anker, partnering with 20 academic and industry partners and ESC. Other authors report no conflicts of interest.

Figures

Figure 1. Incidence rate of SARS-CoV-2 infection by underlying conditions and stages of CKD in one year of COVID-19 pandemic for prevalent (n= 1,934,585) and incident (n= 144,969) chronic kidney disease; after controlling for COVID-19 first dose vaccination.

Legend: *CKD: chronic kidney disease; Dialysis/T: Dialysis/Transplantation; CVD: cardiovascular diseases, PAD: peripheral arterial disease; COPD: chronic obstructive pulmonary disease.*

References

1. Cockwell P, Fisher L-A. The global burden of chronic kidney disease. *Lancet*. 2020;395(10225):662-664.
2. Bikbov B, Purcell CA, Levey AS, 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*. 2020;395(10225):709-733.
3. Stengel B. Chronic kidney disease and cancer: a troubling connection. *J Nephrol*. 2010;23(3):253.
4. Gansevoort RT, Correa-Rotter R, Hemmelgarn BR, et al. Chronic kidney disease and cardiovascular risk: epidemiology, mechanisms, and prevention. *Lancet*. 2013;382(9889):339-352.
5. Henry BM, Lippi G. Chronic kidney disease is associated with severe coronavirus disease 2019 (COVID-19) infection. *Int Urol Nephrol*. Published online 2020:1-2.
6. Extance A. Covid-19 and long term conditions: what if you have cancer, diabetes, or chronic kidney disease? *BMJ*. 2020;368.
7. Saran R, Robinson B, Abbott KC, 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*. 2018;71(3 Suppl 1):A7.
8. Yao H, Chen J-H, Xu Y-F. Patients with mental health disorders in the COVID-19 epidemic. *The Lancet Psychiatry*. 2020;7(4):e21.
9. Pereira-Sanchez V, Adiukwu F, El Hayek S, et al. COVID-19 effect on mental health: patients and workforce. *The Lancet Psychiatry*. 2020;7(6):e29-e30.
10. Anderson RM, Hollingsworth TD, Baggaley RF, Maddren R, Vegvari C. COVID-19 spread in the UK: the end of the beginning? *Lancet*. 2020;396(10251):587-590.
11. Mafham MM, Spata E, Goldacre R, et al. COVID-19 pandemic and admission rates for and management of acute coronary syndromes in England. *Lancet*. 2020;396(10248):381-389.
12. Ghaem-Maghami S, Kehoe S, Burton MK. CEU ID: 11123. *Crisis*. 2020;10:12.
13. Reilev M, Kristensen KB, Pottegaard A, et al. Characteristics and predictors of hospitalization and death in the first 9,519 cases with a positive RT-PCR test for SARS-CoV-2 in Denmark: A nationwide cohort. Int J Epidemiol. 2020 Oct 1;49(5):1468-1481.
14. Lai AG, Pasea L, Banerjee A, et al. Estimating excess mortality in people with cancer and multimorbidity in the COVID-19 emergency. *BMJ Open*. 2020;10(11):e043828.
15. Banerjee A, Pasea L, Harris S, 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*. Published online 2020.
16. Methodology, COVID-19 – high risk shielded patient list identification. *NHS Digit Engl*. Published online 2020:version 9.
17. Who's at higher risk from coronavirus (COVID-19). *NHS Engl*. Published online 2020.
18. Implementing phase 3 of the NHS response to the COVID-19 pandemic. *NHS Engl*. Published online 2020:19-24.
19. Implementing personalised stratified follow-up pathways. *NHS Engl*. Published online 2020.
20. Wood A, Denholm R, Hollings S, et al. Linked electronic health records for research on a nationwide cohort of more than 54 million people in England: data resource. *bmj*. 2021;373.
21. NHS Digital for England. General Practice Extraction Service (GPES) Data for pandemic planning and research: a guide for analysts and users of the data.
22. Thygesen JH, Tomlinson C, Hollings S, et al. Understanding COVID-19 trajectories from a nationwide linked electronic health record cohort of 56 million people: phenotypes, severity, waves & vaccination. *medRxiv*. Published online 2021.

23. David-Neto E, Triboni AHK, Ramos F, et al. Evaluation of MDRD 4, CKD-EPI, BIS-1, and modified Cockcroft–Gault equations to estimate glomerular filtration rate in the elderly renal-transplanted recipients. *Clin Transplant*. 2016;30(12):1558-1563.
24. Denaxas S, Gonzalez-Izquierdo A, Direk K, et al. UK phenomics platform for developing and validating electronic health record phenotypes: CALIBER. *J Am Med Informatics Assoc*. 2019;26(12):1545-1559.
25. Banerjee A, Chen S, Pasea L, Lai AG, Katsoulis M, Denaxas S, Nafilyan V, Williams B, Wong WK, Bakhai A, Khunti K, Pillay D, Noursadeghi M, Wu H, Pareek N, Bromage D, McDonagh TA, Byrne J, Teo JTH, Shah AM, Humberstone B, Tang LV, Shah ASV, Rubboli A, Guo Y, Hu Y, Sudlow CLM, Lip GYH, Hemingway H. Excess deaths in people with cardiovascular diseases during the COVID-19 pandemic. Eur J Prev Cardiol. 2021 Dec 20;28(14):1599-1609.
26. Office for National Statistics. Deaths registered weekly in England and Wales, provisional: week ending 10 April 2020.
27. Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. *Nature*. 2020;584(7821):430-436.
28. Knight SR, Ho A, Pius R, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. *BMJ*. 2020;370.
29. Xie Y, Bowe B, Mokdad AH, et al. Analysis of the Global Burden of Disease study highlights the global, regional, and national trends of chronic kidney disease epidemiology from 1990 to 2016. *Kidney Int*. 2018;94(3):567-581.
30. Docherty AB, Harrison EM, Green CA, et al. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *bmj*. 2020;369.
31. Deaths with COVID-19 on the death certificate by area. *UK Gov*.
32. English Index of Multiple Deprivation. *Off Natl Stat*.
33. Chung EYM, Palmer SC, Natale P, et al. Incidence and outcomes of COVID-19 in people with CKD: a systematic review and meta-analysis. *Am J Kidney Dis*. 2021;78(6):804-815.
34. Forbes A, Gallagher H. Chronic kidney disease in adults: assessment and management. *Clin Med (Northfield Il)*. 2020;20(2):128.
35. Kuan V, Denaxas S, Gonzalez-Izquierdo A, 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*. 2019;1(2):e63-e77.
36. Horton R. Offline: COVID-19 is not a pandemic. *Lancet (London, England)*. 2020;396(10255):874.

Tables and Figures

Figure 1. Incidence rate of SARS-CoV-2 infection by underlying conditions and stages of CKD in one year of COVID-19 pandemic for prevalent (n= 1,934,585) and incident (n= 144,969) chronic kidney disease; after controlling for COVID-19 first dose vaccination.

Table 1. Association between SARS-CoV-2 infection and 1-year mortality by underlying condition for prevalent (n= 1934585) and incident (n= 144969) chronic kidney disease.

Table 2. Estimated one-year excess deaths by population infection rate and relative impact of the pandemic using Lancet 2020 model and prevalent CKD patients in two independent population-based cohorts (NHSD TRE and CPRD).

Table 3. Observed and predicted excess deaths (due to COVID-19) by underlying conditions over 1 year of the pandemic in individuals with prevalent chronic kidney disease (n= 1,934,585).

Table 1. Association between SARS-CoV-2 infection and 1-year mortality by underlying condition for prevalent (n= 1934585) and incident (n= 144969) chronic kidney disease.

		Relative risk (RR), 95% CI									
		Underlying conditions									
	Method	COPD	Asthma	PAD	Heart Failure	Atrial Fibrillation	Diabetes Mellitus	CVD	Cancer	Dialysis/ Transplantation	Chronic Liver Disease
Prevalent	AS	2.63 (2.57-2.69)	3.01 (2.93-3.09)	2.67 (2.57-2.77)	2.25 (2.21-2.30)	2.35 (2.31-2.40)	3.11 (3.06-3.17)	2.59 (2.56-2.63)	2.70 (2.64-2.75)	2.41 (2.30-2.52)	2.09 (1.88-2.31)
	M	1.15 (1.13-1.17)	1.27 (1.24-1.30)	1.16 (1.12-1.19)	1.14 (1.12-1.16)	1.12 (1.10-1.13)	1.32 (1.30-1.33)	1.19 (1.17-1.20)	1.13 (1.11-1.15)	1.18 (1.13-1.22)	1.07 (0.99-1.17)
Incident	AS	3.04 (2.77-3.34)	3.58 (3.21-3.97)	3.63 (3.05-4.29)	2.69 (2.48-2.91)	2.98 (2.76-3.21)	3.66 (3.40-3.93)	3.08 (2.91-3.26)	3.04 (2.80-3.31)	1..54 (1.23-1.90)	1.26 (0.87-1.76)
	M	1.13 (1.04-1.21)	1.31 (1.21-1.42)	1.25 (1.10-1.42)	1.14 (1.07-1.21)	1.15 (1.09-1.22)	1.38 (1.30-1.45)	1.20 (1.15-1.23)	1.15 (1.07-1.22)	0.90 (0.75-1.07	0.85 (0.62-1.15)

"AS" for age and sex; "M" indicates adjusting for age, sex, and vaccination using exact matching. COPD=chronic obstructive pulmonary disease, CVD=cardiovascular disease, PAD = Peripheral arterial disease.

Table 2. Estimated one-year excess deaths by population infection rate and relative impact of the pandemic using Lancet 2020 model and prevalent CKD patients in two independent population-based cohorts (NHSD TRE and CPRD).

Data used in Lancet 2020 model (Date of analysis of prevalent CKD)	**Relative risk of mortality associated with the pandemic**		**Population infection rate (%)**			
			Assumed			**Observed[1]**
			10	**40**	**80**	**6.7**
NHSD TRE (1st Jan 2019)	**Assumed**	**1.5**	14373(41.9)	57492(167.8)	114984(335.6)	9630(28.1)
		2	19164(55.9)	76656(223.7)	153312(447.4)	12840(37.5)
		3	**28746**(83.9)	114984(335.6)	229968(671.1)	19260(56.2)
	Observed[1]	**3.7**	35453(103.5)	141812(413.9)	283624(827.7)	**23754**(69.3)
CPRD (6th April 2014)	**Assumed**	**1.5**	12273(35.8)	49092(143.3)	98184(286.5)	8223(24)
		2	16364(47.8)	65456(191)	130912(382.1)	10964(32)
		3	**24546**(71.6)	98184(286.5)	196368(573.1)	16446(48)
	Observed[1]	**3.7**	20283(59.2)	20283(59.2)	20283(59.2)	**20283**(59.2)

[1]Observed parameters in NHSD TRE data. The value in parentheses shows percentage of observed excess deaths i.e., 34265.

Table 3. Observed and predicted excess deaths (due to COVID-19) by underlying conditions over 1 year of the pandemic in individuals with prevalent chronic kidney disease (n= 1,934,585).

COVID-19 deaths	**COPD**	**Asthma**	**Heart Failure**	**Atrial Fibrillation**	**Diabetes Mellitus**	**CVD**	**Cancer**	**Dialysis/ Transplantation**	**Total excess death** (% Predicted/Observed)
Observed	7890	6822	11394	12166	14617	22839	9979	2043	**34265 (100.0)**
Predicted, using assumed IR 10%/RR 3.0 (% Predicted/Observed)	7152 (90.6)	5251 (77)	10758 (94.4)	11706 (96.2)	11114 (76.0)	20014 (87.6)	9011 (90.3)	2539 (124.3)	**28746 (83.9)**
Predicted, using observed IR 6.7%/RR 3.7 (% Predicted/Observed)	5621 (71.2)	4126 (60.5)	8453 (74.2)	9199 (75.6)	8732 (59.7)	15726 (68.9)	7081 (71.0)	1997 (97.7)	**23754 (69.3)**

COVID-19 deaths	**Underlying conditions (n)**	**≤50**	**50-60**	**60-70**	**70-80**	**>80**	**Total excess death** (% Predicted/Observed)
Predicted, using assumed IR 10%/RR 3.0 (% Total Predicted)		202 (0.7)	516 (1.8)	1506 (5.2)	5314 (18.5)	21208 (73.8)	**28746 (83.9)**
	0	35 (0.1)	66 (0.2)	130 (0.5)	432 (1.5)	2252 (7.8)	2915 (8.5)
	1	66 (0.2)	132 (0.5)	356 (1.2)	1332 (4.6)	6884 (23.9)	8770 (25.6)
	2	67 (0.2)	178 (0.6)	546 (1.9)	1827 (6.4)	6992 (24.3)	9610 (28.0)
	3+	34 (0.1)	140 (0.5)	474 (1.6)	1723 (6.0)	5080 (17.7)	7451 (21.7)
Observed (% Total Observed)		133 (0.4)	543 (1.6)	1786 (5.2)	6109 (17.8)	25694 (75.0)	**34265 (100.0)**
	0	26 (0.1)	64 (0.2)	153 (0.4)	577 (1.7)	2982 (8.7)	3802 (11.1)
	1	43 (0.1)	184 (0.5)	458 (1.3)	1600 (4.7)	8205 (23.9)	10490 (30.6)
	2	45 (0.1)	176 (0.5)	653 (1.9)	1940 (5.7)	8326 (24.3)	11140 (32.5)
	3+	19 (0.1)	119 (0.3)	522 (1.5)	1992 (5.8)	6181 (18.0)	8833 (25.8)

IR: Infection Rate; RR: Relative Risk of COVID-19 pandemic compared with Baseline. Assumed IR/RR is based on Lancet 2020 model (Banerjee et al). Observed IR/RR were observed during pandemic in individuals with CKD.

Supplementary Material