



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



# HIV infection and COVID-19 death: a population-based cohort analysis of UK primary care data and linked national death registrations within the OpenSAFELY platform



Krishnan Bhaskaran, Christopher T Rentsch, Brian MacKenna, Anna Schultze, Amir Mehrkar, Chris J Bates, Rosalind M Eggo, Caroline E Morton, Sebastian C J Bacon, Peter Inglesby, Ian J Douglas, Alex J Walker, Helen I McDonald, Jonathan Cockburn, Elizabeth J Williamson, David Evans, Harriet J Forbes, Helen J Curtis, William J Hulme, John Parry, Frank Hester, Sam Harper, Stephen J W Evans, Liam Smeeth\*, Ben Goldacre\*

## Summary

**Background** Whether HIV infection is associated with risk of death due to COVID-19 is unclear. We aimed to investigate this association in a large-scale population-based study in England.

**Methods** We did a retrospective cohort study. Working on behalf of NHS England, we used the OpenSAFELY platform to analyse routinely collected electronic primary care data linked to national death registrations. We included all adults (aged  $\geq 18$  years) alive and in follow-up on Feb 1, 2020, and with at least 1 year of continuous registration with a general practitioner before this date. People with a primary care record for HIV infection were compared with people without HIV. The outcome was COVID-19 death, defined as the presence of International Classification of Diseases 10 codes U07.1 or U07.2 anywhere on the death certificate. Cox regression models were used to estimate the association between HIV infection and COVID-19 death; they were initially adjusted for age and sex, then we added adjustment for index of multiple deprivation and ethnicity, and then for a broad range of comorbidities. Interaction terms were added to assess effect modification by age, sex, ethnicity, comorbidities, and calendar time.

**Results** 17 282 905 adults were included, of whom 27 480 (0.16%) had HIV recorded. People living with HIV were more likely to be male, of Black ethnicity, and from a more deprived geographical area than the general population. 14 882 COVID-19 deaths occurred during the study period, with 25 among people with HIV. People living with HIV had higher risk of COVID-19 death than those without HIV after adjusting for age and sex: hazard ratio (HR) 2.90 (95% CI 1.96–4.30;  $p < 0.0001$ ). The association was attenuated, but risk remained high, after adjustment for deprivation, ethnicity, smoking and obesity: adjusted HR 2.59 (95% CI 1.74–3.84;  $p < 0.0001$ ). There was some evidence that the association was larger among people of Black ethnicity: HR 4.31 (95% CI 2.42–7.65) versus 1.84 (1.03–3.26) in non-Black individuals ( $p$ -interaction=0.044).

**Interpretation** People with HIV in the UK seem to be at increased risk of COVID-19 mortality. Targeted policies should be considered to address this raised risk as the pandemic response evolves.

**Funding** Wellcome, Royal Society, National Institute for Health Research, National Institute for Health Research Oxford Biomedical Research Centre, UK Medical Research Council, Health Data Research UK.

**Copyright** © 2020 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

## Introduction

Since it emerged in late 2019, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus that causes COVID-19, has infected more than 60 million people worldwide, causing more than 1.4 million deaths as of November, 2020.<sup>1</sup> Older age and male sex have been strongly associated with more severe outcomes; several comorbidities, including those that involve immunosuppression, also seem to be associated with higher risk of COVID-19 death.<sup>2</sup> However, little evidence exists on how HIV infection affects risk of poor outcomes from COVID-19.<sup>3</sup>

There is mixed evidence on the contribution of HIV to previous respiratory virus epidemics. HIV has been associated with a higher risk of severe outcomes from respiratory infections, including seasonal influenza,<sup>4,5</sup>

and people living with HIV at any stage of infection are considered a clinical risk group in seasonal influenza vaccination guidance in the UK.<sup>6</sup> However, the contribution of HIV infection to outcomes during the 2009 H1N1 influenza pandemic was unclear, with no substantive evidence that HIV-infected individuals were at increased risk of being infected or had worse outcomes, unless at an advanced stage of immunosuppression.<sup>7</sup> Little evidence from the SARS-CoV-2 pandemic exists. A large population-based cohort study in South Africa<sup>8</sup> found COVID-19 mortality risk among people living with HIV to be double the risk of those without HIV. A high prevalence of critical illness was observed among HIV-infected patients with COVID-19 in Madrid, Spain, although there was no non-HIV comparison group.<sup>9</sup> Other small studies of hospitalised patients have not

Lancet HIV 2021; 8: e24–32

Published Online

December 11, 2020

[https://doi.org/10.1016/S2352-3018\(20\)30305-2](https://doi.org/10.1016/S2352-3018(20)30305-2)

See [Comment](#) page e2

\*Joint principal investigators

Faculty of Epidemiology and Population Health, London

School of Hygiene & Tropical Medicine, London, UK

(Prof K Bhaskaran PhD,

C T Rentsch PhD, A Schultze PhD,

R M Eggo PhD,

Prof I J Douglas PhD,

H I McDonald PhD,

E J Williamson PhD,

H J Forbes PhD,

Prof S J W Evans MSc,

Prof L Smeeth PhD);

The DataLab, Nuffield

Department of Primary Care

Health Sciences, University

of Oxford, Oxford, UK

(B MacKenna MPharm,

A Mehrkar MBBChir,

C E Morton MBChB,

S C J Bacon BA, P Inglesby MPhil,

A J Walker PhD, D Evans MPhil,

H J Curtis DPhil, W J Hulme PhD,

B Goldacre MRCPsych);

The Phoenix Partnership,

TPP House, Horsforth, Leeds,

UK (C J Bates PhD,

J Cockburn BSc, J Parry MRCPG,

F Hester BSc, S Harper MSc)

Correspondence to:

Prof Krishnan Bhaskaran, Faculty

of Epidemiology and Population

Health, London School of

Hygiene & Tropical Medicine,

London WC1E 7HT, UK

[krishnan.bhaskaran@lshtm.](mailto:krishnan.bhaskaran@lshtm.ac.uk)

[ac.uk](mailto:krishnan.bhaskaran@lshtm.ac.uk)

## Research in context

### Evidence before this study

We searched PubMed on July 20, 2020, for population-based epidemiological studies comparing risk of severe COVID-19 outcomes between people with and without HIV. The search term “HIV AND (COVID OR coronavirus OR SARS-CoV-2) AND (rate OR hazard OR odds OR risk)” was used, with results filtered to articles from the past year with abstracts available, with no language restrictions. 78 papers were identified for screening but none included the comparison of interest. A relevant study separately identified on the medRxiv preprint server found a higher risk of COVID-19 mortality among people living with HIV compared with the general population in Western Cape, South Africa.

### Added value of this study

We used UK primary care data from more than 17 million people linked to national death registrations to compare

risk of COVID-19 death between people with and without HIV. People with HIV were at elevated risk of COVID-19 mortality compared with the general population without HIV, even after accounting for demographic characteristics, lifestyle-associated factors, and comorbidities. The association was particularly marked among people of Black ethnicity.

### Implications of all the available evidence

People with HIV in the UK seem to be at increased risk of COVID-19 mortality. Targeted policies should be considered to address this apparent raised risk as the pandemic response evolves. The monitoring and evaluation of the effects of HIV on COVID-19 outcomes in countries with a higher prevalence of HIV and lower levels of treatment and viral control should be prioritised.

detected any increased risk of severe outcomes in people living with HIV.<sup>10,11</sup>

We therefore aimed to investigate the association between HIV infection and COVID-19 death using population-based data from England.

## Methods

### Study design and population

A retrospective cohort study comparing the risk of COVID-19 death among people living with and without HIV was done within OpenSAFELY, a new data analytics platform in England created to address urgent COVID-19-associated questions, which has been described previously.<sup>2</sup> We used routinely collected electronic data from primary care practices using The Phoenix Partnership (TPP) SystemOne software. These data cover around 40% of the population in England and are linked to Office of National Statistics (ONS) death registrations. We included all adults (aged  $\geq 18$  years) alive and in follow-up on Feb 1, 2020, and with at least 1 year of continuous registration with a general practitioner (GP) before this date to ensure adequate capture of baseline data. We excluded people with missing age, sex, or index of multiple deprivation, because they probably indicate poor quality data.

### Procedures

The outcome was COVID-19 death, defined as a record for death in linked ONS data with the International Classification of Diseases 10 codes U071 (COVID-19, virus identified) or U072 (COVID-19, virus not identified) anywhere on the death certificate.<sup>12</sup> In a sensitivity analysis, we used a stricter outcome definition requiring confirmed COVID-19 (code U071) to be recorded as the underlying cause of death.

The main exposure was HIV status, and covariates considered in the analysis included age at Feb 1, 2020 (grouped as 18–39, 40–49, 50–59, 60–69, 70–79 and

$\geq 80$  years for descriptive analysis, and parametrised as a four-knot restricted cubic spline in regression models), sex, self-reported ethnicity (White, Mixed, South Asian, Black, and other), obesity (body-mass index [BMI] 30–34.9 kg/m<sup>2</sup> [class I], 35–39.9 kg/m<sup>2</sup> [class II], and  $\geq 40$  kg/m<sup>2</sup> [class III]), smoking status (never, former, and current), index of multiple deprivation quintile (derived from the patient's postcode at lower super output area level), and comorbidities considered potential risk factors for severe COVID-19 outcomes. These comorbidities were hypertension, asthma (categorised by use of oral steroids), chronic respiratory diseases other than asthma, chronic heart disease, diabetes (categorised according to the most recent glycated haemoglobin recorded in the 15 months before Feb 1, 2020), non-haematological and haematological cancer (both categorised by recency of diagnosis;  $< 1$ , 1–4.9, and  $\geq 5$  years), reduced kidney function (categorised by estimated glomerular filtration rate derived from the most recent serum creatinine measure; 30– $< 60$  mL/min per 1.73 m<sup>2</sup> and  $< 30$  mL/min per 1.73 m<sup>2</sup>), chronic liver disease, stroke or dementia, other neurological disease (motor neurone disease, myasthenia gravis, multiple sclerosis, Parkinson's disease, cerebral palsy, quadriplegia or hemiplegia, and progressive cerebellar disease), organ transplant, asplenia (splenectomy or a spleen dysfunction, including sickle cell disease), rheumatoid arthritis, lupus, or psoriasis, and other immunosuppressive conditions (permanent immunodeficiency ever diagnosed or aplastic anaemia or temporary immunodeficiency recorded within the past year). Post hoc, we additionally adjusted for household size (1–2, 3–5, 6–10, and  $\geq 11$ ) as a potential confounder among individuals with data available.

Information on HIV and all covariates was obtained by searching TPP SystemOne records before Feb 1, 2020, for

specific coded data, based on a GP subset of SNOMED-CT mapped to Read version 3 codes. 83 codes for HIV were included, but most people living with HIV were identified from four codes: 43C3 HIV positive (73%), X70M6 HIV infection (10%), Xa0ye HIV (4%), and X80bg HIV antibody (4%). People with no HIV code on their primary care record at the study start date formed the control group. All code lists and detailed information on their compilation are available online for inspection and use by the wider research community.

### Statistical analysis

Follow-up time for COVID-19 mortality was from Feb 1, 2020, until the date of COVID-19 death or June 22, 2020, which was the last date for which mortality data were complete. Death from causes other than COVID-19 were censored. Thus, our analysis did not incorporate the effects of competing risks on cumulative incidence, but rather focused on cause-specific hazards,<sup>13</sup> which is an appropriate method for estimating causal effects of covariates.<sup>14</sup> We used Cox regression models to estimate the association between HIV infection and COVID-19 mortality, initially unadjusted, then cumulatively adjusted in sequential models for age and sex, index of multiple deprivation and ethnicity, and smoking and obesity. The model adjusted for all these variables would control for confounding under the assumptions outlined in the directed acyclic graph presented in the appendix (p 3). In a subsequent secondary model, we added adjustment for all aforementioned comorbidities to explore the extent to which any effect of HIV might be mediated through these comorbidities, but they were considered to be causal pathway variables rather than confounders (appendix p 3). Models were stratified by the Sustainability and Transformation Partnership (a UK National Health Service [NHS] administrative region) of the patient's general practice to allow for geographical differences in baseline hazards. Multiple imputation (ten imputations) was used to account for missing ethnicity. A multinomial logistic imputation model including all covariates from the main modelling and an indicator for the outcome was used to generate the imputations, and model estimates from the resulting ten imputed datasets were combined by use of Rubin's rules. Individuals with missing BMI were assumed to be non-obese, and those with missing smoking data were assumed to be never-smokers; we did not use multiple imputation for these variables, because they are expected to be missing not at random in UK primary care.<sup>15</sup> In sensitivity analyses, we excluded individuals with missing data (complete case analysis). Proportional hazards were checked first by testing the slope of the Schoenfeld residuals (among those with complete ethnicity only) and second by fitting an interaction between analysis time and the HIV variable. To fit this interaction, time was categorised as 0–59, 60–89, 90 or more days from Feb 1, 2020. These categories were chosen to

capture the period before social distancing policies in the UK would have affected mortality, the period of peak COVID-19 mortality, and the period during which restrictions began to be eased, because these factors might have differentially affected the risk of acquiring infection (and thus the overall risk of COVID-19 death) in different population subgroups. Collinearity was checked by inspecting standard errors in unadjusted and adjusted models and calculating variance inflation factors in the covariate list in the primary adjusted model. We checked that our conclusions were not sensitive to the decision to use a time-to-event modelling framework by re-fitting the primary adjusted model using logistic regression with the binary outcome of COVID-19 death by June 22.

To investigate variables considered a priori to be potential effect modifiers, we fitted interaction terms (one at a time) between HIV status and age (<60 years and ≥60 years), sex, ethnicity (Black vs other), and presence of comorbidities (any vs none of the comorbidities included in the analysis [a sensitivity analysis excluded hypertension, because a past code for hypertension might not reflect ongoing disease]) to the main Cox model adjusted for age, sex, deprivation, ethnicity, smoking, and obesity. These variables were dichotomised owing to reduced power. Two-sided p values were from Wald tests on the interaction terms.

Cumulative mortality curves, standardised to adjust for different covariate distributions in the HIV and non-HIV groups, were generated. A Royston-Parmar model with the same covariates as the fully adjusted Cox model was fitted, with the baseline hazard modelled using a three-degrees-of-freedom spline.<sup>16</sup> The survival function was predicted from this model for every individual with HIV and averaged to produce the curve for the HIV group. To produce the standardised comparison curve, the survival functions were predicted and averaged again for the same individuals, but with HIV status set to 0.

All statistical analyses were done in Stata version 16.

NHS England is the data controller, TPP is the data processor, and the key researchers on OpenSAFELY are acting on behalf of NHS England. OpenSAFELY is hosted within the TPP environment which is accredited to the ISO 27001 information security standard and is NHS IG Toolkit compliant.<sup>17,18</sup> Patient data are pseudonymised for analysis and linkage using industry standard cryptographic hashing techniques. All pseudonymised datasets transmitted for linkage onto OpenSAFELY are encrypted, and access to the platform is through a virtual private network connection, restricted to a small group of researchers who hold contracts with NHS England and who only access the platform to initiate database queries and statistical models. All database activity is logged; only aggregate statistical outputs leave the platform environment following best practice for anonymisation of results, such as statistical disclosure control for low cell counts.<sup>19</sup> The OpenSAFELY platform adheres to the data protection principles of the UK Data Protection Act 2018 and the EU

General Data Protection Regulation 2016. In March, 2020, the Secretary of State for Health and Social Care used powers under the UK Health Service (Control of Patient Information) Regulations 2002 to require organisations to process confidential patient information for the purposes of protecting public health, providing healthcare services to the public and monitoring and managing the COVID-19 outbreak and incidents of exposure.<sup>20</sup> Taken together, these measures provided the legal bases to link patient datasets on the OpenSAFELY platform. This study was approved by the Health Research Authority (Research Ethics Committee reference 20/LO/0651) and by the London School of Hygiene & Tropical Medicine Ethics Board (reference 21863).

### Role of the funding source

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

### Results

17 282 905 adults were included; 27 480 (0·16%) were living with HIV and 17 255 425 (99·84%) did not have HIV (figure 1). People with HIV had similar median age, but a narrower age distribution overall, compared with those without HIV; there was a substantially lower proportion of people older than 70 years in the HIV group (1146 [4·2%] in those with vs 3070 564 [17·8%] in those without HIV; table 1). People living with HIV were more likely to be male, of Black ethnicity, and from a more deprived area. 921 (3·4%) individuals with HIV had chronic liver disease, compared with 99 214 (0·6%) of those in the comparison group.

There were 25 COVID-19 deaths among people with HIV during 10 680 person-years of follow-up and 14 857 COVID-19 deaths among those without HIV during 6·70 million person-years of follow-up (table 2). 68 638 individuals (0·4%) died from non-COVID-19 causes and were censored. Between Feb 1 and June 22, 2020, estimated cumulative COVID-19 mortality among all study participants, standardised to the covariate distribution of the HIV group, was 0·087% (95% CI 0·056–0·134) in people with HIV and 0·038% (0·036–0·040) in people without HIV (figure 2).

Crude risk of COVID-19 death was similar in people with and without HIV (hazard ratio [HR] 1·03, 95% CI 0·70–1·52), but after adjusting for age and sex, HIV was associated with a nearly three-fold higher risk of COVID-19 death (2·90, 1·96–4·30;  $p < 0·0001$ ; figure 3). This risk attenuated slightly after further adjustment for index of multiple deprivation, ethnicity, obesity, and smoking (2·59, 1·74–3·84;  $p < 0·0001$ ). A post-hoc adjustment for household size among the 15 063 659 individuals (87·2%) with these data available made little difference (2·56, 1·73–3·80). Adjustment for potentially

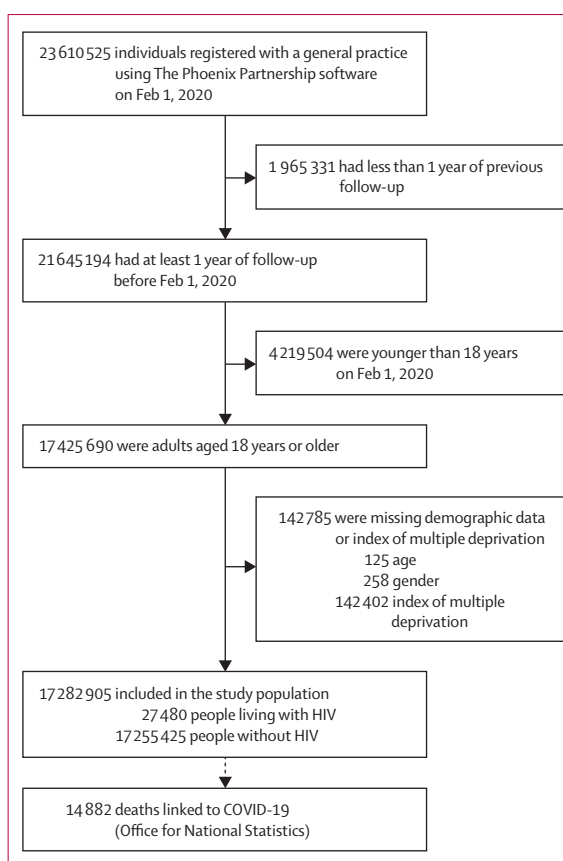


Figure 1: Study profile

mediating comorbidities reduced the HR slightly (2·30, 1·55–3·41).

There was some evidence that the association between HIV and COVID-19 death was larger in Black individuals (HR 4·31, 95% CI 2·42–7·65) than in other ethnic groups (1·84, 1·03–3·26;  $p$ -interaction=0·044; figure 3). There was no statistical evidence of interaction by age, sex, or presence of other comorbidities; the point estimate for the HIV-mortality association was larger for individuals with at least one comorbidity than for those with none, but CIs were wide, and excluding hypertension from the list of comorbidities attenuated the difference. The estimated association between HIV and COVID-19 death appeared to be larger early in the epidemic and reduced over time, suggesting possible non-proportionality of hazards, but it was not statistically significant ( $p=0·26$ ).

Further assessment for non-proportional hazards based on Schoenfeld residuals showed no evidence of non-proportionality by HIV status ( $p=0·32$ ), but there was evidence of non-proportional hazards in several adjustment covariates, namely ethnicity, sex, obesity, smoking, and index of multiple deprivation, so an additional sensitivity analysis was done. This analysis modelled non-proportionality by adding interaction terms between each of these variables and time-updated



	HIV group (N=27 480)	No HIV group (N=17 255 425)
<b>Age</b>		
18–39 years	6625 (24.1%)	5 908 794 (34.2%)
40–49 years	8486 (30.9%)	2 842 155 (16.5%)
50–59 years	8093 (29.5%)	3 043 868 (17.6%)
60–69 years	3130 (11.4%)	2 390 044 (13.9%)
70–79 years	937 (3.4%)	1 938 734 (11.2%)
≥80 years	209 (0.8%)	1 131 830 (6.6%)
Median (IQR)	48 (40–55)	49 (34–64)
<b>Sex</b>		
Male	17 780 (64.7%)	8 614 886 (49.9%)
Female	9 700 (35.3%)	8 640 539 (50.1%)
<b>BMI*</b>		
Not obese	21 539 (78.4%)	13 488 390 (78.2%)
30–34.9 kg/m <sup>2</sup> (obese class I)	3 777 (13.7%)	2 381 893 (13.8%)
35–39.9 kg/m <sup>2</sup> (obese class II)	1 475 (5.4%)	921 864 (5.3%)
≥40 kg/m <sup>2</sup> (obese class III)	689 (2.5%)	463 278 (2.7%)
<b>Smoking status†</b>		
Never	13 398 (48.8%)	8 634 105 (50.0%)
Former	7 922 (28.8%)	5 685 069 (32.9%)
Current	6 160 (22.4%)	2 936 251 (17.0%)
<b>Ethnicity</b>		
White	12 708 (46.2%)	10 859 301 (62.9%)
Mixed	1 390 (5.1%)	168 649 (1.0%)
South Asian	1 175 (4.3%)	1 021 252 (5.9%)
Black	7 143 (26.0%)	332 971 (1.9%)
Other	645 (2.3%)	319 845 (1.9%)
Missing	4 419 (16.1%)	4 553 407 (26.4%)
<b>Index of multiple deprivation</b>		
1 (least deprived)	2 790 (10.2%)	3 494 398 (20.3%)
2	4 127 (15.0%)	3 473 416 (20.1%)
3	5 108 (18.6%)	3 480 657 (20.2%)
4	6 863 (25.0%)	3 475 019 (20.1%)
5 (most deprived)	8 592 (31.3%)	3 331 935 (19.3%)
<b>Comorbidities‡</b>		
Hypertension	5 290 (19.3%)	3 666 002 (21.2%)
Chronic respiratory disease	1 095 (4.0%)	703 335 (4.1%)
Chronic heart disease	1 444 (5.3%)	1 167 008 (6.8%)
Chronic liver disease	921 (3.4%)	99 214 (0.6%)
Stroke or dementia	559 (2.0%)	389 828 (2.3%)
Other neurological disease	239 (0.9%)	170 336 (1.0%)
Organ transplant	72 (0.3%)	19 933 (0.1%)
Asplenia	89 (0.3%)	27 845 (0.2%)
Rheumatoid arthritis, lupus, or psoriasis	1 233 (4.5%)	877 666 (5.1%)
Other immunosuppressive conditions	53 (0.2%)	4 211 (0.0%)

(Table 1 continues in next column)

calendar time and had little effect on the HR for HIV (adjusted HR 2.49 [95% CI 1.60–3.87] after adding interaction terms vs 2.59 [1.74–3.84] in the main analysis).

Results were also similar in sensitivity analyses using a stricter outcome definition (adjusted HR 2.60, 95% CI

	HIV group (N=27 480)	No HIV group (N=17 255 425)
(Continued from previous column)		
<b>Asthma</b>		
With no oral steroid use	3 460 (12.6%)	2 454 236 (14.2%)
With oral steroid use	433 (1.6%)	291 606 (1.7%)
<b>Diabetes</b>		
With HbA <sub>1c</sub> <5.8 mmol/mol	1 521 (5.5%)	1 037 513 (6.0%)
With HbA <sub>1c</sub> ≥5.8 mmol/mol	739 (2.7%)	485 895 (2.8%)
With no recent HbA <sub>1c</sub> measure	449 (1.6%)	193 665 (1.1%)
<b>Cancer (non-haematological)</b>		
Diagnosed <1 year ago	108 (0.4%)	80 004 (0.5%)
Diagnosed 1–4.9 years ago	372 (1.4%)	233 977 (1.4%)
Diagnosed ≥5 years ago	804 (2.9%)	541 921 (3.1%)
<b>Haematological malignancy</b>		
Diagnosed <1 year ago	30 (0.1%)	8 691 (0.1%)
Diagnosed 1–4.9 years ago	114 (0.4%)	27 648 (0.2%)
Diagnosed ≥5 years ago	421 (1.5%)	63 076 (0.4%)
<b>Reduced kidney function</b>		
Estimated GFR 30–60 mL/min per 1.73m <sup>2</sup>	1 427 (5.2%)	1 006 298 (5.8%)
Estimated GFR <30 mL/min per 1.73m <sup>2</sup>	134 (0.5%)	77 964 (0.5%)

Data are n (%) unless otherwise stated. BMI=body-mass index. HbA<sub>1c</sub>=glycated haemoglobin. GFR=glomerular filtration rate. \*Missing BMI included in not obese (HIV group n=4550, non-HIV group n=3 737 207). †Missing smoking included in never smoker (HIV group n=457, non-HIV group n=720 500). ‡The number of individuals with any comorbidity (of those listed) was 12 984 (47.3%) in HIV group and 7 914 272 (45.9%) in non-HIV group. The number with ever hepatitis C was 1491 (5.4%) in HIV group and 38 235 (0.2%) in the non-HIV group.

**Table 1: Demographic characteristics of people with and without HIV in OpenSAFELY**

1.71–3.96), using a logistic modelling framework (adjusted odds ratio 2.59, 95% CI 1.74–3.84), restricting to those with known ethnicity (adjusted HR for HIV 2.49, 1.60–3.87), and slightly attenuated among those with complete BMI and smoking data (adjusted HR 2.19, 1.39–3.43). We found no evidence of collinearity in the main adjusted model (variance inflation factors all <1.3).

## Discussion

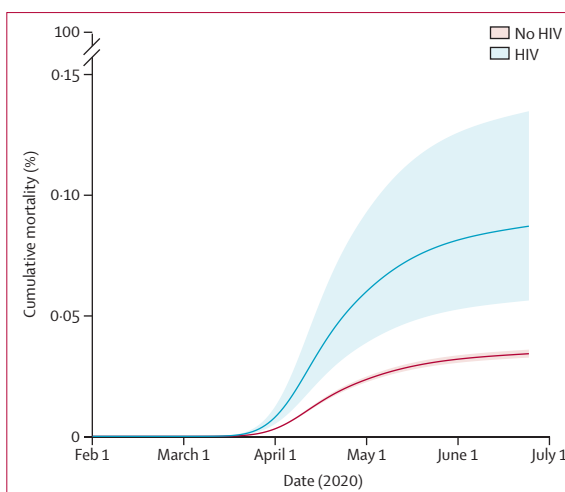
In this large population-based study using data from the OpenSAFELY platform in England, we found people with HIV to be at more than twice the risk of COVID-19 death compared with people without HIV, after accounting for demographic characteristics and lifestyle-associated factors. Absolute cumulative COVID-19 mortality was low, with less than 0.1% of people with HIV dying with COVID-19 as a cause during the study period, reflecting the young age profile of this population. The association between HIV and COVID-19 mortality seemed to be particularly pronounced among people of Black ethnicity, with HIV associated with a 4.3-fold higher risk of COVID-19 death in this group.

	HIV group (N=25)	No HIV group (N=14 857)
<b>Age</b>		
<60 years	10 (40%)	901 (<1%)
≥60 years	15 (60%)	13 956 (94%)
<b>Sex</b>		
Male	18 (72%)	8110 (55%)
Female	7 (28%)	6747 (45%)
<b>Body-mass index</b>		
Not obese	15 (60%)	11 044 (74%)
Obese	10 (40%)	3813 (26%)
<b>Ethnicity</b>		
Black*	11 (44%)	284 (2%)
White, other, or unknown	14 (56%)	14 573 (98%)
<b>Common comorbidities†</b>		
Hypertension	15 (60%)	9752 (66%)
Diabetes	14 (56%)	5372 (36%)
Reduced kidney function‡	9 (36%)	6659 (45%)
Any comorbidity (including hypertension)§	23 (92%)	13 897 (94%)
Any comorbidity (excluding hypertension)§	21 (84%)	13 168 (89%)

\*Self-report as African, Caribbean or other Black. †All other individual comorbidities were present in five or fewer individuals in the HIV group and are redacted as per our protocol. ‡Estimated glomerular filtration rate <30 mL/min per 1.73 m<sup>2</sup>. §From comorbidities listed in table 1.

**Table 2: Characteristics of people who died with COVID-19 as a listed cause of death**

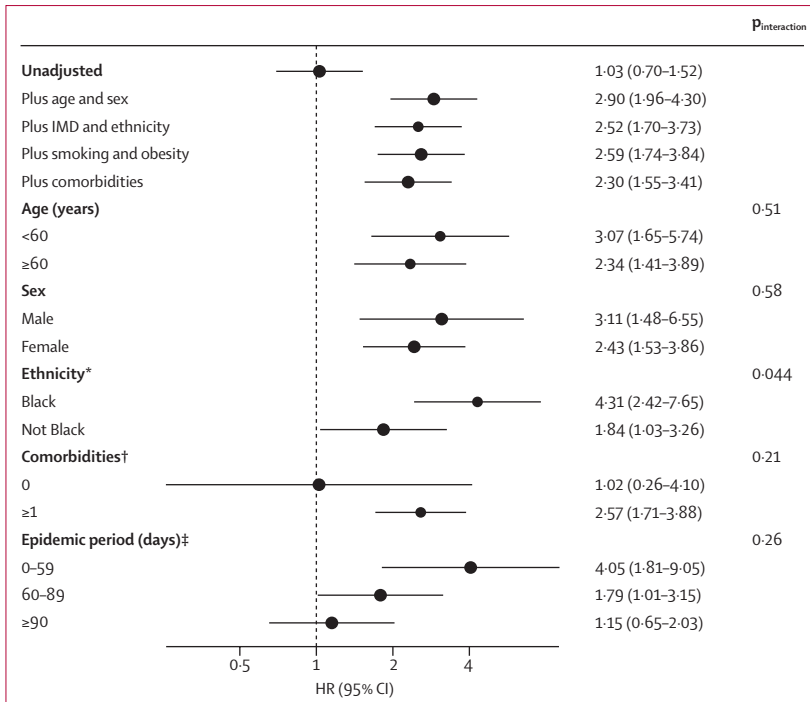
Few comparable data have been published. Our results are similar to those from a large population-based cohort study from South Africa involving 3·5 million individuals,<sup>8</sup> which found an adjusted HR of 2·14 (95% CI 1·98–3·43), compared with 2·30 (1·55–3·41) in our study. It is noteworthy that the association was similar in an English setting in which levels of antiretroviral treatment and viral suppression are higher than in South Africa; however, our observation of this association is consistent with stratified analyses in the South African data, which found a more than doubled risk of COVID-19 death even among those with a recent prescription for antiretroviral treatment and a recent viral load measure of less than 1000 copies per mL, although limitations of that study included a lack of control for socioeconomic status and limited comorbidity data. Early data from small cohorts and case series in other countries included a large proportion of patients with severe disease.<sup>21–23</sup> A single-centre cohort in Madrid, Spain, found that, among 51 people living with HIV and diagnosed with COVID-19, 28 (55%) required hospital admission and 13 (25%) had severe disease; the mortality rate among COVID-19 cases co-infected with HIV in this study was around double that in similarly aged individuals in the general population.<sup>9</sup> Associations between HIV and adverse COVID-19 outcomes seem to be smaller among patients already hospitalised with COVID-19; the HR for mortality in the South African study was attenuated



**Figure 2: Cumulative COVID-19 mortality during the study period by HIV status with 95% CI, standardised to covariate distribution of the HIV group**  
Note that cumulative COVID-19 mortality is not restricted to individuals infected with severe acute respiratory syndrome coronavirus 2 but rather represents the cumulative incidence of acquiring infection and then progressing to death with COVID-19 listed as a cause. Cumulative mortality predicted from a Royston-Parmar model including age, sex, index of multiple deprivation, ethnicity, smoking, and obesity, with the baseline hazard parametrised as a three-degrees-of-freedom cubic spline; predictions standardised to the covariate distribution of the HIV group. This analysis was done for individuals with complete ethnicity data only, because computational limitations prevented implementation in the dataset with multiply imputed ethnicity.

to 1·45 (1·14–1·84) when restricted to hospitalised patients,<sup>8</sup> and emerging data on patients hospitalised in the UK found a similar association,<sup>24</sup> whereas an analysis of patients hospitalised in New York found no difference in adverse outcomes between HIV infected and uninfected individuals.<sup>11</sup> However, any role that HIV might have in increasing the risk of infection or development of severe disease is effectively conditioned out by restriction to hospitalised individuals who are already infected with SARS-COV-2 and who probably have severe disease at the point of inclusion.<sup>25</sup>

The larger association between HIV and COVID-19 death among people of Black ethnicity has not previously been described. We have previously shown a larger overall risk of COVID-19 death in Black and minority ethnic (BAME) compared with white groups in England,<sup>2</sup> and a systematic review suggested that BAME individuals might have a higher risk of acquiring COVID-19 infection and having worse clinical outcomes.<sup>26</sup> HIV infection could plausibly exacerbate both, although it is not possible to delineate from our study which parts of the pathway are more or less affected by HIV status. BAME individuals are also at higher risk of negative HIV outcomes, including virological rebound.<sup>27</sup> Understanding the reasons for the disproportionately large association between HIV and COVID-19 mortality in this group will be a priority if effective policies are to be developed to mitigate any increased risks.



**Figure 3: HRs for the association between HIV and COVID-19 mortality**  
 All stratified models (by age, sex, ethnicity, comorbidities, epidemic period) were adjusted for age, sex, IMD, ethnicity. IMD=index of multiple deprivation. HR=hazard ratio. \*Black is defined as self-report as African, Caribbean, or other Black; a similar pattern was seen in a direct comparison between Black (HR 4.81, 95% CI 2.63-8.80) and white (2.02, 1.05-3.89) among individuals with complete ethnicity data. †Comorbidities refers to diagnosed hypertension, chronic respiratory disease, asthma, chronic cardiac disease, diabetes, non-haematological cancer, haematological cancer, chronic liver disease, stroke and dementia, other neurological disease, reduced kidney function, organ transplant, asplenia, rheumatoid arthritis, lupus, and psoriasis, and other immunosuppressive conditions; the model stratified by comorbidities was additionally adjusted for these comorbidities as individual covariates; excluding hypertension from the list of comorbidities gave stratified HRs of 1.57 (0.59-4.20) for individuals without comorbidities and 2.52 (1.64-3.87) for those with comorbidities (p-interaction=0.39). ‡Days from Feb 1, 2020; the three categories were chosen to capture the period before social distancing policies in the UK would have affected mortality, the period of peak COVID-19 mortality, and the period during which restrictions began to be eased.

Our study was large, with sufficient power to estimate the association between HIV and COVID-19 death, accounting for demographic characteristics, lifestyle-associated factors, including BMI and smoking, and relevant comorbidities. We tested the robustness of our findings to different missing data approaches and to non-proportional hazards. We investigated effect modification by key covariates and were able to detect differences by ethnicity. However, the relatively small number of deaths among individuals with HIV, reflecting the young age distribution of the HIV group, prevented definitive conclusions about the role of comorbidities and changes in the role of HIV over time. Among people with no comorbidities, those with HIV had no observed increase in risk of COVID-19 death, but the non-significant interaction was inconclusive given reduced power. We also noted a suggestion of a larger association between HIV and COVID-19 death early in the epidemic. If confirmed, this finding might suggest a higher risk of infection before widespread social distancing was

introduced, but we could not rule out that it reflected chance variation. Analysis of pre-mortality indicators of severe disease, such as hospitalisation, would have greater statistical power to clarify this, but we could not obtain linked hospital data for people living with HIV, because HIV codes are considered sensitive in NHS data and transfer is legally restricted.<sup>28</sup>

For similar reasons, we were unable to include data on antiretroviral therapy use, viral suppression, CD4 count, or previous AIDS-defining illnesses likely to be captured only in specialist care, so it was not possible to stratify results or quantify how much of the increased risk was driven by the minority of individuals with poorly controlled HIV. HIV in the UK is commonly managed in specialist clinics, but restrictions on the sharing of HIV-associated codes and policy guidance that cautions against sharing HIV-associated information with patients' primary care providers meant that we could not access key data from routine HIV care. Nevertheless, we do not consider the absence of these data to substantially undermine the utility of our findings, given that 94% of diagnosed HIV-positive individuals in the UK are on antiretrovirals and have good viral suppression (97% on therapy, and 97% of these with undetectable viral load).<sup>29</sup> The association between HIV and COVID-19 death is unlikely to have been driven entirely by the 6% assumed untreated or with detectable viral load; if all deaths had been within this group, it would amount to more than 1.5% becoming infected and dying from COVID-19, compared with less than 0.01% of the general population. There are no available routinely collected data on injection drug use, occupation, or contact patterns. HIV diagnoses recorded in primary care are expected to be highly specific, but some under-ascertainment is possible; this should have had little effect, because HIV-positive people misclassified as HIV-negative would have made up only a tiny proportion of our comparator group. Any resulting bias is likely to have been towards the null, particularly if high-risk people not engaged with primary care were more likely to be missed. We included clinically suspected (non-laboratory confirmed) COVID-19 deaths; because there was little testing early in the pandemic, some non-COVID-19 deaths might have been misclassified as COVID-19 deaths. We assume any such misclassification will have been non-differential with respect to HIV status, which could again have led to bias towards the null, but given that more than 90% of COVID-19 deaths in our analysis had a code indicating that SARS-CoV-2 infection was laboratory confirmed, any such bias is likely to be small. Although we included deaths with a COVID-19 code anywhere on the death certificate, more than 90% of outcomes had COVID-19 entered as the underlying cause of death. A sensitivity analysis restricting outcomes to confirmed COVID-19 listed as the underlying cause of death gave similar results to the primary analysis. There might have been some informative censoring if individuals censored



owing to death without COVID-19 were also at higher risk of COVID-19 death, but adjustment for age (the most important risk factor for both COVID-19 and non-COVID-19 death) should have minimised this. Imperfect measurement of adjustment covariates could have led to some residual confounding; for example, index of multiple deprivation quintile is unlikely to have perfectly captured all confounding due to socioeconomic factors. Our study population, derived from one electronic health record software system, is not fully geographically representative of the broader population of England. In particular, London, which has a relatively high prevalence of HIV nationally, is under-represented in our data.

Our findings suggest that people living with HIV might be a high-risk group for COVID-19 death, indicating a need to consider targeted policies for this group. People living with HIV might also need priority consideration if and when a vaccine against SARS-CoV-2 becomes available. Our findings might also have implications globally. Levels of antiretroviral therapy use and viral suppression in England are high, whereas prevalence of HIV is relatively low in global terms;<sup>29</sup> the impact of HIV on the progression of the pandemic in other settings will need to be carefully monitored. Future studies should prioritise better understanding of the drivers of increased COVID-19 mortality among people living with HIV and the role of ethnicity. Although it is of particular interest to separate how HIV status affects risk of SARS-CoV-2 infection versus progression to severe disease or death once infected, this is challenging, because most SARS-CoV-2 testing is focused on individuals at high risk or with symptoms and unlikely to capture those with mild or asymptomatic infection; should mass population-level testing become widespread this will be a more achievable research priority. Future studies might also prioritise assessing how the association between HIV and severe COVID-19 outcomes is affected by comorbidities (including consideration of hypothesised premature ageing among people living with HIV<sup>30</sup>), viral load, CD4 count, and antiretroviral treatment. In England, despite a strong overall infrastructure of linkable electronic health records data, research is hampered by policy guidance that has generally led to restrictions in the sharing and flow of HIV-associated data, despite apparent support for sharing for public health purposes.<sup>31</sup> We propose a review of guidance to ensure that people living with HIV are not unnecessarily excluded from research.

In this large population-based study, people living with HIV in England had an elevated risk of COVID-19 death after accounting for demographic characteristics and lifestyle-associated factors. Absolute cumulative COVID-19 mortality was relatively low, reflecting the young age profile of the population of people living with HIV. Monitoring the role of HIV in COVID-19 outcomes in other settings will be crucial as the pandemic progresses.

#### Contributors

BG conceived the OpenSAFELY platform and the approach. LS and BG led the project overall and are guarantors. KB designed the study, did the analysis, and wrote the first draft. SB led on software development. AM led on information governance. CJB, CEM, SCJB, PI, JC, DE, JP, and SH curated data; KB, CTR, BM, CJB, CEM, IJD, AJW, HIM, JC, HJF, HJC, and JP conceptualised disease categories and code lists; KB and EJW wrote statistical analysis code; EJW, HJC, LS, and BG obtained ethical approvals; CJB, CEM, SCJB, PI, AJW, JC, DE, WJH, and FH contributed to software development; and KB, CTR, BM, AS, AM, RME, CEM, IJD, SJWE, LS, and BG reviewed and edited the manuscript. All authors were involved in design and conceptual development and reviewed and approved the final manuscript.

#### Declaration of interests

We declare no competing interests.

#### Data sharing

All data were linked, stored and analysed securely within the OpenSAFELY platform. All code is shared openly for review and re-use under MIT open license. Detailed pseudonymised patient data is potentially re-identifiable and therefore not shared. We rapidly delivered the OpenSAFELY data analysis platform without previous funding to deliver timely analyses on urgent research questions in the context of the global COVID-19 health emergency. Now that the platform is established, we are developing a formal process for external users to request access in collaboration with NHS England; details of this process will be published shortly on the OpenSAFELY website.

#### Acknowledgments

No dedicated funding has yet been obtained for this work. TPP provided technical expertise and infrastructure within their data centre pro bono in the context of a national emergency. KB holds a Sir Henry Dale fellowship jointly funded by Wellcome and the Royal Society (107731/Z/15/Z). AS is employed by the London School of Hygiene & Tropical Medicine on a fellowship sponsored by GlaxoSmithKline. AM is funded by NHS Digital. IJD holds grants from the National Institute of Health Research (NIHR) and GlaxoSmithKline. HIM is funded by the NIHR Health Protection Research Unit in Immunisation, a partnership between Public Health England and London School of Hygiene & Tropical Medicine. EJW holds grants from the Medical Research Council (MRC). HF holds a UK Research and Innovation (UKRI) fellowship. RME is funded by Health Data Research UK (MR/S003975/1) and MRC (MC\_PC 19065). LS reports grants from Wellcome, MRC, NIHR, UKRI, British Council, GlaxoSmithKline, British Heart Foundation, and Diabetes UK outside this work. BG's work on better use of data in health care more broadly is funded in part by NIHR Oxford Biomedical Research Centre, NIHR Applied Research Collaboration Oxford and Thames Valley, the Mohn-Westlake Foundation, NHS England, and the Health Foundation; all DataLab staff are supported by BG's grants on this work.

The views expressed are those of the authors and not necessarily those of the NIHR, NHS England, Public Health England or the Department of Health and Social Care. Funders had no role in the study design, collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

#### References

- 1 European Centre for Disease Prevention and Control. COVID-19 situation update worldwide, as of 26 November 2020. European Centre for Disease Prevention and Control, 2020. <https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases> (accessed Nov 26, 2020).
- 2 Williamson EJ, Walker AJ, Bhaskaran K, et al. OpenSAFELY: factors associated with COVID-19 death in 17 million patients. *Nature* 2020; **584**: 430–36.
- 3 Fung M, Babik JM. COVID-19 in immunocompromised hosts: what we know so far. *Clin Infect Dis* 2020; published online June 27. <https://doi.org/10.1093/cid/ciaa863>.

For OpenSAFELY see <https://opensafely.org/>

For code see <https://github.com/opensafely/hiv-research>

- 4 Kenmoe S, Bigna JJ, Fatawou Modiyingi A, et al. Case fatality rate and viral aetiologies of acute respiratory tract infections in HIV positive and negative people in Africa: the VARIAFRICA-HIV systematic review and meta-analysis. *J Clin Virol* 2019; **117**: 96–102.
- 5 Neuzil KM, Reed GW, Mitchel Jr EF, Griffin MR. Influenza-associated morbidity and mortality in young and middle-aged women. *JAMA* 1999; **281**: 901–07.
- 6 Public Health England. Influenza: the green book, chapter 19. In: Ramsay M, ed. Immunisation against infectious disease. London: Public Health England, 2019.
- 7 Sheth AN, Patel P, Peters PJ. Influenza and HIV: lessons from the 2009 H1N1 influenza pandemic. *Curr HIV/AIDS Rep* 2011; **8**: 181–91.
- 8 Boulle A, Davies M-A, Hussey H, et al. Risk factors for COVID-19 death in a population cohort study from the Western Cape Province, South Africa. *Clin Infect Dis* 2020; published online Aug 29. <https://doi.org/10.1093/cid/ciaa1198>.
- 9 Vizcarra P, Pérez-Eliás MJ, Quereda C, et al. Description of COVID-19 in HIV-infected individuals: a single-centre, prospective cohort. *Lancet HIV* 2020; **7**: e554–64.
- 10 Cooper TJ, Woodward BL, Alom S, Harky A. Coronavirus disease 2019 (COVID-19) outcomes in HIV/AIDS patients: a systematic review. *HIV Med* 2020; **21**: 567–77.
- 11 Sigel K, Swartz T, Golden E, et al. Covid-19 and people with HIV infection: outcomes for hospitalized patients in New York City. *Clin Infect Dis* 2020; published online June 28. <https://doi.org/10.1093/cid/ciaa880>.
- 12 WHO. Emergency use ICD codes for COVID-19 disease outbreak. World Health Organization, 2020. <https://www.who.int/classifications/icd/covid19/en/> (accessed Aug 6, 2020).
- 13 Pintilie M. Analysing and interpreting competing risk data. *Stat Med* 2007; **26**: 1360–67.
- 14 Bhaskaran K, Racht B, Evans S, Smeeth L. Beta-blockers and prostate cancer survival—interpretation of competing risks models. *Eur Urol* 2013 Oct; **64**: e86–87.
- 15 Bhaskaran K, Smeeth L. What is the difference between missing completely at random and missing at random? *Int J Epidemiol* 2014; **43**: 1336–39.
- 16 Royston P, Parmar MKB. Flexible parametric models for censored survival data with application to prognostic modelling and estimation of treatment effects. *Stat Med* 2002; **21**: 1275–97.
- 17 NHS Digital. BETA—Data security standards. NHS Digital, 2020. <https://digital.nhs.uk/about-nhs-digital/our-work/nhs-digital-data-and-technology-standards/framework/beta--data-security-standards> (accessed Aug 6, 2020).
- 18 NHS Digital. Data security and protection toolkit. NHS Digital, 2020. <https://digital.nhs.uk/data-and-information/looking-after-information/data-security-and-information-governance/data-security-and-protection-toolkit> (accessed Aug 6, 2020).
- 19 NHS Digital. ISB1523: anonymisation standard for publishing health and social care data. NHS Digital, 2020. <https://digital.nhs.uk/data-and-information/information-standards/information-standards-and-data-collections-including-extractions/publications-and-notifications/standards-and-collections/isb1523-anonymisation-standard-for-publishing-health-and-social-care-data> (accessed Aug 6, 2020).
- 20 Secretary of State for Health. Coronavirus (COVID-19): notification to organisations to share information. UK Government, 2020. <https://www.gov.uk/government/publications/coronavirus-covid-19-notification-of-data-controllers-to-share-information> (accessed Aug 6, 2020).
- 21 Blanco JL, Ambrosioni J, Garcia F, et al. COVID-19 in patients with HIV: clinical case series. *Lancet HIV* 2020; **7**: e314–16.
- 22 Childs K, Post FA, Norcross C, et al. Hospitalized patients with COVID-19 and human immunodeficiency virus: a case series. *Clin Infect Dis* 2020; **71**: 2021–22.
- 23 Shalev N, Scherer M, LaSota ED, et al. Clinical characteristics and outcomes in people living with human immunodeficiency virus hospitalized for coronavirus disease 2019. *Clin Infect Dis* 2020; published online May 30. <https://doi.org/10.1093/cid/ciaa635>.
- 24 Geretti AM, Stockdale A, Kelly S, et al. Outcomes of COVID-19 related hospitalisation among people with HIV in the ISARIC WHO Clinical Characterisation Protocol (UK): a prospective observational study. *Clin Infect Dis* 2020; published online Oct 23. <https://doi.org/10.1093/cid/ciaa1605>.
- 25 Griffith G, Morris TT, Tudball M, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nat Commun* 2020; **11**: 5749.
- 26 Pan D, Sze S, Minhas JS, et al. The impact of ethnicity on clinical outcomes in COVID-19: a systematic review. *EclinicalMedicine* 2020; **23**: 100404.
- 27 O'Connor J, Smith C, Lampe FC, et al. Durability of viral suppression with first-line antiretroviral therapy in patients with HIV in the UK: an observational cohort study. *Lancet HIV* 2017; **4**: e295–302.
- 28 NHS Improvement. National cost collection 2019. Technical update: sensitive/legally restricted data. NHS Improvement, 2019. [https://improvement.nhs.uk/documents/5259/2019\\_Technical\\_Update\\_FINAL.pdf](https://improvement.nhs.uk/documents/5259/2019_Technical_Update_FINAL.pdf) (accessed Aug 6, 2020).
- 29 Public Health England. HIV in the UK: towards zero HIV transmission by 2030, 2019 report. Public Health England, 2019. <https://www.gov.uk/government/publications/hiv-in-the-united-kingdom> (accessed July 20, 2020).
- 30 Smith RL, de Boer R, Brul S, Budovskaya Y, van Spek H. Premature and accelerated aging: HIV or HAART? *Front Genet* 2013; **3**: 328.
- 31 British HIV Association. Patient confidentiality in sexual health services. British HIV Association, 2015. <https://www.bhiva.org/file/hyWHoZwCBKRMN/Update-for-disemmination-STI-confidentiality-Final-100915.pdf> (accessed Aug 6, 2020).