medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Estrogen and COVID-19 symptoms: associations in women from the COVID Symptom Study

Ricardo Costeira[1]*, Karla A. Lee[1]*, Benjamin Murray[2], Colette Christiansen[1], Juan Castillo-Fernandez[1], Mary Ni Lochlainn[1], Joan Capdevila Pujol[3], Heather Macfarlane[4], Louise C. Kenny[5], Iain Buchan[6], Jonathon Wolf[3], Janice Rymer[7], Sebastien Ourselin[2], Claire J. Steves[1], Timothy D. Spector[1], Louise R. Newson[8], Jordana T. Bell[1]
*contributed equally to this work

[1] Department of Twin Research and Genetic Epidemiology, King's College London, London, U.K.
[2] School of Biomedical Engineering & Imaging Sciences, King's College London, London, U.K
[3] Zoe Global Limited, London, U.K.
[4] National Health Service, U.K.
[5] Department of Women's and Children's Health, University of Liverpool, U.K.
[6] Department of Public Health, Policy and Systems, University of Liverpool, U.K.
[7] Department of Women's Health, King's College London, London, U.K
[8] Newson Health Menopause & Wellbeing Centre, Stratford-Upon-Avon, U.K.

Correspondence:
Dr. Jordana Bell
Department of Twin Research & Genetic Epidemiology
King's College London
St Thomas' Campus
London SE1 7EH
Email: jordana.bell@kcl.ac.uk

Manuscript word count: 4,207

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Abstract

Background: Men and older women have been shown to be at higher risk of adverse COVID-19 outcomes. Animal model studies of SARS-CoV and MERS suggest that the age and sex difference in COVID-19 symptom severity may be due to a protective effect of the female sex hormone estrogen. Females have shown an ability to mount a stronger immune response to a variety of viral infections because of more robust humoral and cellular immune responses.

Objectives: We sought to determine whether COVID-19 positivity increases in women entering menopause. We also aimed to identify whether premenopausal women taking exogenous hormones in the form of the combined oral contraceptive pill (COCP) and post-menopausal women taking hormone replacement therapy (HRT) have lower predicted rates of COVID-19, using our published symptom-based model.

Design: The COVID Symptom Study developed by King's College London and Zoe Global Limited was launched in the UK on 24th March 2020. It captured self-reported information related to COVID-19 symptoms. Data used for this study included records collected between 7th May - 15th June 2020.

Main outcome measures: We investigated links between COVID-19 rates and 1) menopausal status, 2) COCP use and 3) HRT use, using symptom-based *predicted* COVID-19, tested COVID-19, and disease severity based on requirement for hospital attendance or respiratory support.

Participants: Female users of the COVID Symptom Tracker Application in the UK, including 152,637 women for menopause status, 295,689 for COCP use, and 151,193 for HRT use. Analyses were adjusted for age, smoking and BMI.

Results: Post-menopausal women aged 40-60 years had a higher rate of *predicted* COVID (P=0.003) and a corresponding range of symptoms, with consistent, but not significant trends observed for tested COVID-19 and disease severity. Women aged 18-45 years taking COCP had a significantly lower *predicted* COVID-19 (P=8.03E-05), with a reduction in hospital attendance (P=0.023). Post-menopausal women using HRT or hormonal therapies did not exhibit consistent associations, including increased rates of *predicted* COVID-19 (P=2.22E-05) for HRT users alone.

Conclusions: Our findings support a protective effect of estrogen on COVID-19, based on positive association between *predicted* COVID-19 and menopausal status, and a negative association with COCP use. HRT use was positively associated with COVID-19 symptoms; however, the results should be considered with caution due to lack of data on HRT type, route of administration, duration of treatment, and potential comorbidities.

Trial registration: The App Ethics has been approved by KCL ethics Committee REMAS ID 18210, review reference LRS-19/20-18210

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Introduction

As the COVID-19 pandemic progresses, it has been widely observed that adult men of all ages are at higher risk of developing serious complications. A recent review of biological sex and COVID-19 has described the male bias in COVID-19 mortality in 37 of the 38 countries that have provided sex-disaggregated data (1). Of women who develop COVID-19, being post-menopausal has been independently associated with more severe COVID-19 (2). Epidemiological data from previous coronavirus outbreaks, including SARS-CoV (Severe Acute Respiratory Syndrome Corona Virus)
and MERS-CoV (Middle East Respiratory Syndrome Corona Virus) showed the same pattern: among men, morbidity and fatality rates were markedly higher compared to women (3, 4). When patients aged 70 years or older were examined, the effect of lower morbidity and mortality disappeared among SARS-CoV-2 infected women (5). It has also been noted that pregnant women experience milder COVID-19 than initially expected (6).

Animal model studies of SARS-CoV and MERS suggest that the age and sex difference in COVID-19 symptom severity may be due to protective and acute actions of the female sex hormone estrogen (7). Females have been shown to be able to mount a stronger immune response to a variety of viral infections because of more robust humoral and cellular immune responses (8-10). Anti-mu□llerian hormone (AMH) and estradiol are markers of high ovarian reserve and have been shown to negatively correlate with severity of COVID-19, independent of age, suggesting that pre-menopausal women are somewhat protected against more severe COVID-19 (11). A phase II study testing whether short-course estradiol delivered via transdermal patch will be safe and reduce symptom severity in COVID-19 affected adult men and older women is underway in New York (NCT04359329), with another study investigating oral progesterone in men hospitalised with COVID-19 in California (NCT04365127).

The potential protective effect of estrogen against COVID-19 requires continued and careful evaluation. Here, we investigate whether higher levels of estrogen are linked to a reduction in the rate and severity of COVID-19 among women, based on large-scale self-reported data from the UK. The primary analysis explores if women who have recently gone through menopause are at higher risk of symptomatic COVID-19, compared to women who still report periods over age 40. Follow-up analyses consider exogenous estrogen, exploring if women taking the combined oral contraceptive pill (COCP) have a reduced risk of being COVID-19-positive, with associated symptoms. We also assess if use of HRT is associated with a reduced rate of COVID-19 positivity and severity in post-menopausal women alone.

Materials and Methods

The COVID Symptom Study Smartphone Application ("app") was developed by Zoe Global Limited with scientific input from researchers and clinicians at King's College London and Massachusetts General Hospital. It was launched in the UK on 24th March 2020. It captures self-reported information related to COVID-19 symptoms. On first use, the app records self-reported location, age, and core health risk factors. At this point height and weight are self-reported, allowing calculation of body mass index (BMI). With continued use, participants provide daily updates on symptoms, information on health care visits, COVID-19 testing results, and whether they are self-quarantining or seeking healthcare, including the level of intervention and related outcomes. Individuals without apparent symptoms are also encouraged to use the app.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

On 7 May 2020 we asked all female app-users if they are presently taking any forms of hormonal therapies including hormone replacement therapy (HRT), hormonal contraceptives and testosterone (**S1 Fig**). We also posed questions relating to menstruation and current pregnancy. Patients who indicated they were still having periods were asked about frequency of menstruation. Those who indicated not having periods were asked their age at menopause. The COVID Symptom Study dataset used for this study was obtained from the period of 7 May - 15 June 2020, yielding 40 days of data collection from a total of 1.9M women in the UK. From these, 1.6M women had BMI between 20-35kg/m^2 and were included in downstream analyses.

Ascertainment of exposures, disease outcomes and study covariates

Exposures, outcomes and covariates were ascertained from self-reported app data following quality control with purpose-built scripts (https://github.com/KCL-BMEIS/zoe-data-prep). Exposures used in our analyses included women's menopausal status, and COCP and HRT use. Primary disease outcomes included COVID-19-related symptoms (**S1 Table**) and *predicted* COVID-19 positivity based on symptoms, as described in our recent publication (12). Hospitalisation and respiratory support, defined as supplementary oxygen (+/- ventilation) were used as surrogate markers for disease severity. Self-reported results from nose/throat swab tests were also used as outcome for a subset of the sample who were tested for COVID-19. Reported age, BMI and smoking status were used as covariates in the associations, and the analyses only included female app users.

Analyses of menopausal status included post- and pre-menopausal women aged 40-60 years, with BMI 20–35 kg/m^2, excluding women taking any form of hormonal therapy. We compared post-menopausal women currently reporting no periods and with last period reported after the age of 40 and within 5 years, to pre-menopausal women with regular periods occurring every 3–6 weeks. Analyses of COCP use included post- and pre-menopausal women aged 20-45 years, with BMI 20–35kg/m^2. We compared women taking COCP as their only form of hormonal therapy, to women of the same age taking no form of hormone therapy. Analyses of HRT use were carried out in post-menopausal women aged 50-65 years, with BMI 20–35kg/m^2 and last periods reported at age 45–60. We compared women on HRT alone to women receiving no form of hormonal therapy. Extended hormone therapy analyses also considered women who were on HRT or related hormone therapies, including COCP, progestogen therapy, progestogen containing intrauterine systems, or testosterone (**S1 Table**). Use of estrogen for gender transitioning was excluded from the analyses.

Association analyses

Binomial generalized mixed models with a log-odds/logit link function were used to carry out association analyses. The models included COVID-19 symptom or outcome as a function of exposure variables and covariates including age, BMI, and smoking status. Exposures included menopausal status, COCP use, and HRT use. Exposure data and symptoms of the disease were coded as '1' for positive (TRUE/yes/severe/significant) responses and '0' for negative (FALSE/no/mild) responses or blank (NA) statements from app users. In *predicted* COVID-19 analyses subjects with a *predicted* COVID-19 probability ≥ 50% were considered COVID-19-positive (coded '1', while a predicted COVID-19-negative outcome was coded as '0'). App users in hospital or home from hospital were coded as '1' for hospitalisation. App users who reported going to the hospital and requiring supplementary oxygen (+/- ventilation) were coded as '1' for respiratory support, while all others were coded as '0' for this outcome. App users with at least one positive swab test for COVID-19 were considered COVID-19-positive in the *tested* COVID-19-positive analyses (coded as '1', while a tested COVID-19-negative outcome was coded as '0'). Only a small subset of women was tested for COVID-19, and missing data, failed tests, or tests awaiting outcome were excluded from analyses. Age and BMI were coded as continuous fixed effects. Only women with BMI 20–35kg/m^2 were considered in our study. Smoking was coded as a categorical fixed effect

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

variable with levels '1', '2' and '3' respectively representing never-smokers, ex-smokers and current smokers. Odds ratios of the associations are reported.

Sensitivity analyses

Age is a key risk factor for COVID-19. Age sensitivity analyses were performed to match the mean and median ages of cases and controls for each of the three exposure variables. Analyses were carried out in subsets of app users within 5-year bins for menopausal status and use of COCP and HRT. In menopause, sensitivity analysis was also performed for years passed since time of last period.

TwinsUK cohort data analysis

At the time of data freeze, a subset of app users included research volunteers from the TwinsUK cohort (13). Here, we considered 270 TwinsUK female twins (mean age 66), who also had existing whole blood DNA methylation data profiled using the Illumina Infinium HumanMethylation450 BeadChip. From the 270 female twins, a subset of 84 participants had previously collected TwinsUK questionnaire data on menopausal status, and a further subset of 75 had information on age at last period.

Whole blood DNA extraction and DNA methylation profiling in these samples have been previously described (14). Briefly, DNA methylation levels were determined using Illumina methylation beta-values (15), which range between 0 at unmethylated CpG-sites, and 1 at fully methylated CpG sites. ENmix (16) was used for quality control and minfi (17) was used to exclude samples with median methylated and unmethylated signals below 10.5. Three different epigenetic age calculators were applied to estimate DNA methylation age for each individual, DNAm GrimAge (18), DNAm PhenoAge (19) and the original Horvath methylation age (20). Epigenetic age acceleration measures included estimates obtained from regressing epigenetic age on chronological age. We also considered intrinsic epigenetic age acceleration (IEAA) and extrinsic epigenetic age acceleration (EEAA) (21). IEAA is calculated by regressing the Horvath DNA methylation age and cell blood counts to create an estimate of cell intrinsic methylation ageing, independent of differences in blood cell counts. EEAA is calculated based on epigenetic ageing measures developed by Hannum et al. (22), up-weighted by the relative proportion of three age-related blood immune cells, specifically naïve (CD45RA+CCR7+) cytotoxic T cells, exhausted (CD28-CD45RA-) cytotoxic T cells, and plasmablasts, as previously described (21). EEAA has been proposed to capture aspects of age-related immuno-senescence (21). A one-sided t-test was applied to compare each of the five age acceleration measures to symptoms and *predicted* COVID-19 status, and results are presented at nominal significance.

Results

The COVID Symptom Study dataset used for this study was collected between 7 May - 15 June 2020 and included 1.6M women in the UK (BMI range 20–35kg/m^2), with all analyses adjusted for age, BMI and smoking status.

Menopause

We examined the impact of entering menopause on COVID-19-positivity and related symptoms among 152,637 women aged 40–60 years with BMI 20-35 kg/m^2. Cases were defined as post-menopausal women currently reporting no periods and with last period reported after the age of 40 and within the last 5 years, resulting in altogether 44,268 women (**S1 Table**). Controls were defined as pre-menopausal women with regular periods occurring every 3–6 weeks, resulting in altogether 108,369 controls. Women taking any form of hormonal therapy were excluded from this analysis.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Post-menopausal women had a higher rate of *predicted* COVID-19 (OR = 1.22, 95% CI 1.07–1.39, p = 0.003) and a corresponding range of significant differences in symptoms including hoarse voice, skipped meals, muscle pains, and fever (**Table 1**). Requirement for hospitalisation and respiratory support were not significant, but also showed a positive direction of association in post-menopausal women. Although there was no significant association between menopausal status and testing COVID-19-positive, the direction of association is consistent with *predicted* COVID-19 results (**Table 1**).

Table 1. Association of menopausal status to *predicted* and tested COVID-19 results, self-reported symptoms and indicators of COVID-19 disease severity (hospitalisation/respiratory support) (n = 152,637).

	Beta	**SE**	***p*-value**	**OR**	**95% CI**	
					LL	**UL**
Self-reported symptoms						
Abdominal pain (**)	0.117	0.041	0.005	1.12	1.04	1.22
Chest pain	0.020	0.041	0.631	1.02	0.94	1.11
Delirium	0.114	0.058	0.051	1.12	1.00	1.26
Diarrhoea (**)	0.135	0.042	0.001	1.15	1.06	1.24
Fever (*)	0.119	0.051	0.020	1.13	1.02	1.25
Headache	0.026	0.027	0.337	1.03	0.97	1.08
Hoarse voice (***)	0.207	0.051	4.73E-05	1.23	1.11	1.36
Loss of smell (anosmia)	0.064	0.056	0.250	1.07	0.96	1.19
Persistent cough	0.054	0.046	0.236	1.06	0.97	1.15
Severe fatigue	0.047	0.064	0.467	1.05	0.92	1.19
Severe or significant shortness of breath	0.007	0.106	0.951	1.01	0.82	1.24
Skipped meals (**)	0.128	0.047	0.007	1.14	1.04	1.25
Sore throat	0.038	0.030	0.214	1.04	0.98	1.10
Unusual muscle pains (***)	0.232	0.043	5.91E-08	1.26	1.16	1.37
Hospitalisation	0.043	0.139	0.758	1.04	0.80	1.37
Respiratory support	0.468	0.539	0.386	1.60	0.55	4.59
***Predicted* COVID-19 (**)**	0.198	0.067	0.003	1.22	1.07	1.39
Tested COVID-19^	0.013	0.148	0.929	1.01	0.76	1.35

*: $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

^: in a subset of COVID-19-tested individuals (n = 14,796).
SE: standard error; OR: odds ratio; 95% CI (LL; UL): upper and lower limits for the 95% confidence interval of the odds ratio.

The mean age of post-menopausal and pre-menopausal women included in the analysis above was 53.8 and 45.2 years, respectively. Because of this difference, sensitivity analyses were performed for age within 5-year age bins (40–45, 45–50, 50–55 and 55–60 years old). Upon subgroup analysis by age, we observed that *predicted* COVID-19 results were most driven by the 45–50 age group (OR = 1.35, 95% CI 1.05–1.72, p = 0.017), wherein anosmia, as well as fever and persistent cough, and the need for oxygen treatment in hospital were also significant (**S2 Table**). A sensitivity analysis was also carried for last period within 3 years at time of reporting. Sensitivity analysis for last menstrual period within 3 years from questionnaire showed similar significant results for higher rate of *predicted* COVID-19 in post-menopausal women (OR = 1.17, 95% CI 1.01–1.36, p = 0.036).

Use of Combined Oral Contraceptive Pill

We examined the link between COCP use and COVID-19-positivity and related outcomes in 295,689 women aged 18–45 years (BMI 20-35). Both pre- and post-menopausal women were considered for this analysis, where most participants (85%) were pre-menopausal. Cases were defined as females on the COCP as their only form of hormonal therapy, resulting in 64,253 COCP-users (**S1 Table)**. Controls were women of the same age and BMI group taking no form of hormone therapy, resulting in 231,436 controls (**S1 Table**).

Women using COCP had a lower rate of *predicted* COVID-19 (OR = 0.87, 95% CI 0.81–0.93, p = 8.03E-05) and a corresponding reduced frequency of symptoms, including persistent cough, delirium, anosmia, skipped meals, severe fatigue and pain (p < 0.001) (**Table 2**). Rate of hospitalisation was also significantly lower in COCP users (OR = 0.79, 95% CI 0.64–0.97, p = 0.023). The association between COCP use and tested COVID-19 as outcome was not significant but showed a consistent negative direction of association in the 25–30 and 35–40 age groups (**S2 Table**). Analyses in the subset of 251,786 pre-menopausal women alone, including 52,453 COCP-users and 199,333 controls, showed consistent strong negative associations with *predicted* COVID-19 (OR = 0.87, 95% CI 0.80–0.95, p = 8.38E-04) and a corresponding reduced frequency of a wide range of symptoms (**S2 Table**).

Table 2. Association of COCP use to *predicted* and tested COVID-19 results, self-reported symptoms, and indicators of COVID-19 disease severity (hospitalisation/respiratory support) (n = 295,689).

	Beta	**SE**	***p*-value**	**OR**	**95% CI**	
					LL	**UL**
Self-reported symptoms						
Abdominal pain (**)	-0.072	0.024	0.002	0.93	0.89	0.97
Chest pain (***)	-0.119	0.025	2.29E-06	0.89	0.85	0.93
Delirium (***)	-0.174	0.033	1.13E-07	0.84	0.79	0.90
Diarrhoea	-0.042	0.025	0.092	0.96	0.91	1.01
Fever (*)	-0.070	0.029	0.016	0.93	0.88	0.99
Headache	-0.016	0.016	0.325	0.98	0.95	1.02

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Hoarse voice (***)	-0.119	0.033	3.34E-04	0.89	0.83	0.95
Loss of smell (anosmia) (***)	-0.140	0.036	1.17E-04	0.87	0.81	0.93
Persistent cough (***)	-0.094	0.028	7.03E-04	0.91	0.86	0.96
Severe fatigue (***)	-0.145	0.040	2.57E-04	0.87	0.80	0.94
Severe or significant shortness of breath	-0.106	0.064	0.095	0.90	0.79	1.02
Skipped meals (***)	-0.205	0.027	4.50E-14	0.81	0.77	0.86
Sore throat (***)	-0.078	0.018	1.63E-05	0.93	0.89	0.96
Unusual muscle pains (***)	-0.152	0.028	3.75E-08	0.86	0.81	0.91
Hospitalisation (*)	-0.239	0.105	0.023	0.79	0.64	0.97
Respiratory support	-0.811	0.480	0.091	0.44	0.17	1.14
***Predicted* COVID-19 (***)**	-0.144	0.037	8.03E-05	0.87	0.81	0.93
Tested COVID-19^	0.076	0.086	0.380	1.08	0.91	1.28

*: $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.
^: in a subset of Covid-19-tested individuals (n = 26,871).
SE: standard error; OR: odds ratio; 95% CI (LL; UL): upper and lower limits for the 95% confidence interval of the odds ratio.

The mean ages of COCP users and controls were 29.5 and 34.2 years, respectively. Sensitivity analyses were performed for age, in 5-year bins (18–25, 25–30, 30–35, 35–40 and 40–45-year-olds). Sensitivity analyses showed strongest results for the 25–30 year age group, with COCP-users having lower *predicted* COVID-19-positivity *(*OR = 0.77, 95% CI 0.67–0.90*,* $p = 0.0008$) and avoiding hospitalisation (OR = 0.5, 95% CI 0.31–0.80, $p = 0.004$) (**S2 Table**). Negative association with *predicted* COVID-19 was also significant in the 40–45-year-olds (OR = 0.77, 95% CI 0.63–0.94, $p = 0.01$). Associations in the 18–25-year-olds yielded only two statistically significant symptom associations, suggesting that this age group contributed the least to the results observed in the main analysis.

Use of Hormone Replacement Therapy

The association between HRT use and COVID-19 was assessed in 151,193 post-menopausal women aged 50-65 years, with BMI 20–35 and last periods reported at age 45–60. Controls were post-menopausal women matching these criteria who received no form of hormonal therapy, resulting in 133,395 controls (**S1 Table**). Cases were defined as 17,798 women on HRT only. Extended analyses also considered women on HRT or additional related hormone therapies, excluding use of estrogen in gender transitioning (**S1 Table**).

HRT use was associated with an increased rate of *predicted* COVID-19 (OR = 1.32, 95% CI = 1.16 – 1.49, p = 2.22E-05) and frequency of a wide range of symptoms (**S3 Table**). However, while *predicted* COVID-19 and reporting of symptoms showed positive associations in HRT users, there was no significantly increased rate of hospitalisation in HRT-users. Notably, both the need for respiratory support and testing positive for COVID-19

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

showed a negative trend of association with HRT use, although the results were not nominally significant. The results remained consistent in extended analyses considering use of HRT or other related hormone therapies (**S3 Table**).

The mean ages of HRT users ranged between 56.6 – 56.8 years, while the mean age of controls was 58.2 years. Sensitivity analyses were performed for age, selecting for subgroups of women aged 50–55, 55–60 and 60–65 years. Sensitivity analyses of age were consistent with the overall *predicted* COVID-19 and symptoms results and showed negative directions of association for testing COVID-19-positive, most consistent in 55–60-year-olds (**S2 Table**), which yielded more COVID-19 tests than the 50–55 and 60–65-year-old subgroups (**S1 Table**).

App data validation in TwinsUK questionnaires

A subset of app users included 270 female research volunteers from the TwinsUK cohort (13). Of these, 84 had previously reported questionnaire data on menopausal status. For all 84 women menopausal status in the app response matched the TwinsUK questionnaire data, where women reported that periods had either stopped or that they did not currently have periods. Furthermore, a subset of 75 female twins had reported information on age of last period in TwinsUK questionnaires. Of these, 64% of twins (48 twins) matched age of last period reported from TwinsUK data within 1 year to age at last period reported in the app, and 87% (65) matched within a 3-year range.

Menopause, biological aging, and COVID-19 symptoms

Menopause is a marker of ageing and has previously been linked to accelerated epigenetic ageing (23). To this end, we compared the frequency of COVID-19 symptoms in 270 TwinsUK female twins (**S4 Table**) with available app data to 5 estimates of epigenetic ageing in whole blood, including the original epigenetic age acceleration, GrimAge acceleration, PhenoAge acceleration, blood cell *intrinsic* epigenetic age acceleration (IEAA), and blood *extrinsic* epigenetic age acceleration (EEAA).

Overall, fatigue and unusual muscle pains showed the most (3 or more) nominally significant associations with epigenetic age acceleration measures; followed by hoarse voice, skipped meals and anosmia where significant differences were observed for two age acceleration measures; and fever where a significant difference was observed with GrimAge Acceleration alone (p = 0.01) (**S5 Table**). However, the results should be interpreted with caution as sample sizes for subgroup analyses are modest, and in some cases extremely small. Similarly, the number of individuals with *predicted* COVID-19 was extremely small (3 *predicted* cases), but these individuals as a group had on average accelerated epigenetic age acceleration across all five epigenetic ageing measures (**S5 Table**).

Discussion

Sex is a biological variable that affects immune responses to both self and foreign antigens. The sex of an individual is a multidimensional biological characteristic that shapes infectious disease pathogenesis and is defined by the differential organisation of chromosomes, reproductive organs, and sex steroid levels. Sex is distinct from gender, which includes behaviours and activities that are determined by society or culture in humans. Human biological sex plays a fundamental role in heterogeneous COVID-19 outcomes, with a strong male predominance in mortality. Although gender-related social factors, including smoking, health care-seeking behaviours and some medical comorbidities may impact the outcomes of COVID-19 (24, 25) and contribute to sex-based differences in severity, the cross-cultural

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

emergence of significantly increased risk of mortality for males points to biological risk determinants.

Teasing out the precise drivers of mortality in COVID-19, regardless of sex, is a difficult task. The innate recognition and response to viruses as well as downstream adaptive immune responses during viral infections are known to differ between females and males (26). It has been well-illustrated that females generally mount greater inflammatory, antiviral, and humoral immune responses than males during viral infections (27), which contributes to better clearance of viruses, including SARS-CoV (7). This heightened inflammatory response is advantageous in response to infection and sepsis, but is unfavourable in immune responses against self, leading to an overall increased rate of autoimmune diseases in women compared to men (28, 29). Additionally, enhanced immunity in females can also result in greater immuno-pathology and tissue damage at later stages of viral disease, such as during influenza A virus infection (30). Conversely, maternal physiological adaptations to pregnancy usually predispose pregnant women to a more severe course of many infections, including viral pneumonia, with subsequent higher maternal and fetal morbidity and mortality (31), but observational cohort studies in COVID-19 have reported the risk of severe disease in the pregnant population was approximately half that in the general population of patients presenting with this disease (6). Estrogen levels increase more than 100-fold in pregnancy (32) providing a potential mechanism for the unique resistance to COVID-19 compared with previous pandemic such as H1N1.

With ageing, a general decline in immune function is observed – immune-senescence. Several of these changes are gender specific and affect postmenopausal women. Menopause is a normal part of a woman's lifecycle and consists of a series of body changes that can last from one to ten years. Levels of estrogen, for example, 17β-estradiol (E2), are variable during the menstrual cycle, high during pregnancy and low after menopause in females. E2 affects many components of innate immunity, including the functional activity of innate immune cells that influence downstream adaptive immune responses (26). Loss of sex hormones due to ageing results in a reduction of immune function. For example, in postmenopausal women, a second peak in Human papilloma virus (HPV) prevalence has been reported (33, 34). New HPV infections in older women with no sexual activity are thought to be due to reduced immune responses (35). HIV-1 infection is also increasing in postmenopausal women (36), where a European study found that women over 45 have a 4-fold increased risk of acquiring HIV compared to women under 45 years of age (37).

As we hypothesised, our results show that being pre-menopausal appears to have a protective effect against COVID-19 in a large community survey of female UK app-users. This was supported by a protective effect seen amongst pre-menopausal women taking the COCP but was not seen for post-menopausal women taking HRT. However, HRT results should be considered with caution due to lack of data on HRT type, route of administration and duration of treatment. Historically, HRT provision and prescribing has been poor in the UK and elsewhere, representing a significant area-of-need in women's health. There are different effects on various preparations of HRT, and we did not determine the type of HRT that each woman was taking. The bulk of scientific evidence from preclinical, clinical, and epidemiologic studies, and randomised clinical trials clearly indicates that judiciously selected HRT is usually beneficial and rarely dangerous. The majority of women in the UK are currently given oral estrogen, which has more risks compared to transdermal estrogen, and may also affect immunity differently (38). Transdermal estrogen contains E2, which has more beneficial effects on immunity as well as future health (39). Women taking HRT have been shown to have lower future risk of cardiovascular disease, obesity and type 2 diabetes, which are all known to be associated with more severe symptoms and higher mortality from COVID-19 (40).

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Menopause is a marker of biological ageing in women that has previously been associated with accelerated epigenetic ageing (21). The associations between menopausal status and COVID-19 positivity and symptom severity may in part be related to biological ageing, rather than reduction in estrogen specifically, although the COCP use results suggest that this is unlikely. To explore the possibility that our results may be attributed to ageing, rather than a reduction in estrogen, we also tested the association between COVID-19 symptoms and epigenetic ageing rates in a subset of participants from the TwinsUK cohorts. Our results are consistent with increased frequency of COVID-19 symptoms among subjects with accelerated biological ageing. However, the sample sizes in the epigenetic subgroup analyses is too modest to draw robust conclusions. Interestingly, despite the extremely small sample size, the group of 3 *predicted* COVID-19 individuals showed significantly greater epigenetic age acceleration, including with epigenetic measures that capture aspects of immune-senescence (EEAA), warranting further investigation into the link between biological ageing and COVID-19 positivity and symptom severity.

This study collected data on hundreds of thousands of women resulting in good power to detect effects. However, our study also has several limitations. Data are self-reported, and questions on medication use were non-standard, to ease large-scale app-based reporting. Data on type, route, duration, and dose of hormone therapies, and importantly HRT, were not collected due to difficulties faced collecting very detailed data using an app-based interface. As such, untangling the effect of differing types of HRT was not possible. Most COCPs contain between 20-35 micrograms of ethinylestradiol along with a progestogen, while HRT estradiol doses are generally lower and more physiological. As such, lower estrogen doses and lack of detailed data may have resulted in the lack of effect seen amongst HRT-users. Another limitation relates to reporting bias within both symptoms and test results. Additionally, sampling using an app will under-represent individuals without smartphones, including older participants, and is likely to under-represent those severely affected by COVID-19. Other limitations include the effects of unmeasured confounding and systematic differences between individuals prescribed different types of hormone therapy. There is also potential for selection bias where, for example, experiencing gynaecological problems and menopause may influence likelihood of hormonal therapy use. To ascertain the consistency of our results, we performed multiple cumulative data extracts over the period 7 May 2020 - 15 June 2020, and observed that association results were consistent throughout, including specifically for *predicted* COVID-19 associations. Nonetheless, because prediction of COVID-19 was based on symptoms reported over a two and half-month period since the app launched, predictions could in some cases be influenced by the reporting of symptoms at far-apart dates. Additionally, the individuals on which the prediction model was trained were highly selected as COVID-19 testing was not performed at random when the app was initially launched, although testing criteria have since been extended and more individuals have been able to access COVID-19 testing.

Conclusion

Our findings indicate a protective effect of estrogen from symptomatic COVID-19, based on positive association of menopausal status with *predicted* COVID-19, and negative association of COCP use with *predicted* COVID-19. HRT use was positively associated with COVID-19 symptoms. However, the results should be interpreted with caution due to lack of data on HRT type, route of administration, duration of treatment, and potential comorbidities. Further work focussed on gender with hormone profiling in both pre-clinical and clinical settings, as well as on biological ageing, is needed to uncover novel features of the host immune response to SARS-CoV-2 and ultimately result in more equitable health outcomes.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Acknowledgments:
We thank all participants who entered data into the COVID Symptom Study app, including study volunteers enrolled in the TwinsUK cohort. We thank the staff of the Department of Twin Research at King's College London and Zoe Global Limited for contributing to the running of the study and data collection.

Ethics:
The App Ethics has been approved by KCL ethics Committee REMAS ID 18210, review reference LRS-19/20-18210 and all subscribers provided consent.

Data sharing:
Data used in this study is available to bona fide researchers through UK Health Data Research using the following link https://healthdatagateway.org/detail/9b604483-9cdc-41b2-b82c-14ee3dd705f6

Declaration of Interests:
Zoe Global Limited co-developed the app *pro bono* for non-commercial purposes. Investigators received support from the Wellcome Trust, the MRC/BHF, EU, NIHR, CDRF, and the NIHR-funded BioResource, Clinical Research Facility and BRC based at GSTT NHS Foundation Trust in partnership with KCL. Zoe authors work for Zoe Global Limited and TDS is a consultant to Zoe Global Limited. IB is NIHR Senior Investigator and Chief Data Scientist Advisor for AstraZeneca.

Funding:
This study received no specific funding. However, Zoe Global Limited provided in kind support for all aspects of building, running, and supporting the app and service to all users worldwide. In addition, epigenetic ageing estimates were previously generated with funding from the UK Economic and Social Research Council (ES/N000404/1 to JTB), but the funder had no role in in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Investigators received support from the Wellcome Trust, the MRC/BHF, Alzheimer's Society, EU, NIHR, CDRF. The TwinsUK study was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007–2013); National Institute for Health Research (NIHR)- funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London.

Contribution Statement:
LRN, TDS, CJS and JTB developed the research hypotheses. RC, KAL, LRN and JTB wrote the manuscript. RC performed the primary analyses. BM, MNL, HM, CC, JC-F, JCP, IB, LK, JR and SO contributed to the analyses and interpretation of the results. JCP and JW facilitated the addition of the relevant questions to the app. All authors approved the final manuscript.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Figures

Are you currently having periods? [I've never had periods/I'm currently having periods/I've stopped having periods/I'm pregnant/Prefer not to say]

- Are you taking any of the following forms of hormone treatment? [Choose all that apply]
  - No
  - Combined oral contraceptive pill
  - Progesterone only pill
  - Mirena or other hormone coil
  - Depot injection or implant
  - Hormone Replacement Therapy
  - Estrogen hormone therapy for gender transitioning
  - Testosterone hormone therapy
  - Prefer not to say
- Follow on questions:
  - If "I'm currently having periods":
    - Do your periods usually occur?
    - Regularly every 3-6 weeks
    - Regularly, but less often than every 6 weeks
    - At irregular intervals
  - If "I'm pregnant":
    - How many weeks pregnant are you?
  - If "I've stopped having periods":
    - At what age did your periods stop?

Supplementary Figure 1. Questions posed to female participants with possible answers

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

References:

1. Scully EP, Haverfield J, Ursin RL, Tannenbaum C, Klein SL. Considering how biological sex impacts immune responses and COVID-19 outcomes. Nature Reviews Immunology. 2020;20(7):442-7.
2. Ding T, Zhang J, Wang T, Cui P, Chen Z, Jiang J, et al. A Multi-hospital Study in Wuhan, China：Protective Effects of Non-menopause and Female Hormones on SARS-CoV-2 infection. medRxiv. 2020:2020.03.26.20043943.
3. Karlberg J, Chong DS, Lai WY. Do men have a higher case fatality rate of severe acute respiratory syndrome than women do? Am J Epidemiol. 2004;159(3):229-31.
4. Leong HN, Earnest A, Lim HH, Chin CF, Tan C, Puhaindran ME, et al. SARS in Singapore--predictors of disease severity. Ann Acad Med Singapore. 2006;35(5):326-31.
5. Shaw AC, Goldstein DR, Montgomery RR. Age-dependent dysregulation of innate immunity. Nat Rev Immunol. 2013;13(12):875-87.
6. Chen L, Li Q, Zheng D, Jiang H, Wei Y, Zou L, et al. Clinical Characteristics of Pregnant Women with Covid-19 in Wuhan, China. New England Journal of Medicine. 2020;382(25):e100.
7. Channappanavar R, Fett C, Mack M, Ten Eyck PP, Meyerholz DK, Perlman S. Sex-Based Differences in Susceptibility to Severe Acute Respiratory Syndrome Coronavirus Infection. J Immunol. 2017;198(10):4046-53.
8. Hannah MF, Bajic VB, Klein SL. Sex differences in the recognition of and innate antiviral responses to Seoul virus in Norway rats. Brain Behav Immun. 2008;22(4):503-16.
9. Klein SL, Flanagan KL. Sex differences in immune responses. Nature Reviews Immunology. 2016;16(10):626-38.
10. Robinson DP, Lorenzo ME, Jian W, Klein SL. Elevated 17beta-estradiol protects females from influenza A virus pathogenesis by suppressing inflammatory responses. PLoS Pathog. 2011;7(7):e1002149.
11. Ding T, Zhang J, Wang T, Cui P, Chen Z, Jiang J, et al. Potential Influence of Menstrual Status and Sex Hormones on female SARS-CoV-2 Infection: A Cross-sectional Study from Multicentre in Wuhan, China. Clinical Infectious Diseases. 2020.
12. Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. Nature Medicine. 2020.
13. [Available from: https://twinsuk.ac.uk.
14. Kurushima Y, Tsai P-C, Castillo-Fernandez J, Couto Alves A, El-Sayed Moustafa JS, Le Roy C, et al. Epigenetic findings in periodontitis in UK twins: a cross-sectional study. Clin Epigenetics. 2019;11(1):27-.
15. Du P, Zhang X, Huang C-C, Jafari N, Kibbe WA, Hou L, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics. 2010;11(1):587.
16. Xu Z, Niu L, Li L, Taylor JA. ENmix: a novel background correction method for Illumina HumanMethylation450 BeadChip. Nucleic acids research. 2016;44(3):e20-e.
17. Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics (Oxford, England). 2014;30(10):1363-9.
18. Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging (Albany NY). 2019;11(2):303-27.
19. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging (Albany NY). 2018;10(4):573-91.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

20. Horvath S. DNA methylation age of human tissues and cell types. Genome Biology. 2013;14(10):3156.
21. Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai P-C, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. Aging (Albany NY). 2016;8(9):1844-65.
22. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell. 2013;49(2):359-67.
23. Levine ME, Lu AT, Chen BH, Hernandez DG, Singleton AB, Ferrucci L, et al. Menopause accelerates biological aging. Proceedings of the National Academy of Sciences. 2016;113(33):9327-32.
24. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet. 2020;395(10223):507-13.
25. Meng Y, Wu P, Lu W, Liu K, Ma K, Huang L, et al. Sex-specific clinical characteristics and prognosis of coronavirus disease-19 infection in Wuhan, China: A retrospective study of 168 severe patients. PLoS Pathog. 2020;16(4):e1008520.
26. Klein SL, Flanagan KL. Sex differences in immune responses. Nat Rev Immunol. 2016;16(10):626-38.
27. vom Steeg LG, Klein SL. SeXX Matters in Infectious Disease Pathogenesis. PLoS Pathog. 2016;12(2):e1005374.
28. Olsen NJ, Kovacs WJ. Gonadal steroids and immunity. Endocr Rev. 1996;17(4):369-84.
29. Talal N. Sex steroid hormones and systemic lupus erythematosus. Arthritis Rheum. 1981;24(8):1054-6.
30. Vermillion MS, Ursin RL, Kuok DIT, Vom Steeg LG, Wohlgemuth N, Hall OJ, et al. Production of amphiregulin and recovery from influenza is greater in males than females. Biol Sex Differ. 2018;9(1):24.
31. Chen Y-H, Keller J, Wang IT, Lin C-C, Lin H-C. Pneumonia and pregnancy outcomes: a nationwide population-based study. Am J Obstet Gynecol. 2012;207(4):288.e1-.e2887.
32. Abbassi-Ghanavati M, Greer LG, Cunningham FG. Pregnancy and laboratory studies: a reference table for clinicians. Obstet Gynecol. 2009;114(6):1326-31.
33. Franceschi S, Herrero R, Clifford GM, Snijders PJ, Arslan A, Anh PT, et al. Variations in the age-specific curves of human papillomavirus prevalence in women worldwide. International journal of cancer. 2006;119(11):2677-84.
34. Smith JS, Melendy A, Rana RK, Pimenta JM. Age-specific prevalence of infection with human papillomavirus in females: a global review. J Adolesc Health. 2008;43(4 Suppl):S5-25, S.e1-41.
35. González P, Hildesheim A, Rodríguez AC, Schiffman M, Porras C, Wacholder S, et al. Behavioral/lifestyle and immunologic factors associated with HPV infection among women older than 45 years. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2010;19(12):3044-54.
36. Meditz AL, Moreau KL, MaWhinney S, Gozansky WS, Melander K, Kohrt WM, et al. CCR5 expression is elevated on endocervical CD4+ T cells in healthy postmenopausal women. J Acquir Immune Defic Syndr. 2012;59(3):221-8.

medRxiv preprint doi: https://doi.org/10.1101/2020.07.30.20164921; this version posted August 2, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

37. Comparison of female to male and male to female transmission of HIV in 563 stable couples. European Study Group on Heterosexual Transmission of HIV. Bmj. 1992;304(6830):809-13.
38. L'Hermite M. Bioidentical menopausal hormone therapy: registered hormones (non-oral estradiol±progesterone) are optimal. Climacteric. 2017;20(4):331-8.
39. Trenti A, Tedesco S, Boscaro C, Trevisi L, Bolego C, Cignarella A. Estrogen, Angiogenesis, Immunity and Cell Metabolism: Solving the Puzzle. International journal of molecular sciences. 2018;19(3):859.
40. Davey DA. Menopausal hormone therapy: a better and safer future. Climacteric. 2018;21(5):454-61.