Brain imaging before and after COVID-19 in UK Biobank

Gwenaëlle Douaud[1✉], Soojin Lee[1], Fidel Alfaro-Almagro[1], Christoph Arthofer[1], Chaoyue Wang[1], Frederik Lange[1], Jesper L.R. Andersson[1], Ludovica Griffanti[1,2], Eugene Duff[1,3], Saad Jbabdi[1], Bernd Taschler[1], Anderson Winkler[4], Thomas E. Nichols[5], Rory Collins[6], Paul M. Matthews[7], Naomi Allen[6], Karla L. Miller[1], Stephen M. Smith[1]

[1]FMRIB Centre, Wellcome Centre for Integrative Neuroimaging (WIN), Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

[2]OHBA, Wellcome Centre for Integrative Neuroimaging (WIN), Nuffield Department of Clinical Neurosciences, University of Oxford, Oxford, UK

[3]Department of Paediatrics, University of Oxford, Oxford, UK

[4]National Institutes of Mental Health, National Institutes of Health, Bethesda, MD, USA

[5]Big Data Institute, University of Oxford, Oxford, UK

[6]Nuffield Department of Population Health, University of Oxford, Oxford, UK

[7]UK Dementia Research Institute and Department of Brain Sciences, Imperial College, London, UK

✉ Correspondence to: Prof. Gwenaëlle Douaud

FMRIB Centre, Wellcome Centre for Integrative Neuroimaging,

Nuffield Department of Clinical Neurosciences,

John Radcliffe Hospital, Oxford OX3 9DU, UK

+44 1865 610 478

gwenaelle.douaud@ndcn.ox.ac.uk

Abstract

There is strong evidence for brain-related pathologies in COVID-19, some of which could be a consequence of viral neurotropism. The vast majority of brain imaging studies so far have focused on qualitative, gross pathology of moderate to severe cases, often carried out on hospitalised patients. It remains unknown however whether the impact of COVID-19 can be detected in milder cases, in a quantitative and automated manner, and whether this can reveal a possible mechanism for the spread of the disease. UK Biobank scanned over 40,000 participants before the start of the COVID-19 pandemic, making it possible to invite back in 2021 hundreds of previously-imaged participants for a second imaging visit. Here, we studied the effects of the disease in the brain using multimodal data from 782 participants from the UK Biobank COVID-19 re-imaging study, with 394 participants having tested positive for SARS-CoV-2 infection between their two scans. We used structural and functional brain scans from before and after infection, to compare longitudinal brain changes between these 394 COVID-19 patients and 388 controls who were matched for age, sex, ethnicity and interval between scans. We identified significant effects of COVID-19 in the brain with a loss of grey matter in the left parahippocampal gyrus, the left lateral orbitofrontal cortex and the left insula. When looking over the entire cortical surface, these results extended to the anterior cingulate cortex, supramarginal gyrus and temporal pole. We further compared COVID-19 patients who had been hospitalised (n=15) with those who had not (n=379), and while results were not significant, we found comparatively similar findings to the COVID-19 vs control group comparison, with, in addition, a greater loss of grey matter in the cingulate cortex, central nucleus of the amygdala and hippocampal cornu ammonis (all $|Z|>3$). Our findings thus consistently relate to loss of grey matter in limbic cortical areas directly linked to the primary olfactory and gustatory system. Unlike in *post hoc* disease studies, the availability of pre-infection imaging data helps avoid the danger of pre-existing risk factors or clinical conditions being mis-interpreted as disease effects. Since a possible entry point of the virus to the central nervous system might be via the olfactory mucosa and the olfactory bulb, these brain imaging results might be the *in vivo* hallmark of the spread of the disease (or the virus itself) via olfactory and gustatory pathways.

Introduction

While the global pandemic of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has now claimed millions of lives across the world, there has been increased focus by the scientific and medical community on the effects of mild-to-moderate COVID-19 in the longer term. There is strong evidence for brain-related pathologies, some of which could be a consequence of viral neurotropism[1,2]: neurological and cognitive deficits demonstrated by patients (with an incidence of neurological symptoms in more than 80% of the severe cases[3]), radiological and *post mortem* tissue analyses demonstrating the impact of COVID-19 on the brain[4] (doi.org/10.3389/fnagi.2021.646908), and the presence of virus itself in the central nervous system[5-7].

In particular, one consistent clinical feature, which can appear before the onset of respiratory symptoms, is the disturbance in olfaction and gustation in COVID-19 patients[8,9]. In a recent study, 100% of the patients in the subacute stage of the disease were displaying signs of gustatory impairment (hypogeusia), and 86% either hyposmia or anosmia[10]. Olfactory cells, concentrated in the nose, are particularly vulnerable to coronavirus invasion and this is also the case with SARS-CoV-2[11,12]. Despite the sometimes conflicting and varied findings both on COVID-19 neurological presentation and neuropathology based on autopsy reports, ranging from inflammatory processes to cerebrovascular events or endothelial damages, and also to no COVID-19-specific pathology, one hypothesis is that an entry point of the virus to the central nervous system might be achieved by crossing the neural–mucosal interface in the olfactory mucosa, and progress from there onto the olfactory bulb in the brain[5]. Within the olfactory system, direct neuronal connections from and to the olfactory bulb encompass regions of the piriform cortex (the primary olfactory cortex), the parahippocampal gyrus/perirhinal cortex, entorhinal cortex and the amygdala[13,14]. Further connections of the primary olfactory cortex include the insula and medial and lateral orbitofrontal areas[13,14]. Notably, the insula, and more consistently its more dorsal part around the junction with the fronto-parietal operculum ("area G"), is the centre of the primary gustatory area, which then projects onto the orbitofrontal cortex[15] (doi.org/10.1016/B978-012547626-3/50032-6).

Most brain imaging studies of COVID-19 to date have focused on acute cases and radiological reports of single cases or case series based on CT, PET or MRI scans, revealing a broad array of gross cerebral abnormalities ranging from white matter hyperintensities, hypoperfusion

and signs of ischaemic events spread throughout the brain, but found more consistently in the cerebrum (doi.org/10.3389/fnagi.2021.646908). Of the few larger studies focusing on cerebrovascular damage using CT or MRI, some have either found no clear marker of abnormalities in the majority of their patients, or importantly no spatially consistent pattern for the distribution of white matter hyperintensities or microhaemorrhages, except perhaps in the middle or posterior cerebral artery territories and the basal ganglia. Imaging cohort studies of COVID-19, quantitatively comparing data across subjects through automated preprocessing and realignment of images are much rarer. For instance, a recent cohort study focusing on correlates of cognitive impairment has demonstrated in 29 COVID-19 patients at a subacute stage the involvement of fronto-parietal areas revealed as ^{18}F-FDG hypometabolism[10]. Another glucose PET study has shown bilateral hypometabolism in the bilateral orbital gyrus rectus and the right medial temporal lobe[16]. One multi-organ imaging study in over 50 COVID-19 previously hospitalised patients revealed modest abnormalities in T2* of the left and right thalami compared with matched controls[17], while the study's brain-focused follow-up confirmed this result and further showed in particular an increase in white matter hyperintensity volume (www.medrxiv.org/content/10.1101/2021.05.19.21257316v1.full). It remains unknown however whether any of these abnormalities *predates* the infection by SARS-CoV-2. These findings could define an increased brain vulnerability to the deleterious effects of COVID-19 and a higher probability to show more pronounced symptoms, rather than being a consequence of the COVID-19 disease process.

UK Biobank (UKB) offers a unique resource to elucidate these issues. With the data from this large, multi-modal brain imaging study, we use for the first time a longitudinal setting whereby participants had been already scanned as part of UKB *before* contracting the virus. They were then imaged again, on average 37 months later, after some had tested positive for COVID-19, and those participants were then matched with controls who had undergone the same longitudinal imaging protocol but tested negative for COVID-19. In total, 394 COVID-19 patients with usable imaging data at both timepoints were included in this study, as well as 388 controls, matched for age, sex, ethnicity and time elapsed between the two scans. These large numbers may allow us to detect subtle, but consistently spatially distributed sites of damage caused by the disease, thus underlining *in vivo* the possible spreading pathways of the virus within the brain. The longitudinal aspect of the study should in turn make it possible

to tease apart which, of the observed effects post infection, were actually likely caused by the disease itself.

Our general approach in this study was therefore as follows: (i) utilise brain imaging data from 782 participants who have visited the UKB imaging centres for two scanning sessions, on average 3 years apart, with 394 of these having contracted COVID-19 in between their two scans; (ii) estimate — from each subject's multimodal brain imaging data — hundreds of distinct brain imaging-derived phenotypes (IDPs), each IDP being a distinct measure of one aspect of brain structure or function; (iii) model confounding effects, and estimate the longitudinal change in IDPs between the two scans; and (iv) identify significant patient vs. control group differences in these longitudinal brain change estimates, correcting for multiple comparisons across IDPs. We did this for both a focussed set of *a priori* defined IDPs testing the hypothesis that the olfactory system is a major route of infection into the brain, as well as an exploratory set of analyses considering all IDPs. In both cases we identified significant (corrected-P < 0.05) effects of COVID-19, primarily relating to loss of grey matter in cortical areas directly connected to primary olfactory and gustatory cortex.

Methods

Study Design

As part of UK Biobank imaging[18], thousands of subjects had received brain scans before the start of the COVID-19 pandemic. In total, quality-controlled multimodal brain imaging data, collected at four sites with identical imaging hardware, scanner software and protocols, was obtained from 42,729 participants over the age of 45y, and made available to researchers worldwide.

Before the COVID-19 pandemic, longitudinal (first- and second-timepoint scanning) had already begun in UKB imaging, with around 3,000 participants visiting for their second scan prior to scanning being paused in 2020 as a result of the pandemic. More recently, starting in February 2021, hundreds of UK Biobank participants who had already taken part in UKB imaging before the pandemic were invited back for a second scan. This COVID-19 re-imaging study was set up to investigate the effects of SARS-CoV-2 infection on internal organs by comparing imaging scans taken from participants before vs. after infection. Amongst those, participants were identified as having had COVID-19 based on either their primary care (GP) data, hospital records, or results of their antigen tests identified through record linkage to the Public Health datasets in England, Wales and Scotland. Another source of information was drawn from the home-based lateral flow kits sent by UKB to all participants who agreed to receive one. A second kit was sent to all participants who recorded an initial positive result, and who had indicated they had not yet been vaccinated, in order to double-check the result and reduce the number of false positives.

More specifically, the full list of inclusion criteria is as follows:

- had already attended an imaging assessment at one of the three imaging sites (the fourth opened just before the pandemic began),
- still lived within the catchment area of the clinic they attended for their first imaging assessment,
- no incidental findings were identified from their scans taken at the first imaging visit,
- had not withdrawn or died
- had a valid email and postal address,
- high-quality scans were obtained from the first imaging visit,

- lived within 60 km of the clinic (extended to 75 km in Feb 2021), due to travel restrictions during the lockdown period.

(See for more details the online documentation: https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/casecontrol_covidimaging.pdf)

Controls were then selected by identifying, from the remaining UK Biobank participants with negative results (from the home-based lateral flow kits, and no history of positive tests from primary care, hospital records or Public Health Wales/England/Scotland), those who could be paired to each of the COVID patients. Pairing was done following five criteria:

- sex
- ethnicity (white/non-white, as numbers were too low to allow for a finer distinction)
- date of birth (+/-6 months)
- location of imaging assessment
- date of first imaging assessment (+/-6 months).

Image Processing

For this work, we primarily used the IDPs generated by our team on behalf of UKB, and made available to all researchers by UKB. The IDPs are summary measures, each describing a different aspect of brain structure or function, depending on what underlying imaging modality is used[18,19].

The protocol includes three structural MRI scans: T1, T2 fluid attenuation inversion recovery (FLAIR) and susceptibility-weighted MRI, as well as diffusion MRI and resting and task functional MRI. T1 scans make it possible to derive global measures of brain and CSF volumes, as well as localised measures of grey matter volume and cortical thickness and area. The T2 FLAIR scan identifies differences that might be indicative of inflammation or tissue damage. Susceptibility-weighted MRI is sensitive to iron content. Diffusion MRI measurements give insight into the tissue microstructure integrity. Resting-state functional MRI is performed on an individual who is not engaged in any particular activity or task and can provide indices related to the functional connectivity between brain regions[20]. Individual functional connectivity information contains valuable signal, but is intrinsically noisy when each

connection is considered individually, so we focused here our analysis on 6 dimensionally-reduced functional connectivity networks[21]. For the UKB COVID-19 second-scan imaging, for the first time, arterial spin labelling (blood flow) imaging was added to the protocol; we did not use this data here in our longitudinal study, as this ASL data is only available from the second timepoint scans.

We therefore used 2,105 IDPs from those disseminated by UKB, including: regional and tissue volume, cortical parcels' surface area, volume and thickness, regional and tissue intensity, cortical grey-white contrast, white matter hyperintensity volume, white matter microstructural measures such as anisotropy and diffusivity, resting-state connectivity and amplitude measures, and task fMRI activation. We also generated 255 new IDPs, as described below.

We computed additional IDPs obtained using Quantitative Susceptibility Mapping (QSM), which has been recently added into our UK Biobank processing pipeline, and is briefly described below. Magnitude and phase data from the susceptibility-weighted MRI acquisitions were processed to provide quantitative measures reflecting clinically-relevant tissue susceptibility properties. The core UK Biobank pipeline already takes the magnitude data from two echoes to provide a quantitative mapping of T2* signal decay. Here, we adapted this following recent work correcting for background field bias[22], providing slightly more robust measures of T2*. Median T2* was calculated within 17 subcortical structures (estimated from the T1) as IDPs; 14 of these are the same regions already estimated by the core UK Biobank pipeline, and here we added 3 more subcortical regions-of-interest (ROIs): left and right substantia nigra[23] and regions of white matter hyperintensities (lesions)[24]. Second, susceptibility-weighted MRI phase data were processed for QSM following a pipeline being developed for UK Biobank[25] (www.medrxiv.org/content/10.1101/2021.05.19.21257316v1.full). Briefly, phase data of individual channels were combined following the MCPC-3D-S approach[26], channel-combined phase data was unwrapped using a Laplacian-based algorithm[27] and filtered using the V-SHARP algorithm[28] to remove background field, susceptibility maps were calculated using the iLSQR algorithm[29]. Susceptibility in CSF was used as the reference for the susceptibility maps. QSM IDPs (CSF-referenced) IDPs were calculated in the same 17 subcortical structures as the T2* IDPs.

Additional IDPs were created via refined sub-segmentations of the hippocampus and amygdala, and of the thalamic nuclei, as implemented in FreeSurfer[30-33]. We extracted these ROI masks from the FreeSurfer processing and applied them to the T2* and diffusion images to generate additional subcortical IDPs. Substantia nigra was also used to extract grey matter volume, T2* and diffusion measures.

Finally, we generated new IDPs tailored to the olfactory system, as described below.

Olfactory and gustatory systems hypothesis-driven approach

Based on prior expectations based on animal models and *post mortem* findings, we chose to focus *a priori* our primary analyses on a subset of 332 regions-of-interest (297 of which passed the reproducibility thresholding) from the available IDPs that anatomically corresponds to the telencephalic primary and secondary connections of the olfactory and gustatory cortex. Briefly, these include the piriform cortex, parahippocampal gyrus/perirhinal cortex, entorhinal cortex, amygdala, insula, frontal/parietal operculum, medial and lateral orbitofrontal cortex, hippocampus and basal ganglia[13-15] (doi.org/10.1016/B978-012547626-3/50032-6). As no labelling of the piriform cortex exists in any of the atlases used in the UK Biobank imaging processing, we refined a previously published ROI made openly available of the piriform cortex (frontal and temporal), anterior olfactory nucleus and olfactory tubercle by editing this primary olfactory ROI in our standard space (https://github.com/zelanolab/primaryolfactorycortexparcellation[34]). We further used maps from the same study's resting-state fMRI analysis of each of the four parts of this ROI (piriform frontal, piriform temporal, anterior olfactory nucleus and olfactory tubercle) to the rest of the brain, to generate four additional extended ROIs of the functionally-connected cortical and subcortical regions to these primary olfactory areas[34]. For this, we masked their connectivity t-value maps to keep only significant voxels, and further extracted grey matter volume (GM), T2* and diffusion values by (i) regressing each of these maps into the GM, T2* or diffusion images in their respective native spaces, or (ii) by binarising the maps and extracting mean and 95th percentile values.

Separate masks for the left and right olfactory bulbs were generated by manually drawing a binary mask for the right olfactory bulb on an averaged template-space T2 FLAIR volume generated from 713 subjects, and mirroring this to obtain the mask for the left (having confirmed by visual

inspection that symmetry in this region allowed for this to be effective). Both masks were then modulated by the T2 intensities in their respective ROIs, and normalised to account for partial volume effects, which provided the final "label" maps with values ranging between 0-1. For the hypothalamus, we combined and refined ROIs from two previously published and publicly available atlases of a probabilistic hypothalamus map (https://neurovault.org/collections/3145/[23]) and hypothalamic subregions[35]. Both the probabilistic hypothalamus map and the binarised map obtained from fusing the 26 hypothalamic subregions were transformed to our standard space where the probabilistic map was then masked by the binarised map. We then extracted volume, and T2 mean and 95th percentile intensity measurements in subjects' native spaces, using the olfactory bulb and hypothalamus maps (unthresholded and thresholded at 0.3). For the hypothalamus, we additionally extracted these metrics from T2* and diffusion images (diffusion tensor model MD and FA; and NODDI model OD, ISOVF, ICVF).

The full list of pre-determined and reproducible ROIs is available in **Supplementary Table 1.**

Exploratory Approach

The full set of 2,360 IDPs described above (including 2,105 already released by UKB) were used for a more exploratory, inclusive analysis of disease effects on brain structure and function (see full list in **Supplementary Table 2**).

Statistical Modelling

The following modelling was applied in the same way to both the hypothesis-driven analyses of a subset of IDPs, and the all-IDPs exploratory analyses.

All IDPs from all subjects were pooled for initial processing: 42,729 Scan 1 datasets (all pre-pandemic), 2,943 pre-pandemic Scan 2 datasets, and 798 Scan 2 datasets acquired after the beginning of the COVID-19 pandemic (COVID patients and controls combined). Extreme outlier values (individual IDPs from individual scanning sessions) were removed, on the basis of being more extreme than 5 times the median absolute deviation from the median (for a given IDP).

The IDPs from the 798 subjects imaged during the pandemic (COVID patients and controls combined, at both timepoints) were then treated as a single dataset comprising 1,596 imaging sessions, and cross-sectional deconfounding (treating all scans equivalently) was carried out for head size, scanner table position, and image motion in the diffusion MRI data (these confound variables first had outlier removal applied as described above)[19,36]. The difference between the IDP values at the two timepoints (ΔIDP = IDP2 - IDP1) was then computed, to feed into the main disease-effect analysis. Subjects were kept if at least the T1-weighted structural image was usable from both timepoints, resulting in ΔIDPs from 782 subjects.

At this point, we evaluated the scan-rescan reproducibility of IDPs, primarily in order to discard IDPs that were not highly reproducible between scans, and to offer a measure of contrast between true subject variability and session variability and noise. For each IDP, we correlated the Scan 1 with Scan 2 values, separately for COVID patients and controls, resulting in two reproducibility measures (Pearson correlation r) for each IDP. This assumes that potential ageing and disease effects are subtle compared with between-subject variability and IDP noise, and indeed the (N_{IDPs} x 1) vectors of reproducibility for patients and controls are very highly correlated (r=0.98); hence, we averaged these patient and control values to give a single reproducibility measure for each IDP. At this point, from the initial set of 2,360 IDPs, the least reproducible IDPs (r<0.5) were discarded, leaving 2,023 IDPs. Finally, IDPs with high levels of missing data (values from fewer than 50 subjects) were discarded, leaving in total 2,022 IDPs and ΔIDPs.

Despite initial between-group subject pairing, missing data disrupted the exact matching, and thus we included in the modelling those variables originally used as pairing criteria. Each ΔIDP (pooled across all patients and controls) was thus deconfounded for: difference between the subject's ages at each of their two scans, the difference of the squares of the ages (to account for quadratic dependencies of IDPs on age), genetic sex, and ethnicity (white vs non-white). Each deconfounded ΔIDP was then correlated against a patient-vs-control binary variable modulated by age at Scan 2 (that being a close proxy for age at COVID-19 infection for the entire patient group, with less than a year's error). As the effects of COVID-19 increase dramatically with age[37], for example, exponentially[38], we used a high power of age at Scan 2 (age^6) to modulate the demeaned binary patient-vs-control variable. The regressor is therefore:

PatientvsControlRegressor = demeaned(PatientvsControlBinaryIndicator) * (AgeScan2^6)

The specific choice of age-dependency was not critical: for example, it gave an almost identical (r>0.95) regressor to other age-increasing options such as age^{10} and $10^{age/20}$. This regressor also correlated highly with an "uninformed" unmodulated binary regressor (r=0.8), and likewise the final IDP-disease associations (Z-statistics) were also highly similar (r=0.87) between using modulated vs. unmodulated regressors.

The patient-vs-control regressor was then adjusted for the same confounds as applied to the ΔIDP values; in some circumstances this could affect validity of exchangeability in permutation testing, and we verified that here the corrected p-values (without vs with deconfounding of the patient-vs-control regressor) were identical (r>0.999).

Pearson correlation was then computed to find the strength of the association between each ΔIDP and the patient-vs-control variable, followed by the Fisher transformation to Z-statistics (including standardisation by scaling by sqrt(N-3), and hence taking into account the different patterns of missing data for different IDPs). For each IDP's correlation, any missing data is ignored (that subject is left out for that sub-analysis). The Z statistics are valuable because they provide a standardised measure of effect strength accounting for varying degrees of freedom in different IDPs.

We then used permutation testing to correct for the multiple comparisons across IDPs while accounting for the strong dependences among these responses. We randomly permuted the patient-vs-control regressor relative to the ΔIDPs, with 5,000 random permutations. At each permutation we computed the association Z value for each ΔIDP, and recorded the maximum absolute value across all IDPs. By taking the absolute value, we corrected for the two-tailed nature of the test (we did not pre-assume the direction of any disease effect). After building up the null distribution of the maximum across IDPs, we then tested the original |Z| values against this, to obtain family-wise error corrected-p-values Pfwe fully correcting for multiple comparisons across IDPs.

We refer below to this main group comparison analysis (comparing ΔIDPs between COVID patients and controls) as Model 1. As secondary analyses, we also applied the same approach to compare hospitalised COVID patients against controls (Model 2), and to compare

hospitalised COVID patients against non-hospitalised patients (Model 3), although power to detect effects in these tests would be lower, given the relatively small number of known hospitalised cases.

For all 3 models, testing was carried out twice: first using the *a priori* focussed set of IDPs identified for the hypothesis-driven analyses, and then using the full set of IDPs for the more exploratory analyses. In both cases IDPs were identified as having significant group difference results fully corrected for multiple comparisons.

Results

Participants

UK Biobank (UKB) has been releasing data from the COVID-19 re-imaging study on a rolling basis. As of 1 June 2021, 404 participants met the re-imaging study inclusion criteria and were identified as having had COVID-19 based on either their primary care (GP) data, hospital records, results of their antigen tests identified through record linkage to the Public Health datasets in England, Wales and Scotland, or two concordant home-based lateral flow kits positive results. Of these 404 COVID patients, a total of 394 had usable brain scans at both timepoints (**Tables 1 and 2**).

In total, 388 controls met the inclusion criteria and had usable brain scans at both timepoints (**Table 1**). COVID status was identified using UKB Showcase variable 41000.

Table 1. Main demographics of the COVID patients and controls. Non-parametric tests were used whenever a variable for each group was not normally distributed (Lilliefors P = 1e-03). Two-sample Kolmogorov-Smirnov test was used for age at Scan 1 or Scan 2 and Years between Scan 1 and Scan 2. Chi-square test was used for sex, ethnicity, and diagnosed diabetes. Mann-Whitney U-test was used for the systolic and diastolic blood pressures, weight, waist/hip ratio, and BMI.

	COVID patients	**Controls**	**P**
Number of subjects	394	388	
Age at Scan 1, mean±SD (range)	59.1±7.0 (46.9–80.2)	60.4±7.5 (47.1–79.8)	0.13
Age at Scan 2, mean±SD (range)	62.2±6.7 (51.3–81.4)	63.5±7.1 (51.3–81.3)	0.09
Sex, male/female	170 (43.1%) / 224 (56.9%)	166 (42.8%) / 222 (57.2%)	0.92
Ethnicity, white/non-white*	382 (97.0%) / 12 (3.0%)	377 (97.2%) / 11 (2.8%)	0.86
Years between Scans 1 and 2, mean±SD (range)	3.1±1.6 (1.0–7.0)	3.1±1.6 (1.0–6.9)	0.96
Systolic blood pressure [mmHg]	115.1±30.2	114.1±31.1	0.72
Diastolic blood pressure [mmHg]	79.8±8.8	79.8±9.0	0.76
Diagnosed diabetes	15	20	0.36
Weight [kg]	76.1±15.4	75.9±15.3	0.74
Waist/Hip ratio	0.85±0.09	0.85±0.09	0.84
BMI [kg/m^2]	26.4±4.1	26.4±4.3	0.99

*The white/non-white distinction was made as numbers were unfortunately too low to allow for a finer distinction

Despite the original matched-pairing of the COVID-19 patients and controls, their age distributions were slightly — though not statistically significantly — different, due to different patterns of missing/usable data (**Supplementary Figure 1**). Note that the control group is on average slightly (not significantly) older than the COVID group, which would be expected to make any change between the two timepoints harder to detect in the group comparisons, rather than easier (for example, this is unambiguously the case with results in IDPs measuring cortical thickness and volume, all of which decrease with aging and most pathologies).

Table 2. Main clinical information available for the COVID patients. Of note, of the 394 COVID patients used in our main analyses, 60 were identified as cases via the two different home-based lateral flow kits and thus do not have additional clinical information (or date of diagnosis) in their primary care or hospital records.

	N or mean (SD)
Total number of covid+ patients	394
Origin of diagnosis	
- GP	14
- Hospital	47
- Public Health England, Wales, Scotland	273
- Home-based lateral flow kits	60
Total	394
Number of COVID patients with available information on date of diagnoses	334
Date of diagnosis	
- 2020/01/01 – 2020/03/31	4
- 2020/04/01 – 2020/06/30	22
- 2020/07/01 – 2020/09/30	17
- 2020/10/01 – 2020/12/31	194
- 2021/01/01 – 2021/04/01	97
Total	334
ΔDays of COVID-19 infection before Scan 2, mean±SD (range)	134.9±72.6 (34–402)
Total number of hospitalized subjects	15
- COVID-19 as primary cause	11
- COVID-19 as secondary cause	4
- Days of hospitalization, mean±SD (range)	10.3±10.6 (0–39)
- Critical care unit	2
- Invasive ventilation	1
- Continuous positive airway pressure	1
- Non-invasive ventilation	1
- Unspecified oxygen therapy	1
- No procedure recorded	7

Of the 334 patients for whom clinical information was available through primary care (GP) or hospital records, 15 were hospitalised with COVID-19, including 2 in critical care (**Tables 2 and 3**). These hospitalised patients were on average older, had higher blood pressure and BMI, and were more likely to be men and have diabetes (**Table 3**).

Table 3. Comparison between COVID patients (hospitalised vs. non-hospitalized). For statistical procedures, please refer to **Table 1**.

	Hospitalized	Non-hospitalized	P
Number of subjects	15	379	
Age at Scan 1, mean±SD (range)	67.0±8.4 (52.2–80.2)	58.8±6.8 (46.9–77.0)	<0.001
Age at Scan 2, mean±SD (range)	69.6±7.9 (57.4–81.4)	61.9±6.5 (51.3–80.0)	<0.001
Sex, male/female	11 (73.3%) / 4 (26.7%)	159 (42.0%) / 220 (58.0%)	0.016
Ethnicity, white/non-white*	15 (100%) / 0 (0%)	367 (96.8%) / 12 (3.2%)	0.48
Years between Scan 1 and 2, mean±SD (range)	2.5±1.4 (1.0–5.8)	3.2±1.6 (1.1–7.0)	0.21
Systolic blood pressure [mmHg]	130.0±32.6	114.5±30.0	0.08
Diastolic blood pressure [mmHg]	84.4±9.1	79.6±8.8	0.04
Diagnosed diabetes	3	12	<0.001
Weight [kg]	86.0±11.2	76.1±15.5	0.0044
Waist/Hip ratio	0.92±0.08	0.84±0.09	0.0012
BMI [kg/m^2]	28.9±3.7	26.3±4.1	0.010

*The white/non-white distinction was made as numbers were unfortunately too low to allow for a finer distinction

Olfactory and gustatory systems hypothesis-driven longitudinal results

The time x group interaction analyses between the 394 COVID patients and 388 controls (Model 1) on 297 IDPs yielded 8 significant results surviving correction for multiple comparisons (**Table 4**, **Figure 1, Supplementary Table S1** for full list of results). All significant longitudinal effects were found in the primary or secondary cortical gustatory and olfactory areas, in the left hemisphere, using grey matter information (thickness and volume). Four of these IDPs covered a similar region of the parahippocampal gyrus/perirhinal cortex in the left hemisphere (slightly differently defined depending on each the three predefined atlas parcellations used here), while three IDPs covered the region of the left lateral orbitofrontal

cortex. The final region was in the left superior insula (precisely in the superior segment of the circular sulcus of the insula). All three regions showed greater reduction of grey matter thickness or volume over time in the 394 COVID patients compared with the 388 controls.

While these results were not significant for the comparison between the 15 hospitalised COVID patients and 388 controls (Model 2), or between the 15 hospitalised COVID patients and 379 non-hospitalised ones (Model 3), likely due to the reduction in degrees of freedom in these two models, Z values obtained for these other two group comparisons suggest that similar effects were seen in the parahippocampal gyrus (max Z=-4.3 for the Model 2; Z=-3.2 for Model 3) and superior insula to a lesser extent (max Z=-3.3; Z=-2.1), i.e. the first cortical relays of the olfactory and gustatory systems.

Table 4. Hypothesis-driven olfactory and gustatory approach: longitudinal group comparison results. The top 10 results out of 297 IDPs, based on their corrected family-wise error (FWE) p-values, showing where the 394 COVID patients and 388 controls differed over time are presented. In bold, the findings surviving correction for multiple comparisons based on this contrast. All significant results were found in the primary or secondary cortical gustatory and olfactory areas, in the left hemisphere, using grey matter information (volume, thickness). As indicated by the Z values in the other two group comparisons, those results are similar for the comparison between the 15 hospitalised COVID patients and controls, and between hospitalised COVID patients and non-hospitalised ones, especially in the parahippocampal gyrus and superior insula (which are both first cortical relays of the olfactory and gustatory system, respectively). The reproducibility measurements show the scan-rescan similarities of IDPs across both COVID patients and controls (see **Methods**).

	COVID vs controls			COVID/Hosp+ vs controls		COVID/Hosp+ vs COVID/Hosp-			
Imaging-Derived Phenotype (IDP) Description	**r**	**Z**	**Pfwe**	**Z**	**Pfwe**	**Z**	**Pfwe**	**L/R**	**Repro**
Parahippocampal gyrus - thickness (DKT atlas)	**-0.16**	**-4.2**	**0.01**	-4.4	0.07	-3.2	0.42	**L**	**0.83**
Lateral orbitofrontal cortex - thickness (DKT atlas)	**-0.16**	**-4.2**	**0.01**	-2.5	0.93	-1.2	1	**L**	**0.58**
Parahippocampal gyrus - thickness (Desikan atlas)	**-0.14**	**-3.9**	**0.03**	-4.3	0.08	-3.3	0.36	**L**	**0.83**
Lateral orbitofrontal cortex - volume (Desikan atlas)	**-0.15**	**-4.0**	**0.03**	-2.7	0.79	-1.6	1	**L**	**0.93**
Lateral orbitofrontal cortex - thickness (Desikan atlas)	**-0.14**	**-3.9**	**0.03**	-2.1	1	-0.8	1	**L**	**0.63**
Perirhinal cortex - thickness (Brodmann's area atlas)	**-0.15**	**-4.0**	**0.03**	-2.1	1	-0.6	1	**L**	**0.60**
Parahippocampal gyrus - thickness (Destrieux atlas)	**-0.15**	**-4.0**	**0.03**	-2.2	0.98	-0.6	1	**L**	**0.71**
Superior insula - thickness (Destrieux atlas)	**-0.15**	**-3.9**	**0.03**	-3.8	0.14	-2.1	1	**L**	**0.82**
Entorhinal cortex - thickness (DKT atlas)	-0.14	-3.7	0.06	-0.3	1	1.1	1	L	0.58
Lateral orbitofrontal cortex - volume (DKT atlas)	-0.14	-3.7	0.07	-2.3	0.97	-1.3	1	L	0.92

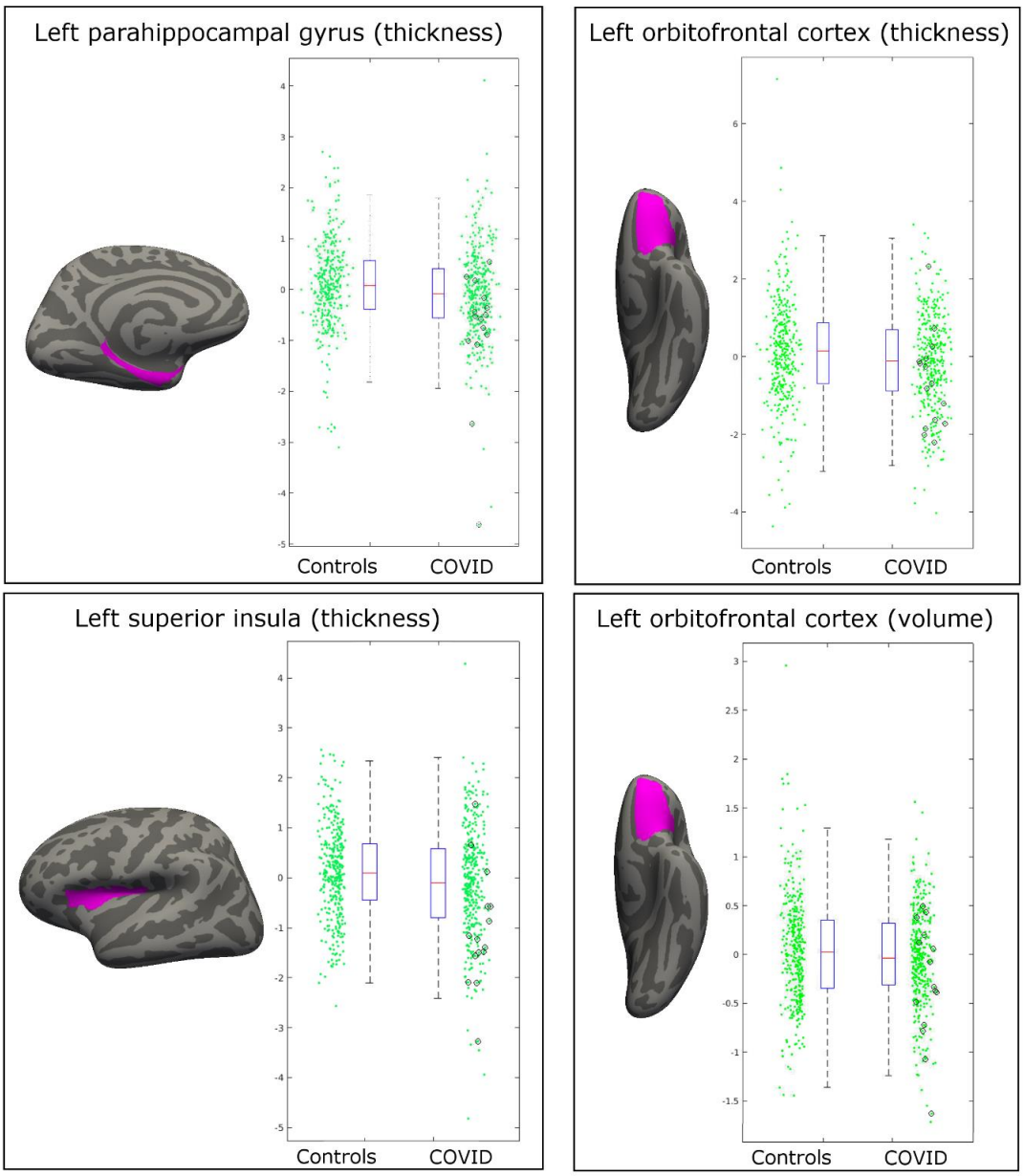


Figure 1. Most significant longitudinal group comparison results. The three main regions showing significant loss of grey matter (thickness, volume) between the two timepoints specifically for the COVID patients are the parahippocampal gyrus, the lateral orbitofrontal cortex, and the superior insula. All results were localised to the left hemisphere. For each region, the IDP spatial region of interest shown in magenta, overlaid on the FreeSurfer average inflated cortical surface; to the right are the scatter and box plots showing the difference in cortical thickness or volume between the two timepoints for the 388 controls and 394 COVID patients. In black circles are the 15 hospitalised COVID patients. All y axes are arbitrary units proportional to the original measures, due to the normalisation steps in the IDP preprocessing.

Exploratory approach

2,022 IDPs passed the initial tests of reproducibility and data completeness (see **Supplementary Table S2** for the complete list of reproducible IDPs and results). The main analysis (Model 1) revealed four measures with significant longitudinal difference between the 394 COVID patients and 388 controls after correction for multiple comparisons: two global measures of volume (ratio of the volume of the segmented brain over the estimated total intracranial volume generated by FreeSurfer, and volume of right lateral ventricle), and two measures that independently confirmed the involvement of olfactory-related cortical areas (the left parahippocampal gyrus and left lateral orbitofrontal cortex)(**Table 5**). Due to the very limited degrees of freedom in our second and third models (15 hospitalised COVID patients vs controls, and those 15 hospitalised COVID patients vs non-hospitalised), these 4 results were not significant for these two other group comparisons. However, Z values suggest once again that the impact of COVID-19 is still measurable in the left parahippocampal gyrus (max Z=-4.4 for Model 2; max Z=-3.3 for Model 3).

Table 5. Exploratory approach: longitudinal group comparison results. The top 10 results showing where the 394 COVID patients and 388 controls differed over time are presented. In bold, the findings surviving correction for multiple comparisons based on this contrast. Apart from a few global measures relating to loss of brain volume (increase of CSF volume), most of the top exploratory results implicate again the primary and secondary connections of the olfactory and gustatory system (in the left hemisphere). As indicated by the Z values in the other two group comparisons, those results are similar for the comparison between the hospitalised COVID patients and controls, and between hospitalised COVID patients and non-hospitalised ones, especially in the parahippocampal gyrus.

	COVID+ vs COVID-			COVID+Hosp+ vs COVID-		COVID+Hosp+ vs COVID+Hosp-			
Imaging-Derived Phenotype (IDP) Description	**r**	**Z**	**Pfwe**	**Z**	**Pfwe**	**Z**	**Pfwe**	**L/R**	**Repro**
Ratio brain volume/estimated total intracranial volume	**-0.19**	**-5.0**	**0.001**	-2.9	0.99	-2.0	1	**N/A**	**0.92**
Lateral ventricle - volume	**0.17**	**4.4**	**0.02**	2.2	1	0.7	1	**R**	**0.98**
Parahippocampal gyrus - thickness (DKT atlas)	**-0.16**	**-4.2**	**0.04**	-4.4	0.19	-3.2	0.93	**L**	**0.83**
Lateral orbitofrontal cortex - thickness (DKT atlas)	**-0.16**	**-4.2**	**0.04**	-2.5	1	-1.2	1	**L**	**0.58**
Unnormalised CSF volume	0.15	4.1	0.07	1.7	1	-0.3	1	N/A	0.98
Lateral orbitofrontal cortex - volume (Desikan atlas)	-0.15	-4.0	0.09	-2.7	1	-1.6	1	L	0.93
Perirhinal cortex - thickness (Brodmann's area atlas)	-0.15	-4.0	0.09	-2.1	1	-0.6	1	L	0.6
Parahippocampal gyrus - thickness (Destrieux atlas)	-0.15	-4.0	0.11	-2.2	1	-0.6	1	L	0.71
Superior insula - thickness (Destrieux atlas)	-0.15	-3.9	0.12	-3.8	0.49	-2.1	1	L	0.82
Parahippocampal gyrus - thickness (Desikan atlas)	-0.14	-3.9	0.14	-4.3	0.23	-3.3	0.90	L	0.83

As so many of the top exploratory and hypothesis-driven results included IDPs of cortical thickness particularly in the left hemisphere (e.g., 11 out of the top 20 for the exploratory approach, and 10 out of 20 for the ROI-based approach), we further conducted a visual analysis of the vertex-wise longitudinal thickness differences between the 394 COVID patients and 388 controls over the entire cortical surface (**Figure 2**). This again suggested that the strongest deleterious effects of COVID-19 could be seen predominantly in the left hemisphere, extending beyond the parahippocampal gyrus, orbitofrontal cortex and insula to the left supramarginal gyrus and anterior cingulate cortex, and in the right temporal pole — also weakly present in the left hemisphere.

Of note, when looking at the maximum |Z| values for Model 3, which compares longitudinal changes between hospitalised and non-hospitalised COVID patients, the top 20 largest effects could be seen reliably in olfactory- and memory-related areas (**Supplementary Table 2**), including the left parahippocampal gyrus, the left anterior, posterior and isthmus of the cingulate cortex, and the right CA1 region of the hippocampus and right cortical nucleus of the amygdala.

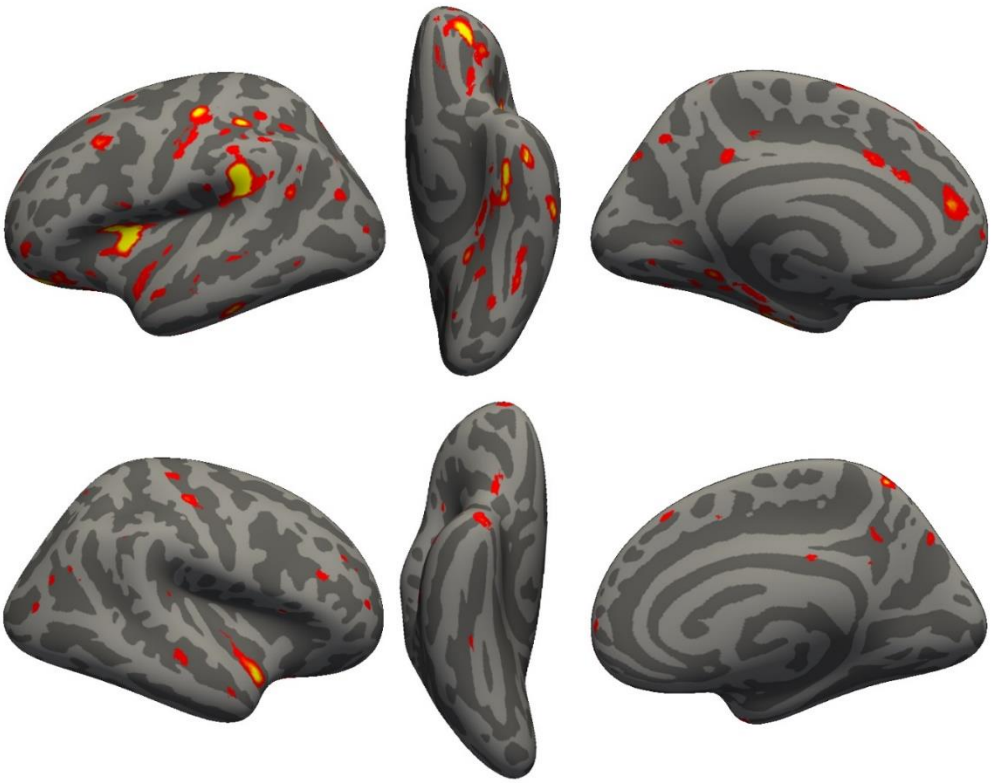


Figure 2. Vertex-wise longitudinal group comparison results in grey matter thickness. Thresholded map (|Z|>3) showing the strongest longitudinal differences between the 394 COVID participants and 388 controls. More prominent reduction in cortical thickness between the two scans for the COVID patients can be seen particularly in the left hemisphere in the anterior parahippocampal gyrus (perirhinal cortex), the anterior and lateral orbital gyrus, superior insula, supramarginal gyrus and anterior cingulate cortex. In the right hemisphere, there is also a notable reduction of grey matter thickness in the temporal pole (also weakly present in the left hemisphere).

Discussion

To our knowledge, this is the first longitudinal imaging study in COVID-19 where COVID patients were initially scanned *before* they contracted the disease. This multimodal exploration of the impact of SARS-CoV-2 on the brain was made possible by leveraging the unique resource of the UK Biobank project, and its COVID-19 re-imaging study. In total, we were able to include 782 participants with two usable brain scans: 394 COVID patients, and 388 controls who were well-matched for age, sex, ethnic background and interval between the two scans. The converging longitudinal analyses revealed a significant, deleterious impact of COVID-19 on the olfactory and gustatory cortical systems, with a more pronounced reduction of grey matter thickness and volume in the left parahippocampal gyrus, the left superior (dorsal) insula and the left lateral orbitofrontal cortex in the COVID patients. Further investigation demonstrated that this effect extended to the left anterior cingulate cortex, the left supramarginal gyrus and the right temporal pole. Comparing the few patients (n=15) who had been hospitalised with COVID-19 with those who had not also showed a similar pattern of olfactory- and memory-related brain regions, particularly centred around the left cingulate cortex, and the right hippocampus and amygdala.

Much has been made of the benefit of using a longitudinal design to estimate in particular trajectories of brain ageing and cognitive decline[39] (www.biorxiv.org/content/10.1101/2021.02.08.428915v2). The longitudinal nature of the UK Biobank COVID-19 re-imaging study, with the baseline scan acquired *before* infection from SARS-CoV-2 and the second scan *after* infection, makes it possible to more confidently disentangle the contribution of the disease process from pre-existing differences in the brain (or risk factors) of future COVID-19 patients. An illustrative example is that, based solely on the cross-sectional group comparisons between COVID patients and controls at just the second timepoint post infection (i.e., equivalent to what *post hoc* studies would do), the strongest effect is seen in the volume of the thalamus ($r=-0.15$, $Z=-4.3$, $P_{fwe}=0.03$). This effect entirely disappears when taking into account the baseline scans however, since the thalamus of the participants who will later contract COVID-19 appears to already differ from the controls years before infection ($r=-0.15$, $Z=-4.3$, $P_{fwe}=0.04$). Whether these differences played a subsequent role in those patients being more likely to contract the virus, or to develop symptoms from them, would need to be investigated further by linking them to either genetic information, or the many lifestyle,

cognitive, physical and other health outcome measures available in UK Biobank. This highlights the difficulties in interpreting cross-sectional post-infection differences as being necessarily the consequence of COVID-19. The longitudinal aspect of this study design, combined with the matching of the patients and controls for age, sex, scan interval and ethnicity, increases the interpretability of the results, and confidence that the reduction in thickness and volume of those olfactory-related brain regions is an impact of the disease process. As this study is observational (as opposed to a randomised, controlled, interventional study), one cannot make claims of disease causality with absolute certainty, but interpretational ambiguities are greatly reduced compared with *post hoc* cross-sectional studies.

Our cohort-based, quantitative imaging study, unlike the majority of single case and case series studies published so far, does not focus on gross abnormalities that could be observed at the single subject-level with a naked eye, such as microhaemorrhages or (sub)acute ischaemic infarctions (doi.org/10.3389/fnagi.2021.646908). It does however rely on a *consistent* pattern of abnormalities caused by the disease process, a common spatial distribution of these pathological alterations across the patients, which could be uncovered by aligning all the images together in a common standard space, followed by applying a pipeline of modality-specific image processing algorithms. This automated, objective and quantitative processing of the images facilitates the detection of subtle changes that would not be visible at the individual level, but which point to a possible mechanism for the spread of the disease inside the central nervous system.

Our hypothesis-driven analyses revealed a clear involvement of the olfactory and gustatory systems, which was further confirmed by the exploratory analyses and the vertex-wise map considering the entire cortical surface. While no significant longitudinal differences were seen in the piriform cortex or olfactory bulbs (both located in a region above the sinuses prone to susceptibility distortions to the brain images, and both being difficult to segment in MRI data), we identified consistent abnormalities in the left parahippocampal gyrus (**Tables 4 and 5**, **Figure 2**) — a limbic region of the brain involved in supporting memory of events — specifically in its most anterior part also known as the perirhinal cortex. The anterior parahippocampal gyrus' function is distinct from its posterior part, and is thought to relate to familiarity, and to specific encoding of object information[22,40], as well as to play a crucial,

integrative role for the relative temporal order of events in episodic memory[41]. Importantly, it is directly connected to the piriform cortex and entorhinal cortex, which are both part of the primary olfactory cortex[13,42]. Similarly, the orbitofrontal cortex, which we also found altered by COVID-19, is often referred to as the secondary olfactory cortex as it possesses direct connection to both entorhinal and piriform cortex[42], as well as to the anterior olfactory nucleus[14,34]. In fact, in a recent functional connectivity study of the primary olfactory cortex, the orbitofrontal cortex was found to be connected to all four primary olfactory regions investigated (frontal and temporal piriform cortex, anterior olfactory nucleus and olfactory tubercle), possibly explaining why it is reliably activated even in basic and passive olfactory tasks[34]. The insula, which completes the regions found to be significantly different between COVID and control groups, is not only directly connected to the primary olfactory cortex[14], but is also considered to be the primary gustatory cortex. "Area G" (i.e. the dorsal part of the insula at the junction with the frontal and parietal operculum), as it is coined, in turn connects with the orbitofrontal cortex[15] (doi.org/10.1016/B978-012547626-3/50032-6). The vertex-wise visualisation on the cortical surface of this loss of grey matter in the insula spatially corresponds in particular to the area of consistent activation to all basic taste qualities[43]. Interestingly, early neurological signs in COVID-19 are hyposmia and hypogeusia, which appear to precede the onset of respiratory symptoms in the majority of affected patients[2,11,44]. In addition, a popular hypothesis has been that an entry point to the central nervous system is via the olfactory bulb[2,5]. Whether our consistent findings, which seem to delineate a clear limbic network of the primary and secondary olfactory and gustatory cortex, reflect a brain-related pathological process of COVID-19, or a direct effect of the spread of the virus itself, remains unknown.

The majority of our results were found using imaging biomarkers of grey matter thickness or volume, which are very sensitive to detect the effects of a neurodegenerative process compared with other imaging modalities[45], and are particularly robust measurements that makes them ideal in a longitudinal setting[46]. In fact, the longitudinal effects between the COVID-19 and control groups, while markedly and significantly localised in a limbic olfactory and gustatory network, seemed also (at a lower level) to be generalised, as illustrated in the shift in the distribution of Z values over the entire cortical surface (which was more pronounced on the left than on the right, **Supplementary Figure 2**). Our strongest results are

found in the left hemisphere, although it remains to be investigated whether this effect is significant. This apparent asymmetry might relate to the asymmetry of the olfactory system itself, such as that of the olfactory bulb or its neuronal firing[47], or to the olfactory processes, where the left hemisphere seem to be more involved in the emotional aspect of olfactory memory[47] (the famous madeleine de Proust association that seemed particularly targeted in COVID-19[48]). This asymmetry of the olfactory memory has been successfully exploited by using for instance a simple but powerful olfactory test in Alzheimer's disease[49]. Remarkably, Alzheimer's disease patients demonstrated an asymmetrical — left greater than right — decrement of odour detection sensitivity[49].

The limbic nature of the regions of the olfactory system, and their physical proximity to the hippocampus in particular, raise the possibility that longer-term consequences of SARS-CoV-2 infection (for which some suggest that the coronavirus itself enters the brain via the olfactory route) might in time contribute to Alzheimer's disease or other forms of dementia[2]. This has lead to the creation of an international consortium including the Alzheimer's Association and representatives from more than 30 countries to investigate these questions[2]. The comparison, in "Model 3", of hospitalised vs non-hospitalised patients, further revealed more pronounced loss of grey matter in limbic regions, more specifically the central nucleus of the amygdala (primary olfactory cortex), the cingulate cortex and the cornu ammonis of the hippocampus, which all contribute to the olfactory system downstream of the olfactory bulb. This might represent a more deleterious impact of COVID-19, or be due to risk factors (as hospitalised patients were older, had higher BMI and blood pressure, and higher risk of having diabetes), or an interaction of both. The loss of grey matter in memory-related regions of the brain may in turn increase the risk of these patients of developing dementia in the longer term[2].

Limitations of this study include the lack of stratification of the patients beyond the information of whether they had been hospitalised (information on O_2 saturation levels and details of treatment or hospital procedures is currently available on only a few participants), lack of clinical correlates, small number of participants from Asian, Black or other ethnic background other than White, (minority of) cases and (all of) controls' COVID-19 status identified through lateral flow test kits that have varied diagnostic accuracy[50]. For those cases, no distinction is possible at present to determine whether a positive test is due to COVID-19

or thanks to vaccination, so potential cases identified only through lateral flow test in vaccinated participants were not included. Another issue inherent to the recruitment strategy of UK Biobank, based on participants volunteering after being contacted at home for a possible re-imaging session, is the high number of mild cases. This can be seen however as a strength of this study: the majority of the brain imaging publications so far having focussed on moderate to severe cases of COVID-19 (doi.org/10.3389/fnagi.2021.646908), there is a fundamental need for more information on the cerebral effects of the disease even in its mildest from. It is also worth noting that the UK Biobank COVID-19 re-imaging study is an ongoing effort, and self-report on symptoms and treatments, as well as further distinction between vaccination and COVID-19 status, will be made available.

This is the first longitudinal imaging study comparing brain scans acquired from individuals before and after SARS-CoV-2 infection to those scans from a well-matched control group. It also is one of the largest COVID-19 brain imaging studies, with 782 participants including 394 individuals with COVID-19. Its unique design makes it possible to tease apart the contribution of the disease from pre-existing risk factors. By using automated, objective and quantitative methods, we were able to uncover a consistent spatial pattern of loss of grey matter in limbic brain regions forming an olfactory and gustatory network. Whether these abnormal changes are the hallmark of the spread of the disease (or the virus itself) in the brain, which may prefigure a future vulnerability of the limbic system, including memory, for these patients, remains to be investigated.

Acknowledgements

This work was primarily supported by a Wellcome Trust Collaborative Award 215573/Z/19/Z. KLM is supported by a Wellcome Trust Senior Research Fellowship 202788/Z/16/Z. The Wellcome Centre for Integrative Neuroimaging (WIN FMRIB) is supported by centre funding from the Wellcome Trust (203139/Z/16/Z). This research has been conducted in part using the UK Biobank Resource under Application Number 8107. We are grateful to UK Biobank for making the data available, and to all UK Biobank study participants, who generously donated their time to make this resource possible. Analysis was carried out at the Oxford Biomedical Research Computing (BMRC) facility. BMRC is a joint development between the Wellcome Centre for Human Genetics and the Big Data Institute, supported by Health Data Research UK and the NIHR Oxford Biomedical Research Centre.

Ethics

Human subjects: UK Biobank has approval from the North West Multi-centre Research Ethics Committee (MREC) to obtain and disseminate data and samples from the participants (http://www. ukbiobank.ac.uk/ethics/), and these ethical regulations cover the work in this study. Written informed consent was obtained from all participants.

References

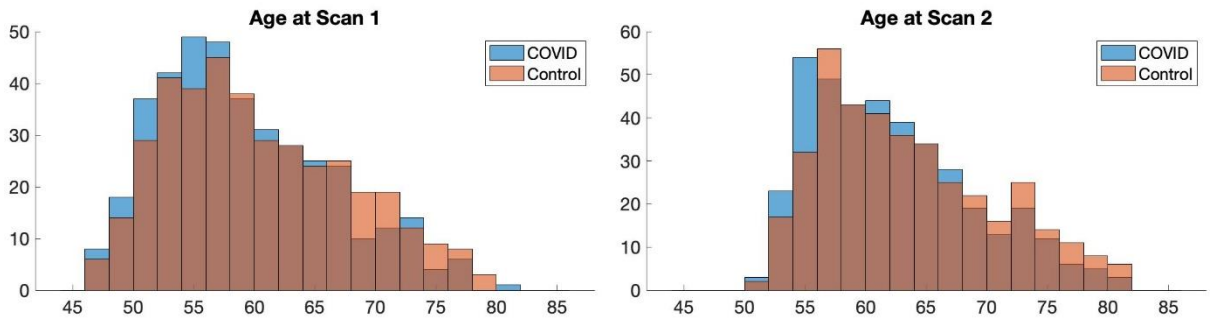
1. Paterson RW, Brown RL, Benjamin L, et al. The emerging spectrum of COVID-19 neurology: clinical, radiological and laboratory findings. *Brain* 2020; **143**(10): 3104-20.
2. de Erausquin GA, Snyder H, Carrillo M, et al. The chronic neuropsychiatric sequelae of COVID-19: The need for a prospective study of viral impact on brain functioning. *Alzheimers Dement* 2021.
3. Helms J, Kremer S, Merdji H, et al. Neurologic Features in Severe SARS-CoV-2 Infection. *N Engl J Med* 2020; **382**(23): 2268-70.
4. Mukerji SS, Solomon IH. What can we learn from brain autopsies in COVID-19? *Neuroscience letters* 2021; **742**: 135528.
5. Meinhardt J, Radke J, Dittmayer C, et al. Olfactory transmucosal SARS-CoV-2 invasion as a port of central nervous system entry in individuals with COVID-19. *Nat Neurosci* 2021; **24**(2): 168-75.
6. Puelles VG, Lutgehetmann M, Lindenmeyer MT, et al. Multiorgan and Renal Tropism of SARS-CoV-2. *N Engl J Med* 2020; **383**(6): 590-2.
7. Matschke J, Lutgehetmann M, Hagel C, et al. Neuropathology of patients with COVID-19 in Germany: a post-mortem case series. *Lancet Neurol* 2020; **19**(11): 919-29.
8. Lechien JR, Chiesa-Estomba CM, De Siati DR, et al. Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): a multicenter European study. *Eur Arch Otorhinolaryngol* 2020; **277**(8): 2251-61.
9. Cooper KW, Brann DH, Farruggia MC, et al. COVID-19 and the Chemical Senses: Supporting Players Take Center Stage. *Neuron* 2020; **107**(2): 219-33.
10. Hosp JA, Dressing A, Blazhenets G, et al. Cognitive impairment and altered cerebral glucose metabolism in the subacute stage of COVID-19. *Brain* 2021; **144**(4): 1263-76.
11. Butowt R, Bilinska K. SARS-CoV-2: Olfaction, Brain Infection, and the Urgent Need for Clinical Samples Allowing Earlier Virus Detection. *ACS Chem Neurosci* 2020; **11**(9): 1200-3.
12. Netland J, Meyerholz DK, Moore S, Cassell M, Perlman S. Severe acute respiratory syndrome coronavirus infection causes neuronal death in the absence of encephalitis in mice transgenic for human ACE2. *J Virol* 2008; **82**(15): 7264-75.
13. Palouzier-Paulignan B, Lacroix MC, Aime P, et al. Olfaction under metabolic influences. *Chem Senses* 2012; **37**(9): 769-97.
14. Carmichael ST, Clugnet MC, Price JL. Central olfactory connections in the macaque monkey. *The Journal of comparative neurology* 1994; **346**(3): 403-34.
15. Avery JA, Liu AG, Ingeholm JE, Riddell CD, Gotts SJ, Martin A. Taste Quality Representation in the Human Brain. *J Neurosci* 2020; **40**(5): 1042-52.
16. Guedj E, Million M, Dudouet P, et al. (18)F-FDG brain PET hypometabolism in post-SARS-CoV-2 infection: substrate for persistent/delayed disorders? *Eur J Nucl Med Mol Imaging* 2021; **48**(2): 592-5.
17. Raman B, Cassar MP, Tunnicliffe EM, et al. Medium-term effects of SARS-CoV-2 infection on multiple vital organs, exercise capacity, cognition, quality of life and mental health, post-hospital discharge. *EClinicalMedicine* 2021; **31**: 100683.
18. Miller KL, Alfaro-Almagro F, Bangerter NK, et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nature neuroscience* 2016; **19**(11): 1523-36.
19. Alfaro-Almagro F, Jenkinson M, Bangerter NK, et al. Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. *NeuroImage* 2018; **166**: 400-24.
20. Littlejohns TJ, Holliday J, Gibson LM, et al. The UK Biobank imaging enhancement of 100,000 participants: rationale, data collection, management and future directions. *Nat Commun* 2020; **11**(1): 2624.
21. Elliott LT, Sharp K, Alfaro-Almagro F, et al. Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature* 2018; **562**(7726): 210-6.
22. Diana RA, Yonelinas AP, Ranganath C. Imaging recollection and familiarity in the medial temporal lobe: a three-component model. *Trends in cognitive sciences* 2007; **11**(9): 379-86.

23. Pauli WM, Nili AN, Tyszka JM. A high-resolution probabilistic in vivo atlas of human subcortical brain nuclei. *Sci Data* 2018; **5**: 180063.
24. Griffanti L, Zamboni G, Khan A, et al. BIANCA (Brain Intensity AbNormality Classification Algorithm): A new tool for automated segmentation of white matter hyperintensities. *NeuroImage* 2016; **141**: 191-205.
25. Wang C, Foxley S, Ansorge O, et al. Methods for quantitative susceptibility and R2* mapping in whole post-mortem brains at 7T applied to amyotrophic lateral sclerosis. *NeuroImage* 2020; **222**: 117216.
26. Eckstein K, Dymerska B, Bachrata B, et al. Computationally Efficient Combination of Multi-channel Phase Data From Multi-echo Acquisitions (ASPIRE). *Magn Reson Med* 2018; **79**(6): 2996-3006.
27. Schofield MA, Zhu Y. Fast phase unwrapping algorithm for interferometric applications. *Opt Lett* 2003; **28**(14): 1194-6.
28. Schweser F, Deistung A, Lehr BW, Reichenbach JR. Quantitative imaging of intrinsic magnetic tissue properties using MRI signal phase: an approach to in vivo brain iron metabolism? *NeuroImage* 2011; **54**(4): 2789-807.
29. Wei H, Dibb R, Zhou Y, et al. Streaking artifact reduction for quantitative susceptibility mapping of sources with large dynamic range. *NMR in biomedicine* 2015; **28**(10): 1294-303.
30. Iglesias JE, Insausti R, Lerma-Usabiaga G, et al. A probabilistic atlas of the human thalamic nuclei combining ex vivo MRI and histology. *NeuroImage* 2018; **183**: 314-26.
31. Iglesias JE, Van Leemput K, Augustinack J, et al. Bayesian longitudinal segmentation of hippocampal substructures in brain MRI using subject-specific atlases. *NeuroImage* 2016; **141**: 542-55.
32. Iglesias JE, Van Leemput K, Bhatt P, et al. Bayesian segmentation of brainstem structures in MRI. *NeuroImage* 2015; **113**: 184-95.
33. Saygin ZM, Kliemann D, Iglesias JE, et al. High-resolution magnetic resonance imaging reveals nuclei of the human amygdala: manual segmentation to automatic atlas. *NeuroImage* 2017; **155**: 370-82.
34. Zhou G, Lane G, Cooper SL, Kahnt T, Zelano C. Characterizing functional pathways of the human olfactory system. *Elife* 2019; **8**.
35. Neudorfer C, Germann J, Elias GJB, Gramer R, Boutet A, Lozano AM. A high-resolution in vivo magnetic resonance imaging atlas of the human hypothalamic region. *Sci Data* 2020; **7**(1): 305.
36. Alfaro-Almagro F, McCarthy P, Afyouni S, et al. Confound modelling in UK Biobank brain imaging. *NeuroImage* 2021; **224**: 117002.
37. Papst I, Li M, Champredon D, Bolker BM, Dushoff J, DJ DE. Age-dependence of healthcare interventions for COVID-19 in Ontario, Canada. *BMC Public Health* 2021; **21**(1): 706.
38. Levin AT, Hanage WP, Owusu-Boaitey N, Cochran KB, Walsh SP, Meyerowitz-Katz G. Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications. *Eur J Epidemiol* 2020; **35**(12): 1123-38.
39. Ronnlund M, Nyberg L, Backman L, Nilsson LG. Stability, growth, and decline in adult life span development of declarative memory: cross-sectional and longitudinal data from a population-based study. *Psychol Aging* 2005; **20**(1): 3-18.
40. Staresina BP, Duncan KD, Davachi L. Perirhinal and parahippocampal cortices differentially contribute to later recollection of object- and scene-related event details. *J Neurosci* 2011; **31**(24): 8739-47.
41. Naya Y, Suzuki WA. Integrating what and when across the primate medial temporal lobe. *Science (New York, NY* 2011; **333**(6043): 773-6.
42. Doty RL. Olfaction: Smell of Change in the Air. *Cerebrum* 2017; **2017**.
43. Chikazoe J, Lee DH, Kriegeskorte N, Anderson AK. Distinct representations of basic taste qualities in human gustatory cortex. *Nat Commun* 2019; **10**(1): 1048.
44. Tsai ST, Lu MK, San S, Tsai CH. The Neurologic Manifestations of Coronavirus Disease 2019 Pandemic: A Systemic Review. *Front Neurol* 2020; **11**: 498.

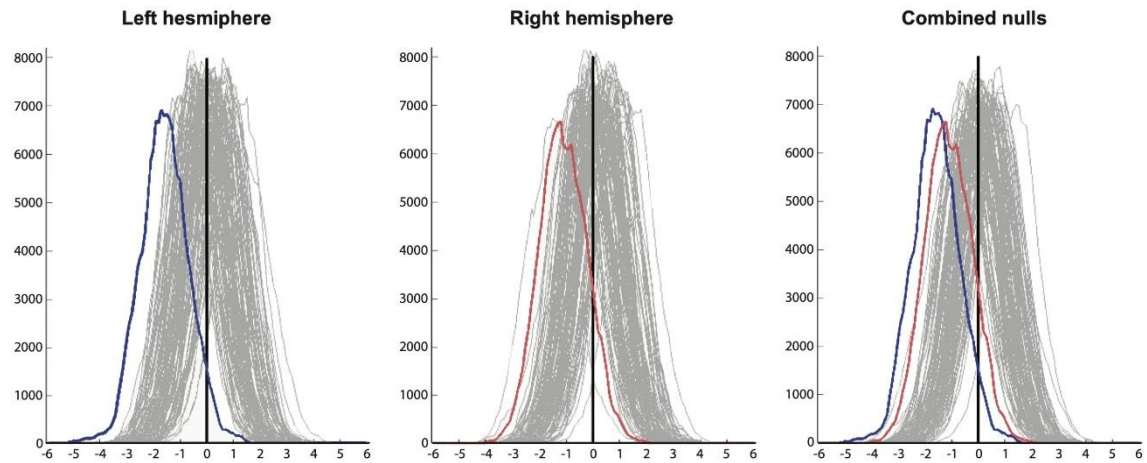
45. Douaud G, Jbabdi S, Behrens TE, et al. DTI measures in crossing-fibre areas: increased diffusion anisotropy reveals early white matter alteration in MCI and mild Alzheimer's disease. *NeuroImage* 2011; **55**(3): 880-90.
46. Douaud G, Refsum H, de Jager CA, et al. Preventing Alzheimer's disease-related gray matter atrophy by B-vitamin treatment. *Proceedings of the National Academy of Sciences of the United States of America* 2013; **110**(23): 9523-8.
47. Royet JP, Plailly J. Lateralization of olfactory processes. *Chem Senses* 2004; **29**(8): 731-45.
48. Guedj E, Lazarini F, Morbelli S, et al. Long COVID and the brain network of Proust's madeleine: targeting the olfactory pathway. *Clin Microbiol Infect* 2021.
49. Stamps JJ, Bartoshuk LM, Heilman KM. A brief olfactory test for Alzheimer's disease. *J Neurol Sci* 2013; **333**(1-2): 19-24.
50. Dinnes J, Deeks JJ, Berhane S, et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. *Cochrane Database Syst Rev* 2021; **3**: CD013705.

Supplementary Material

Supplementary Figure 1. Age distributions for COVID patients and controls at each timepoint do not differ significantly. Two-sample Kolmogorov-Smirnov was used to compute the P values for age comparisons, since age for each group was not normally distributed (Lilliefors P = 1e-03 for each group, and both age at Scan 1 or Scan 2). This showed no significant difference in age distribution between COVID patients and controls at Scan 1: p = 0.13 or at Scan 2: p = 0.09.



Supplementary Figure 2. Histograms of group comparison Z statistics of longitudinal change in cortical thickness. Left, histogram of Z-statistics (blue) across cortical vertices in left hemisphere (with grey lines showing 100 null histograms created through random permutations of the group variable). Middle: right hemisphere Z-statistics (orange) and matched nulls. Right: the same left and right hemisphere Z-statistics histograms overlaid (with a pooled null histogram in grey).



Supplementary Table 1. Full list of reproducible IDPs used in the hypothesis-driven approach, and corresponding statistics.

Please see additional attached PDF document.

Supplementary Table 2. Full list of reproducible IDPs used in the exploratory approach, and corresponding statistics. **Version 1**, ordered by IDPs. **Version 2**, ordered by decreasing |Z| from Model 3 group comparison (hospitalised vs non-hospitalised COVID).

Please see additional attached PDF documents.