Pre-Registration of Exploratory Questions for 'The effect of social race and socio-economic status on COVID-19 chemosensory loss'

BACKGROUND

SARS-CoV2 attaches to air pollutants ($PM^{2.5}$, PM^{10}, and NO_2) and differentially affects risk of infection [1–9] and mortality in individuals exposed to higher rates of pollution [4–10]. Loss of smell and/or taste in the absence of other symptoms is a key indicator of COVID-19 infection [11,12] but may be useless for communities already suffering from pre-existing smell and taste disorders due to the effects of environmental racism [13–15]—the pollution burden carried by racial minorities and socio-economically disadvantaged groups on local [16–21] and global scales [22] is disproportionate to the general population. US and UK COVID-19 mortality data race indicate that black populations are more than twice as likely to die than other groups [23] but there are no data for other aspects of infection or disease progression. Race is not genetically or biologically discrete but the human social environment is a key predictor of adult morbidity and mortality [24] and demographic differences in allostatic load (e.g., race, gender) predispose marginalized populations to greater sensitivity to chronic disease [25,26]. Pandemic management and planning must include plans for at-risk groups but despite these known differences, European countries ignore social race and clinical studies on COVID-19 in the UK and America detached race data from clinical status. The proposed study would be the first to attempt to understand the global experience of race and socio-economic status (SES) on pandemic dynamics.

Two GCCR papers analyzed data for a large body of case studies and clinical findings regarding taste and smell loss relative to COVID-19 but these papers did not include questions regarding race or SES. While there are limited data in the UK and US on race, these data are not systematically collected in hospitals and there are no representative samples that can inform hypothesis building. Further, these data do not explain the experience prior to death and there are no data on the impact of SES on COVID-19, beyond the general knowledge that lower SES equates to poorer health outcomes. One published study proposes that COVID-19 is a syndemic, interaction and exacerbation of pre-existing health inequities relative to a pandemic [27] and an online report indicates that marginalized groups are linked to lower indices of deprivation and health [28]. Further, given that this will be an unprecedented attempt to understand social race in a global context via self-reported data on individual local perceptions of race and self-reported relative status, an analysis plan will have to be open-ended. Due to the limited data available, the global scale of the study (with various understandings of race and SES), the study is not hypothesis driven other than the US sample and predicting COVID based on smell loss. The survey questions collect a variety of data that can be further explored to develop hypotheses to test in future efforts.

The primary research question is: Is smell loss predictive for COVID-19 when considering race and SES? We aim to identify if there are differences attributable to social race and SES and their interaction relative to COVID-19 and chemosensory loss by country or region (depending on sample sizes). We also ask questions about sexual orientation and gender due to recent reports that LGTBQIA+ populations may be experiencing greater rates of severe COVID-19

disease progression [29]. These questions are not required to answer due to the sensitivity of those topics in some cultures/countries.

DATA COLLECTION METHODS

The only method used for data collection is a survey. The survey considers cross-cultural variation in race and SES. We ask questions about income and resources and whether the participant is a marginalized member of their larger community. We provide definitions for all terms that are not common or may need clarification. We have two survey groups.

1) The first will be a FormR-based survey that will be advertised via a mass blind cc email to participants from a previous survey conducted by the Global Consortium for Chemosensory Research. The email will contain a link to our survey which will not collect their email address. The survey will also be advertised via social media to expand the sample size. We will contact various partners via email or social media to request they boost our signal. The social media text and email are included as part of the advertising/recruitment efforts in this package along with the electronic informed consent without signature and the survey questions themselves. The social media platforms we will use will include Twitter, Facebook, Whatsapp, and Instagram. Other social media platforms may be included for countries where these are not available. The only screening requirement for the survey is that participants must be 18 years of age or older.
2) The second survey group will be via Qualtrics, a professional survey company that we have paid ($3,000) for a guaranteed sample of 520 participants that each have an income below $41,000, with the sample split equally across our demographics of interest: white, black, and Hispanic. The same 18+ age requirement applies to this survey. Qualtrics pays participants a token (~$4.50) in exchange for their time). Because we have additional screening requirements for this survey, the FormR survey will be copied into Qualtrics survey software in two parts: The first part will contain questions directed to Americans (income in three brackets and race) and the second part will be the remainder of the survey questions.

For both forms, participants cannot retrieve answers to past responses. For FormR, participants get an access token for the study, which functions as a strong password. The access token is stored on participant's devices/browsers by default, so the only possible breach of protection is access to the device used to consent to the survey—but once the survey is submitted, the token is no longer active and the survey and answers cannot be retrieved. IP addresses are not automatically collected in either platform.

- The Qualtrics survey is operated by Qualtrics and will run until we get the 520 participants.
- The FormR survey will run until 15 January or until we have 520 participants. We will advertise for the FormR survey once a week. We will share the survey via GCCR Slack channels (key_announcemtns) and the GCCR newsletter.

PARTICIPANT CHARACTERISTICS

The only screening requirement for both surveys is that participants must be 18 years of age or older. The second survey group administered via Qualtrics, a professional survey company, has additional requirements: income below $41,000, race self-identified as white, black, or Hispanic.

1. Inclusion Criteria. All surveys: Adults over the age of 18 years only. Participants must click on a button to accept the terms and confirm they are 18+.
2. Qualtrics Specific inclusion criteria: must be white, black, or Hispanic. Must also have an income under 41,000 for a household of 3 or more.

INFORMED CONSENT

The conditions for consenting to participate in the survey are provided on the initial page of the survey. Once participants agree to continue, they can end their participation at any time and we will exclude partial submissions. Once the submit survey button is clicked, the data are recorded anonymously and we cannot retrieve and delete individuals responses nor can participants edit their responses. A copy of the language is found here: https://github.com/kchoover14/covidSESrace/blob/main/Informed%20Consent.pdf

ETHICS

A copy of the approval letter is in the Github repository for this study and that link is shared in the informed consent/information part of the survey. https://github.com/kchoover14/covidSESrace/blob/main/IRB%20Approval.pdf

ESTIMATE OF REQUIRED SAMPLE SIZE

We cannot estimate effect size for our analysis because there is no precedent for a global analysis of COVID symptoms relative to social race and SES, even if there are some studies in the US and the UK on those specific country demographics. The closest analysis available for estimation of sample size is a global study (also survey-based) on COVID-19 symptoms but that did not consider demographic factors such as race or SES [11]. Our target sample of 520 exceeds two separate calculations of power.

1) If we sample from a population of 100,000 (depending on the celebrities we engage or other social media groups we target) and target a 95% confidence level and 5% confidence interval and only 50% of population participate, the size of the required sample is 384 (https://www.surveysystem.com/sample-size-formula.htm).
2) If we sample from a population of 15,000, using a hypothesized frequency of smell loss among COVID positive people of 88% (based on [11]), the required sample size is 280 for 99% confidence level, and 456 for 99.9% confidence level. (OpenEpi - Toolkit Shell for Developing New Applications)

GLOBAL DATA ANALYSIS METHODS (item numbers are from Codebook)

Data Cleaning

We will exclude participants with the following implausible combinations of answers:

- reported smell and taste loss (32) but no change on 10-point scales before and during illness (40-41)

- reported smell loss (32) but no “smell changes” as a symptom during illness (37)
- reported “smell changes” as a symptom during illness (37), but no change on 10-point scales before and during illness (40,41)
- reported smell loss as a factor in seeking diagnosis (38), but no change on 10-point scales before and during illness (40,41)
- reported smell loss as a factor in seeking diagnosis, but no “smell changes” as a symptom during illness (38,37)
- reported more than two generations in household (19), but then gives a number lower than 3 on the question how many (20).
- Checked the box of being over 18 years old on the consent form and answered being under 18 years old (22, formR only)

Global Data Transformations

- Items 9-17: Eight questions on household income and resources are will be used to create two variables
    - Needs Index (69): Household needs relative to income and/or resources
        - Income that meets needs
        - Income with resources that meets needs
        - Income that does not meet needs
        - Income with resources that does not meet needs.
    - SES Index (70): Relative SES ranking (distribution, dimension reduction)
        - Within group with income that meets needs
        - Within country with income and resources that meets needs
        - Within group with income that does not meet needs
        - Within country with income and resources that does not meet needs
- Items 19-21 and 23: Depending on bin size, information on household (two+ generations, how many generations, rural versus urban) will be collapsed into one new variable with four categories.
    - Household Index (71)
        - Urban household with 1-2 generations
        - Urban household with 3+ generations
        - Rural household with 1-2 generations
        - Rural household with 3+ generations
- Item 21: Environmental exposure to agents that impede olfaction will be used to create a new variable
    - Environmental Index (72)
        - Count of options selected (0-4): a higher number indicates a higher risk
- Items 29-30: Depending on bin size, these may be collapsed into new variables. We may have too small a sample of non-cis gendered individuals and we may have too small a sample of non-heterosexual individuals.
- Item 34-36 on smell/taste loss cause and diagnosis will be collapsed into one new variable with three categories.
    - ST Loss Cause (73)
        - Loss attributable to non-viral illness

        - Loss attributable to COVID-19 viral illness
        - Loss attributable to non-COVID-19 viral illness
- Item 37 on viral illness will be split into two variables:
    - Symptom Index (74)
        - Count of symptom boxes ticked (1-13): a higher number indicates greater severity of illness
    - ST Loss (75)
        - smell/taste loss
        - no smell/taste loss
- Items 59-62 on odor awareness will be collapsed into one variable
    - Odor Awareness Index (77)
        - We will perform a factor analysis and use a scree plot to determine the number of factors underlying this set of items.
            - On items 60-62, the answers “not applicable” and “never” will both be coded as 1.
            - The answers “Seldom”, “Sometimes”, “Often”, “Always” will be recoded to 2,3,4,5 respectively.
            - If a single underlying factor is found, we will sum the four recoded 4 odor awareness items 59-62 into a single score from 4-20.
            - If more than one factor is found, we will sum the items that load onto the same factor and proceed with odor awareness subscales in our analyses.
- Items 63-67 on health hazards that impede olfaction will be used to create a new variable
    - Health Index (78)
        - Count of options selected (0-5): a higher number indicates a higher risk
- Items 40 (olfactory ability before illness) and 41 (olfactory ability during illness) will be collapsed into one variable
    - Olfactory Change Index (80)
        - Olfactory ability during illness MINUS olfactory ability before illness
- Items 41 (olfactory ability during illness) and 42 (current olfactory ability) will be collapsed into one variable
    - Recovery index (81)
        - Current olfactory ability MINUS olfactory ability during illness
- Items 40 (olfactory ability before illness) and 43 (current olfactory ability) will be collapsed into one variable
    - Long haul index (82)
        - Current olfactory ability MINUS olfactory ability before illness

Global Analysis Parameters

- Alpha = 0.05
- Standard Bonferroni correction in the case of multiple dependent tests, within paper.

COVID GEOGRAPHY AND HOUSEHOLD

Justification

Significant variation in host predisposition to COVID-19 has been documented around the globe.[30,31] Many studies have found that race and SES are important factors in this predisposition, however it has also been shown that household size distribution and regional dynamics can cause significant differences in incidence.[30–33] Further elucidating the household and regional dynamics that influence this variation, beyond what can be described by differences in viral variants, will provide insights into ameliorating the factors that lead to differences of incidence.

Research Questions

1. Does the Household Index affect the incidence of COVID-19differentially?
    a. We predict that multi-generational households and larger households will have a greater risk for COVID-19 infection.
    b. Regression on infection dependent on Household Index
2. Is there a difference in COVID-19 rates based on Location (urban vs. rural) and Environment Index?
    a. We expect both independent factors will have an impact and there will be an interaction between independent factors
        i. The impact of environmental exposure is greater than location because urban and rural places vary internally in pollution levels.
        ii. Location may have an independent impact but also may not. While rural areas are less crowded which reduces disease transmission, rural areas are less likely to be vaccinated which might mitigate the advantage of a more sparse population distribution.
    b. Logistic regression on infection dependent on Location and Household Index (and interaction).

Factors to Consider in Analysis

- In our sample, how do race and SES or location (urban/rural) relate? Is the relationship reflective of trends in the US population?
- What proportion of respondents that select each race category reside in an urban location? In a rural location? Is there enough individuals for comparisons
-

ORAL HEALTH AND COVID CONSIDERING SES

Xerostomia (dry mouth report) is one of prevalent and persistent oral symptoms associated with COVID-19, however, unlike taste dysfunction, xerostomia tends to be overlooked despite the potential negative impact of xerostomia on the oral-health related quality of life [34,35].

Good oral health has been linked to risk reduction of viral acute respiratory diseases[36,37] and the lower severity of COVID [38]. Studies have shown an impact of SES on oral dryness and the oral health of an individual [39,40].

Factors to Consider in Analysis

- Create a new variable from Item 37: Symptoms (Dry Mouth) with two levels: those with and those without dry mouth as a symptom (Dry Mouth Index, item 79).

Research Questions and Hypotheses

1. To what degree does oral health (item 68) and SES (Needs Index 69, SES Index 70) influence viral illness (ST Loss Cause 73)?
    a. We predict viral illness rates will be higher in those with poor oral health compared to good oral health.
    b. Regression with viral illness as dependent on oral health and SES and interaction between oral health and SES.
    c. Regression with viral illness as dependent on oral health and SES.
2. To what degree does oral health (item 68) and SES (Needs Index 69, SES Index 70) influence COVID-19 (ST Loss Cause 73)?
    a. We predict COVID-19 rates will be higher in those with poor oral health compared to good oral health.
    b. Regression with COVID-19 as dependent on oral health and SES and interaction between oral health and SES.
    c. Regression with COVID-19 as dependent on oral health and SES.
3. Is there a relationship between SES (Needs Index 69, SES Index 70) and the symptom of *dry mouth (item 79)* ?
    a. We predict that lower SES (Needs Index 69, SES Index 70) has greater frequency of dry mouth.
    b. Regression with Dry Mouth as dependent on SES.

TASTE PERCEPTION AND COVID

Loss of taste and smell is associated with COVID-19 infection [11,12]. A study using self-reported data from a web-based questionnaire of COVID-19 patients (n=128) reported that 52% changed in their spicy taste perception, 54 in the salty taste, 53 in sour taste, and 61 in sweet taste [41]. The reported changes in taste perception were gustation (taste) and trigeminal effects (spicy effects) but results did not include complex foods that are dominated by a specific taste quality like bitter/sour or sweet.

Research Question

1. What patterns of change do we see in food liking (Items 44-57) as a result of losing your sense of smell and/or taste (Item 32)?
2. What patterns of change do we see in food liking (Items 44-57) as a result of the amount of change in smell (Item 80) and recovery from illness (Item 81)?