

## Supplementary materials

### Table of Contents

Definitions of predictor variables.....	1
Derivation of sampling weights.....	3
Supplementary Table 1.....	4
Demographic composition of staff responding to survey vs. composition of population.....	4
CHERRIES checklist.....	5
References.....	9

### Definitions of predictor variables

We explored six groups of predictor variables:

#### 1. *Demographics and prior mental health*

Demographic variables included age, gender, ethnic group, and birthplace. Participants were able to define their gender as ‘Female,’ ‘Male,’ or ‘Other,’ as well as indicating they would ‘Prefer not to say.’ For the purposes of modelling, where small numbers in the latter two categories would be problematic, we have randomly allocated these responses to ‘Female’ or ‘Male’ based on the sample proportions. To avoid small cell sizes, ethnicity was coded into five categories following recommendations of the Office for National Statistics [1]: White, Mixed, Asian/Asian British, Black/African/Caribbean/Black British, or Other ethnic group. Birthplace was categorised as ‘United Kingdom,’ ‘European Union (not United Kingdom)’ or ‘Other.’

Prior mental health was assessed using the question ‘Which of the following have you ever experienced (include only those that have been diagnosed by a doctor, psychiatrist or other health professional)?’ Using responses of ‘Depression’ for history of depression and ‘Generalised anxiety disorder,’ ‘Panic attacks’ or ‘Post-traumatic stress disorder’ for history of an anxiety disorder.

#### 2. *Living arrangements*

Participants were asked “Which of the following best describes your current living arrangement?” Responses were dichotomised to indicate whether participants were ‘Living

alone' (the participant reported living "Alone") vs. 'living with others.' Housing tenure was measured as any 'renting' category vs. all other categories.

### 3. *Caring roles*

The number of children living in the household was measured as an ordinal variable (coded as 0, 1, 2 or 3+ children) and other caregiving responsibilities besides childcare was measured as a binary variable (0 = 'No'; 1 = 'Yes').

### 4. *Healthcare*

We considered three variables measured healthcare. (i) Difficulties accessing healthcare from 'Have you needed to access healthcare services since the beginning of the pandemic for reasons other than COVID-19 (coronavirus)?' and 'Were you able to access that service?' categorised as 'Did not need to access healthcare,' 'Needed healthcare and was able to access' and 'Needed healthcare and was not able to access.' (ii) Presence of a long-term condition from 'Do you have any physical conditions or illnesses lasting or expected to last 12 months or more?' (iii) Reported self-isolating or shielding were also included in the healthcare category. 'Self-isolating' was defined as "avoiding contact with others and not leaving the home for any reason". 'Shielding' was defined in the questionnaire as "a type of self-isolation, which involves not leaving your home for any reason for at least 12 weeks to reduce your risk of contracting COVID-19 (coronavirus)".

### 5. *Occupational factors*

Four occupational factors were measured for staff but not students. These included (i) whether the respondent was in a key worker role ('Are you currently fulfilling a 'key worker' role as identified by the government?') and (ii) whether they were on a temporary contract (either 'Fixed term' or 'Casual' contract types). (iii) Participants newly working remotely were identified using questions about working remotely before the pandemic ('No' or 'Occasionally' vs. more often) and since the pandemic. (iv) Staff role, based on participants' stated primary role ("Which of the following best describes your primary role as a member of staff at King's College London?"), was organised into three categories reflecting seniority and the degree of contact with the public: 'Academic, specialist and management,' 'Research, clerical and technical,' and 'Teaching, facilities, and clinical.' These categorises were defined as shown below:

Responses recorded in questionnaire		Derived category used in analyses
Academic	→	"Academic, specialist and management"
Specialists and professionals		

Management		
Research	→	“Research, clerical and technical”
Clerical		
Technical		
Teaching	→	“Teaching, facilities and clinical”
Facilities		
Clinical		

6. ‘COVID-19’ perceived infection status was defined using the question ‘Do you think that you have had COVID-19 (coronavirus) at any time?’ with responses dichotomised as ‘Definitely’ or ‘Probably’ vs. ‘No’ or ‘Unsure.’

### Derivation of sampling weights

Sampling weights were calculated for staff and PGR students separately. Information on age group, gender, and ethnicity for the target population (all KCL staff and PGR students, respectively) were extracted from administrative records. The procedure to derive weights was:

1. We harmonised the categories in administrative records with those measured in the survey.
2. Missing information in the survey data (for age group, gender, and ethnicity) was imputed with k-nearest neighbours (with k=5 nearest neighbours) using the kNN function from the VIM [2] package for R [3].
3. Weights (W) were derived based on age group, gender, and ethnicity using raking (iterative proportional fitting) using the rake function of the survey package [4] for R.
4. Extreme values were trimmed to  $W_T$ , where:

$$W_T = \text{median}(W) + (5 * \text{IQR}(W))$$

(IQR = inter quartile range). This involved setting any weight greater than  $W_T$  to  $W_T$ .

**Supplementary Table 1***Demographic composition of staff responding to survey vs. composition of population*

Unweighted percentages.

		Staff		PGR students	
		Admin. records	Baseline survey	Admin. records	Baseline survey
		%	%	%	%
Gender	Female	55	69	57	72
	Male	45	31	43	28
Age group (Staff)	16 - 25	5	4	--	--
	26 - 35	35	34	--	--
	36 - 45	28	29	--	--
	46 - 55	18	19	--	--
	56 - 65	11	11	--	--
	>=66	3	3	--	--
Age group (PGR students)	18 - 24	--	--	18	14
	25 - 29	--	--	39	47
	30 - 39	--	--	31	25
	40+	--	--	12	13
Ethnic group	White	70	87	61	78
	Black	5	1	4	2
	Asian	12	6	23	12
	Mixed	4	4	5	3
	Other	8	2	7	5

## CHERRIES checklist

A CHERRIES checklist (Checklist for Reporting Results of Internet E-Surveys) [5] is provided below.

Checklist Item	Explanation	Page Number
Describe survey design	Describe target population, sample frame. Is the sample a convenience sample? (In “open” surveys this is most likely.)	4, 5.
IRB approval	Mention whether the study has been approved by an IRB.	Described in [6].
Informed consent	Describe the informed consent process. Where were the participants told the length of time of the survey, which data were stored and where and for how long, who the investigator was, and the purpose of the study?	Described in [6].
Data protection	If any personal information was collected or stored, describe what mechanisms were used to protect unauthorized access.	Described in [6].
Development and testing	State how the survey was developed, including whether the usability and technical functionality of the electronic questionnaire had been tested before fielding the questionnaire.	Described in [6].
Open survey versus closed survey	An “open survey” is a survey open for each visitor of a site, while a closed survey is only open to a sample which the investigator knows (password-protected survey).	This was an open survey, with the restriction that invitations to participate were sent to participants’ university email account.
Contact mode	Indicate whether or not the initial contact with the potential participants was made on the Internet. (Investigators may also send out questionnaires by mail and allow for Web-based data entry.)	See page 5.
Advertising the survey	How/where was the survey announced or advertised? Some examples are offline media (newspapers), or online (mailing lists – If yes, which ones?) or banner ads (Where were these banner ads posted and what did they look like?). It is important to know the wording of the announcement as it will heavily influence who chooses to participate. Ideally the survey announcement should be published as an appendix.	The survey was advertised internally via newsletters and internal social media. Please refer to [6] for details.
Web/E-mail	State the type of e-survey (eg, one posted on a Web site, or one sent out through e-mail). If it is an e-mail survey, were the responses entered manually into a database, or was there an automatic method for capturing responses?	See page 5. Responses were automatically captured using online survey software. Please refer to [6] for details.
Context	Describe the Web site (for mailing list/newsgroup) in which the survey was posted. What is the Web site about, who is visiting it, what are visitors normally looking for? Discuss to	Not applicable.

	what degree the content of the Web site could pre-select the sample or influence the results. For example, a survey about vaccination on a anti-immunization Web site will have different results from a Web survey conducted on a government Web site	
Mandatory/voluntary	Was it a mandatory survey to be filled in by every visitor who wanted to enter the Web site, or was it a voluntary survey?	Voluntary.
Incentives	Were any incentives offered (eg, monetary, prizes, or non-monetary incentives such as an offer to provide the survey results)?	No.
Time/Date	In what timeframe were the data collected?	5.
Randomization of items or questionnaires	To prevent biases items can be randomized or alternated.	No.
Adaptive questioning	Use adaptive questioning (certain items, or only conditionally displayed based on responses to other items) to reduce number and complexity of the questions.	Yes.
Number of Items	What was the number of questionnaire items per page? The number of items is an important factor for the completion rate.	This was dependent on the questionnaire instrument and device being used. On average, $\approx 3$ items were displayed per page.
Number of screens (pages)	Over how many pages was the questionnaire distributed? The number of items is an important factor for the completion rate.	This was dependent on the questionnaire instrument and device being used. On average, $\approx 10$ pages were displayed.
Completeness check	It is technically possible to do consistency or completeness checks before the questionnaire is submitted. Was this done, and if “yes”, how (usually JavaScript)? An alternative is to check for completeness after the questionnaire has been submitted (and highlight mandatory items). If this has been done, it should be reported. All items should provide a non-response option such as “not applicable” or “rather not say”, and selection of one response option should be enforced.	Yes, this was provided by the survey software ( <i>Qualtrics</i> ).
Review step	State whether respondents were able to review and change their answers (eg, through a Back button or a Review step which displays a summary of the responses and asks the respondents if they are correct).	Yes.
Unique site visitor	If you provide view rates or participation rates, you need to define how you determined a unique visitor. There are different techniques available, based on IP addresses or cookies or both.	Not applicable, view rates were not provided.
View rate (Ratio of	Requires counting unique visitors to the first page of the survey, divided by the number of unique site visitors (not page	Not applicable.

unique survey visitors/unique site visitors)	views!). It is not unusual to have view rates of less than 0.1 % if the survey is voluntary.	
Participation rate (Ratio of unique visitors who agreed to participate/unique first survey page visitors)	Count the unique number of people who filled in the first survey page (or agreed to participate, for example by checking a checkbox), divided by visitors who visit the first page of the survey (or the informed consents page, if present). This can also be called “recruitment” rate.	Not applicable.
Completion rate (Ratio of users who finished the survey/users who agreed to participate)	The number of people submitting the last questionnaire page, divided by the number of people who agreed to participate (or submitted the first survey page). This is only relevant if there is a separate “informed consent” page or if the survey goes over several pages. This is a measure for attrition. Note that “completion” can involve leaving questionnaire items blank. This is not a measure for how completely questionnaires were filled in. (If you need a measure for this, use the word “completeness rate”.)	Not applicable.
Cookies used	Indicate whether cookies were used to assign a unique user identifier to each client computer. If so, mention the page on which the cookie was set and read, and how long the cookie was valid. Were duplicate entries avoided by preventing users access to the survey twice; or were duplicate database entries having the same user ID eliminated before analysis? In the latter case, which entries were kept for analysis (eg, the first entry or the most recent)?	Yes, cookies were assigned to unique users upon consenting to take part in the study. Cookies were valid for 7 days, and were checked on each page to validate the questionnaire completer.
IP check	Indicate whether the IP address of the client computer was used to identify potential duplicate entries from the same user. If so, mention the period of time for which no two entries from the same IP address were allowed (eg, 24 hours). Were duplicate entries avoided by preventing users with the same IP address access to the survey twice; or were duplicate database entries having the same IP address within a given period of time eliminated before analysis? If the latter, which entries were kept for analysis (eg, the first entry or the most recent)?	Not applicable.
Log file analysis	Indicate whether other techniques to analyze the log file for identification of multiple entries were used. If so, please describe.	No.
Registration	In “closed” (non-open) surveys, users need to login first and it is easier to prevent duplicate entries from the same user. Describe how this was done. For example, was the survey never displayed a second time once the user had filled it in, or was the username stored together with the survey results and later eliminated? If the latter, which entries were kept for	Not applicable.

	analysis (eg, the first entry or the most recent)?	
Handling of incomplete questionnaires	Were only completed questionnaires analyzed? Were questionnaires which terminated early (where, for example, users did not go through all questionnaire pages) also analyzed?	Incomplete questionnaires were analysed.
Questionnaires submitted with an atypical timestamp	Some investigators may measure the time people needed to fill in a questionnaire and exclude questionnaires that were submitted too soon. Specify the timeframe that was used as a cut-off point, and describe how this point was determined.	Respondents had to complete the questionnaire within two weeks of consenting to participate and receiving the survey link via email.
Statistical correction	Indicate whether any methods such as weighting of items or propensity scores have been used to adjust for the non-representative sample; if so, please describe the methods.	See page 8 and supplementary materials.

## References

- 1 ONS. Ethnic group, national identity and religion - Office for National Statistics. Ethnic group, national identity and religion. 2020.<https://www.ons.gov.uk/methodology/classificationsandstandards/measuringequality/ethnicgroupnationalidentityandreligion#ethnic-group> (accessed 10 Dec 2020).
- 2 Kowarik A, Templ M. Imputation with the R package VIM. *Journal of Statistical Software* 2016;**74**:1–16. doi:10.18637/jss.v074.i07
- 3 R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: 2020. <https://www.R-project.org/>
- 4 Lumley T. Analysis of complex survey samples. *Journal of Statistical Software* 2004;**9**:1–19.
- 5 Eysenbach G. Improving the Quality of Web Surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *J Med Internet Res* 2004;**6**. doi:10.2196/jmir.6.3.e34
- 6 Davis KAS, Stevelink SAM, Al-Chalabi A, *et al*. The King's College London Coronavirus Health and Experiences of Colleagues at King's Study (KCL CHECK) protocol paper: a platform for study of the effects of coronavirus pandemic on staff and postgraduate students. *medRxiv* 2020;:2020.06.16.20132456. doi:10.1101/2020.06.16.20132456