

# Testing behaviour and positivity for SARS-CoV-2 infection: Insights from web-based participatory surveillance

Scott A. McDonald (✉ [scott.mcdonald@rivm.nl](mailto:scott.mcdonald@rivm.nl))

Netherlands National Institute for Public Health and the Environment

**Loes Soetens**

Netherlands National Institute for Public Health and the Environment

**Maarten Schipper**

Netherlands National Institute for Public Health and the Environment

**Ingrid H. M. Friesema**

Netherlands National Institute for Public Health and the Environment

**Cees C. van den Wijngaard**

Netherlands National Institute for Public Health and the Environment

**Anne Teirlinck**

Netherlands National Institute for Public Health and the Environment

**Nienke Neppelenbroek**

Netherlands National Institute for Public Health and the Environment

**Susan van den Hof**

Netherlands National Institute for Public Health and the Environment

**Jacco Wallinga**

Netherlands National Institute for Public Health and the Environment

**Albert Jan van Hoek**

Netherlands National Institute for Public Health and the Environment

---

## Research Article

**Keywords:** COVID-19, participatory surveillance, testing, positivity, behaviour, Netherlands

**Posted Date:** July 6th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-688708/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at BMJ Open on December 1st, 2021. See the published version at <https://doi.org/10.1136/bmjopen-2021-056077>.

## **Testing behaviour and positivity for SARS-CoV-2 infection: Insights from web-based participatory surveillance**

Scott A. McDonald<sup>\*</sup>, Loes Soetens<sup>1</sup>, Maarten Schipper<sup>1</sup>, Ingrid H. M. Friesema<sup>1</sup>, Cees C. van den Wijngaard<sup>1</sup>, Anne Teirlinck<sup>1</sup>, Nienke Neppelenbroek<sup>1</sup>, Susan van den Hof<sup>1</sup>, Jacco Wallinga<sup>1,2</sup>, Albert Jan van Hoek<sup>1</sup>

<sup>1</sup> Centre for Infectious Disease Control, Netherlands National Institute for Public Health and the Environment, Bilthoven, the Netherlands

<sup>2</sup> Department of Biomedical Data Sciences, Leiden University Medical Center, Leiden, the Netherlands

\* Corresponding author.

Email: [scott.mcdonald@rivm.nl](mailto:scott.mcdonald@rivm.nl)

*Conflict of interest declaration:*

The authors declare that they have no competing interests.

## **ABSTRACT**

### **Background**

Voluntary testing for SARS-CoV-2 infection is an integral component of an effective response to the COVID-19 pandemic. It is essential to identify populations at a high risk for infection but who are less likely to present for testing. Here, we use internet-based participatory surveillance data from the Netherlands to identify sociodemographic and household factors that are associated with a lower propensity to be tested and, if tested, with a higher risk of a positive test result.

### **Methods**

Multivariable analyses using generalised estimating equations for binomial outcomes were conducted to estimate the adjusted odds ratios of testing and of positivity associated with participant and household characteristics.

### **Results**

Based on five months (17 November 2020 to 18 April 2021) of weekly surveys obtained from 12,026 participants, males (adjusted odds ratio for testing ( $OR_t$ ): 0.92; adjusted odds ratio for positivity ( $OR_p$ ): 1.30, age-groups <20 ( $OR_t$ : 0.89;  $OR_p$ : 1.27) 50-64 years ( $OR_t$ : 0.94;  $OR_p$ : 1.06) and 65+ years ( $OR_t$ : 0.78;  $OR_p$ : 1.24), diabetics ( $OR_t$ : 0.97;  $OR_p$ : 1.06), and sales/administrative employees ( $OR_t$ : 0.93;  $OR_p$ : 1.90) were distinguished as lower propensity/higher positivity factors.

### **Conclusions**

The factors identified using this approach can help identify potential target groups for improving communication and encouraging testing among those with symptoms and thus increase the effectiveness of testing, which is essential for the response to the COVID-19 pandemic and for public health strategies in the longer term.

## INTRODUCTION

The pillars of effective control of the acute respiratory disease COVID-19 are testing, contact tracing, and isolation after a positive test [1-3]. It therefore is vital to identify those characteristics that are associated with willingness (or aversion) to undergo testing, given one or more symptoms compatible with COVID-19, especially for those persons not already identified through contact tracing. Furthermore, knowledge of the characteristics that are most strongly associated with SARS-CoV-2 positivity will optimise interpretation of test data and assist in helping public health communication in the targetting of advice to individuals that should get tested and/or self-isolate.

For studies of the factors associated with infectious disease risk in the community, internet-based participatory surveillance is a valuable tool, as demonstrated by studies of influenza-like illness in the Netherlands and elsewhere [4,5] and new app-based monitoring COVID-19 of symptoms and other outcomes [9,10,11]. To gain insight into sociodemographic, participant and household factors associated with testing for SARS-CoV-2 infection, we rely on data from the Infectieradar web-based participatory surveillance system for COVID-19 in the Netherlands, which was set up to monitor a number of epidemiological outcomes, including the occurrence of symptom(s), PCR/antigen test behaviour and test results. Other existing surveillance systems in the Netherlands can only collect information on those persons who get tested, as persons who would be eligible, but do not present for testing, are not reached.

Using Infectieradar data, we investigated the associations between participant/household characteristics and the propensity to get tested for SARS-CoV-2 infection, given symptoms. We also quantified the associations between the same factors and a positive test result among those tested, irrespective of symptoms. By combining the results of both analyses, it is possible to identify persons in certain demographic categories or with other risk factors who are less likely to present for testing, but have a higher risk of testing positive.

## METHODS

### **Study setting, analysis period and participants**

We conducted a cohort study based on approximately five months of Infectieradar data collected between 17 November 2020 and 18 April 2021. As of the study period end date, the total number of active users (defined as those who had filled out at least two weekly surveys since the date when Infectieradar data collection was re-started in the autumn of 2020 (i.e., 2 October 2020) was 17,054. We began analysis on 17 November 2020, as testing eligibility policy prior to this data excluded children younger than 12 years from testing unless severely ill.

### **Data source**

Participants were recruited for the Infectieradar surveillance system (a member of the Influenzanet consortium, a European collaboration involving universities and public health partners) was conducted via a web announcement [6]. Registration was open to all residents of the Netherlands. Children younger than 16 years could participate under supervision of their legal guardian, or if their legal guardian acted on their behalf. Registration involved filling out an intake questionnaire, which asks for sociodemographic data (age, sex, education level, occupation, partial postal code, number and age make-up of persons in household, etc), and medical history (e.g., pre-existing health conditions such as allergies/hay fever and chronic diseases). These elicited data were selected in line with previous research conducted by the Influenzanet consortium [5]. In addition, the intake survey queried if the participant had ever received a positive PCR/rapid test result, and the date the test was conducted.

After completing the intake survey, participants were requested to fill out a standard questionnaire (and every week thereafter), to report any symptoms experienced during the past 7-day period. Namely, in each weekly survey participants were asked to report the occurrence of one or more of a set of 21 symptoms (fever, chills, runny nose, sneezing, sore throat, cough, dyspnea (shortness of breath), headache, muscle/joint pain, chest pain, malaise, loss of appetite, coloured phlegm, watery or bloodshot eyes, nausea, vomiting, diarrhoea, stomach-ache, loss of sense of smell, loss of sense of taste, or nosebleed). If a participant had been tested for SARS-CoV-2 infection since their previous survey (irrespective of symptoms), they were asked to report the type of test and the test outcome.

Additional weekly survey data included information regarding healthcare-seeking behaviour and suspected cause of symptom(s) (if reported). All data were pseudonymised, with individual participants assigned a unique identifier used in subsequent analysis.

### **Ethical approval**

The research protocol was shared with the Medical Ethics Review Committee Utrecht, and an official waiver for ethical approval was obtained, which stated that ethical assessment was unnecessary given the non-invasive nature of data collection (reference number: WAG/avd/20/008757).

### **Statistical analysis**

*Data inclusion/exclusion criteria.* We attempt to reduce the impact of selection bias caused by persons who registered or only participated once because they recently experienced symptoms, by (i) excluding participants who had contributed fewer than two weekly surveys, and (ii) removing the first weekly survey submitted by each participant. Prevalent illness at time of registration was addressed by excluding all weekly surveys in which symptom onset was indicated as before the date of intake survey. Surveys for persons with missing age or sex ( $n=24$ ) were excluded.

*Regression analysis.* We fitted univariate and multivariable binomial generalized estimating equations to the two distinct outcomes (i) test propensity (among all participants who reported at least one symptom from the set of 21 symptoms queried), and (ii) test positivity (among all participants reporting being tested since their last survey). We did not restrict (ii) to only those surveys in which symptoms were reported because the participant could have recovered by the time they reported having been tested (the test date may have been more than one week previously). For (i) we were interested in the associations between participant/household characteristics and test propensity (conditioning on symptoms reduces the influence of factors such as requests for testing generated by contact tracing or to be released from quarantine); for (ii) we were interested in the associations between participant/household characteristics and a positive result, conditioned on being tested.

Besides sex, age-group, education level (higher, middle, none/lower or missing) and underlying conditions (asthma, allergies/hay fever, chronic lung disease, diabetes, cardiovascular disease), smoker status (never; current/ex-smoker), children aged <5 years in household, children 5-18 years in household, and occupation category (see below), the covariate set for outcome (i) only included covariates for test history (defined with respect to the date of each weekly survey): *Never*, *previously reported negative result*, *previously reported positive result* (including a positive result reported at time of intake), and 'suspected non-COVID-19 cause' of reported symptom(s). The latter variable was defined as *Yes* if any of the answers 'flu/flu-like illness', 'common cold', or 'allergies/hay fever/asthma', 'gastrological complaints/stomach flu', or 'other' were selected for the question 'Do you have any idea what caused your symptoms?', and as *No* otherwise (thus 'No' included the responses 'coronavirus (COVID-19)' and 'don't know'). For outcome (ii) only, the covariate logarithm-transformed total number of reported symptoms (in that weekly survey) was added. This transformation was selected based on prior assessment of model fit. We defined six categories for the occupation covariate: *Not applicable* (preschool, pupil, student, household, unemployed, retired); *healthcare, education* (including day-care staff), *knowledge worker* (e.g., manager, accountant, scientist), *sales/administration* (e.g., shop/supermarket staff, receptionist, administrator, financial assistant), and *all other occupations*. In regression analysis, *knowledge worker* – the occupation category with the greatest proportion of participants – was defined as the reference category.

One model for each outcome was fitted retaining only participant/other factors (excluding symptoms) as covariates. A natural spline was fitted in each model to capture temporal trends in test uptake (five knots) and test positivity (three knots) that is not explained by the other covariates. We used a generalised estimating equations approach with exchangeable correlation structure to correct standard errors for repeated observations per participant.

The predicted absolute risk of getting tested for each variable was estimated from the fitted regression model using marginal standardisation [7]; i.e., producing the predicted proportion tested, taking into account the distribution of other variables in the model.

To bring together both sets of association measures for test propensity and positivity, we graphically categorised participant characteristics associated with a low propensity to get tested for SARS-CoV-2, but a relatively high risk of a positive outcome. Finally, we calculated the positive predictive value (PPV) for each of the 21 possible symptoms that could be reported, to determine if the efficiency of the above-identified participant factors could be improved by adding reported symptom(s).

All analyses were conducted using R statistical software, version 3.6.0 [8].

## RESULTS

During the five-month analysis period between 17 November 2020 and 18 April 2021, 282,277 weekly surveys were submitted by 16,807 unique Infectieradar participants. 12,026 individuals reported at least one symptom in 50,946 weekly surveys. Among these 50,946 surveys, undergoing a PCR/antigen test and receipt of the test outcome since the previous survey was reported in 8997 surveys (51% of participants) (Table 1); 48% of ever-tested individuals were tested twice or more. A previous negative test result (in 42.0% of surveys) or previously testing positive (in 7.6%) was reported. Females (57.2%) and those with a higher education level (60.7%) were over-represented in Infectieradar and the age-distribution was older (median 53.7 years) compared with the Netherlands general population (Supplementary Materials, Table S1).

The adjusted odds ratios (ORs) for the propensity to test associated with participant/household factors from all surveys with at least one reported symptom are provided in Table 1 (see also Fig. 1). The strongest positive associations were observed for: suspected non-COVID-19 cause (OR=1.45 (95% CI: 1.37-1.53)), children <5 years in household (1.41 (1.29-1.54)), children 5-18 years in household (1.22 (1.14-1.31)), *occupation: healthcare* (1.19 (1.08-1.30) (with *knowledge worker* as reference), age 20-29 years (1.19 (1.05-1.35)), and age 30-39 years (1.18 (1.08-1.29) (compared with the reference category 40-49 years). The lowest odds of being tested were for: previously tested positive (0.35 (0.30-0.42)), allergy sufferers (0.77 (0.73-0.81)), lower education level (0.78 (0.63-0.98), compared with middle education level), and age 65+ years (0.78 (0.69-0.89)). To estimate of the absolute risk

of testing per participant/household category among Infectieradar participants, we calculated the model-predicted test propensity using marginal standardisation. Among all participant/household characteristics, the highest predicted propensity was observed for age-group 20-29 years (22.5%), with the lowest predicted propensity for lower education level and age-group 65+ (16.2% for both).

Among all those who reported having been tested since their previous weekly survey, irrespective of the presence of symptoms ( $n=13,219$  surveys), 6.5% ( $n=854$ ) had a positive outcome; among only those tested with reported symptom(s), 8.7% ( $n=783/9008$  surveys) were positive. The associations between participant/household factors and a positive test result among all tested participants (whether or not reporting symptoms) are shown in Fig. 2. The highest odds of testing positive were for: (log) total number of symptoms (6.71 (5.58-8.07)), *occupation: sales/admin* (1.90 (1.35-2.68), compared with *knowledge worker*), children 5-18 years in household (1.34 (1.07-1.68), males (1.30 (1.05-1.59)), and age-group <20 years (1.27 (0.88-1.85)), though OR was not significant. The lowest odds of testing positive were for: ever-smoker (0.64 (0.46-0.91)), age-group 30-39 (0.65 (0.48-0.89), compared with 40-49 years), and higher education level (0.77 (0.63-0.94)).

To bring together both sets of association measures for test propensity and positivity, we categorised participant characteristics associated with a low propensity to get tested for SARS-CoV-2, but a relatively high risk of a positive outcome (Fig. 3). These factors were identified as: males, age-groups 65+, 50-64, and <20 years (compared with 40-49), diabetes, and *occupation: sales/admin* (compared with *knowledge worker*).

These identified risk factors are by definition associated with a higher probability of a positive outcome. We next calculated the PPV for each symptom, conditional on each of the identified participant factors (Fig. S1). For all factors, the highest PPVs were estimated for *loss of taste* (range of 50.0% to 78.6% over the six factors) and for *loss of smell* (44.4% to 77.8%). Other symptoms with higher PPVs were *loss of appetite* for males and the 50-64 and 65+ years age-groups (34.5%, 38.2% and 46.2%, respectively) and *fever* for *occupation: sales/admin* and diabetics (42.2%, 35.9%). Across the subpopulations defined by our identified risk factors,

very low PPVs were calculated for the mild respiratory symptoms *runny nose* (8.0 to 14.9%), *sneezing* (5.4 to 17.2%), and *sore throat* (4.0 to 11.1%).

Finally, in sensitivity analysis we restricted the dataset to participants who fell within the approximate age-range of the working population, defined as 20-64 years ( $n=41,017$  weekly surveys). An overlapping set (excluding the two extreme age-groups) of low propensity/high positivity risk factors was identified (Supplementary Materials, Fig. S2); the main difference was that *diabetes* fell outside of the area of interest.

## DISCUSSION

We identified the participant characteristics male sex, younger and older age, diabetes, and the occupation category *sales/administration* (relative to *knowledge worker*) to be associated with a lower propensity to test given symptoms, but with a higher risk of a positive outcome. Among all subpopulations characterised by these factors, the symptoms *loss of taste* and *loss of smell* had the highest predictive value of a positive outcome (Supplementary Materials, Fig. S1), confirming previous results for the general population [9,10], with *loss of appetite* (from age 50 years) also improving predictions.

Ever-smokers and allergies/hay-fever were participant characteristics that identified a low propensity to test with a relatively low risk of a positive outcome, suggesting that these persons correctly self-identified their symptoms as unlikely to be attributable to COVID-19. Interestingly, a higher propensity to test (OR of 1.45) was associated with the variable suspected non-COVID cause, the definition of which included allergy. If allergy is removed from the definition of this variable, the OR increases to 1.89, indicating that causes other than allergy drive this result. A higher willingness to be tested (or alternatively, a greater degree of cautiousness) accompanied by a low risk of positivity was observed among those with children aged <5 years in the household, age-group 30-39 years, asthma sufferers, and those with an underlying lung condition. Such information may not necessarily help guide testing policy, but nevertheless is valuable in understanding the participant and household factors associated with other aspects of the data.

The low odds of getting tested given symptom if had previously tested positive (OR of 0.35) has implications for the detection of re-infections within test-based COVID-19 surveillance data. This lower test propensity could lead to severe under-ascertainment of re-infection, if there is a reduced willingness to get tested – even given symptoms – when one had previously tested positive.

Previous research on factors related to SARS-CoV-2 test uptake was conducted using ecological analysis to compare state-level testing rates in the USA with prevalence of underlying conditions [12], which suggested associations between testing rates and certain COVID-19 risk factors. For states with higher prevalences of hypertension, diabetes, or lung cancer, overall testing rates were lower and positivity rates were higher compared with states in which the prevalence of these risk factors was lower. Ecological associations between testing rate and ethnic composition, and between positivity rate and socio-economic status were also observed at neighbourhood level in New York City [13]. A large study using individual-level data collected from a mobile/web app during April/May 2020 in the USA (when screening was not universally available) reported a higher odds of being tested associated with healthcare and other essential occupations, and a higher odds for ages 55-64 and 65+ years and a lower odds for females [10]; these age and sex findings are inconsistent with our results. Our study builds on this previous research and will be useful in contexts beyond the Netherlands.

Our approach has a number of limitations. First, position in the testing/positivity space (Fig. 3) for categorical variables (e.g., occupation) is affected by the reference category selected. Second, as with all participatory surveillance systems, observed associations with participant or other factors and survey responses cannot distinguish the effects of the measured variable from reporting behaviour. For instance, if participants reporting underlying allergies are more prone to be aware of their symptoms (and thus report them) compared with non-allergy sufferers, associations between this risk factor and test propensity may be distorted. Third, the dependence between lower test propensity and higher positivity rate for certain groups is consistent with risk behaviours associated with both unwillingness to test and with infection, or with between-group differences in

symptom severity threshold for getting tested, or with differences in the ability to accurately assess infection risk and so avoid unnecessary testing. Our data cannot distinguish between these accounts. We did not have sufficient data to investigate whether ORs of a positive result for certain groups (such as age) changed over our study period, in line with group-specific infection risk varying over time.

Although those individuals who had previously tested positive appear much less likely to get tested again – even with eligible symptoms – behaviour may depend on the severity of these symptoms; we cannot easily assess this. Also, having previously tested positive might increase the chance of reporting (continued) symptoms in following surveys (i.e., either because of experienced prolonged symptoms ('long-COVID') or because the participant's threshold for symptom reporting may be lowered due to confirmation bias), which would strengthen the inverse relation with positive test history. We note that within our data collection period there were no restrictions regarding access to testing for symptomatic persons.

Finally, certain levels of the occupation covariate are not represented among all age categories; hence adjustment for both of these variables may lead to biased ORs, or the occupation category may be too restrictive or too heterogeneous with respect to age (e.g. very few persons younger than 20 years are in the *healthcare* or *education* occupation categories; as well, *occupation: N/A* includes ages from infants through retirees). However, the sensitivity analysis restricting to the working age population largely corroborated the principal results.

The level of willingness to undergo testing for SARS-CoV-2 infection in the Netherlands is undesirably low. Cross-sectional behavioural surveys demonstrated that despite 85-89% of respondents with COVID-19-compatible symptoms being supportive of getting tested, 44-64% nevertheless did not do so (period 10 November 2020 to 6 April 2021) [14]. This behaviour, also observed in other countries [15], occurs within the context of public health communications urging voluntary, cost-free testing, and thus constitutes a challenge for the pandemic response.

Estimation of the statistical relationships between test propensity and sociodemographic, background health status, and household situation variables will help in ascertaining risk factors and thus identify subpopulations for whom testing should be facilitated or otherwise encouraged, and/or are likely under-represented in the case notification and testing surveillance systems. This analysis using participatory surveillance data has assisted in determining these potential barriers to testing. Moreover, this study has enabled identification of symptoms that are strong predictors of a positive test outcome amongst the identified subpopulations with higher probability of a positive result. Improving communication to the public (and thus their motivation) to present for SARS-CoV-2 testing, especially to those individuals less willing or able to do so, poses challenges for public health planning, but will be important for effective and sustainable control and management of what is expected to become an endemic infection [16].

## REFERENCES

1. Kucharski, A. J., Klepac, P., Conlan, A. J., Kissler, S. M., Tang, M. L., Fry, H., ... & Simons, D. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. *Lancet Inf Dis.* 2020;20(10):1151-1160.
2. Aleta, A., Martin-Corral, D., Pastore y Piontti, A., et al. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nature Human Behav.* 2020;4(9):964-971.
3. Rubin GJ, Smith LE, Melendez-Torres GJ, Yardley L. Improving adherence to 'test, trace and isolate'. *J Royal Soc Med.* 2020;113(9):335-338.
4. Guerrisi C, Turbelin C, Souty C, et al. The potential value of crowdsourced surveillance systems in supplementing sentinel influenza networks: the case of France. *Eurosurv.* 2018;23:1700337.
5. Van Noort SP, Codeco CT, Koppeschaar CE, et al. Ten-year performance of Influenzanet: ILI time series, risks, vaccine effects, and care-seeking behaviour. *Epidemics* 2015;13:28-36.
6. McDonald SA, van den Wijngaard CC, et al. Risk factors associated with the incidence of self-reported COVID-19-like illness: Data from a web-based syndromic surveillance system in the Netherlands. *Epidemiol. Inf.* 2021.
7. Muller, C. J., & MacLehose, R. F. Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. *Int J Epidemiol.* 2014;43(3):962-970.
8. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. 2019. Vienna, Austria. URL <https://www.R-project.org/>.
9. Menni, C., Valdes, A. M., Freidin, M. B., Sudre, C. H., Nguyen, L. H., Drew, D. A., ... & Spector, T. D. Real-time tracking of self-reported symptoms to predict potential COVID-19. *Nature Med.* 2020;26(7):1037-1040.
10. Allen, W. E., Altae-Tran, H., Briggs, J., et al. Population-scale longitudinal mapping of COVID-19 symptoms, behaviour and testing. *Nature Hum Behav.* 2020;4(9):972-982.

11. Zens M, Brammertz A, Herpich J, Südkamp N, Hinterseer, M. App-based tracking of self-reported COVID-19 symptoms: Analysis of questionnaire data. *J Med Internet Res.* 2020;22(9):e21956
12. Souch, J. M., & Cossman, J. S. A commentary on rural-urban disparities in COVID-19 testing rates per 100,000 and risk factors. *J Rural Health* 2021; 37(1):188-190. DOI: 10.1111/jrh.12450.
13. Lieberman-Cribbin W, Tuminello S, Flores RM, Taioli E. Disparities in COVID-19 Testing and Positivity in New York City. *Am. J. Prev. Med.* 2020;59:326–332.
14. RIVM Behaviour Unit. Compliance and Support. Data until 9 March 2021. <https://coronadashboard.government.nl/landelijk/gedrag>. (Accessed 16 May 2021).
15. Smith, L. E., Potts, H. W., Amlôt, R., Fear, N. T., Michie, S., & Rubin, G. J. Adherence to the test, trace, and isolate system in the UK: results from 37 nationally representative surveys. *BMJ* 2021;372.
16. Lavine JS, Bjornstad ON, Antia R. Immunological characteristics govern the transition of COVID-19 to endemicity. *Science.* 2021 Feb 12;371(6530):741-5.

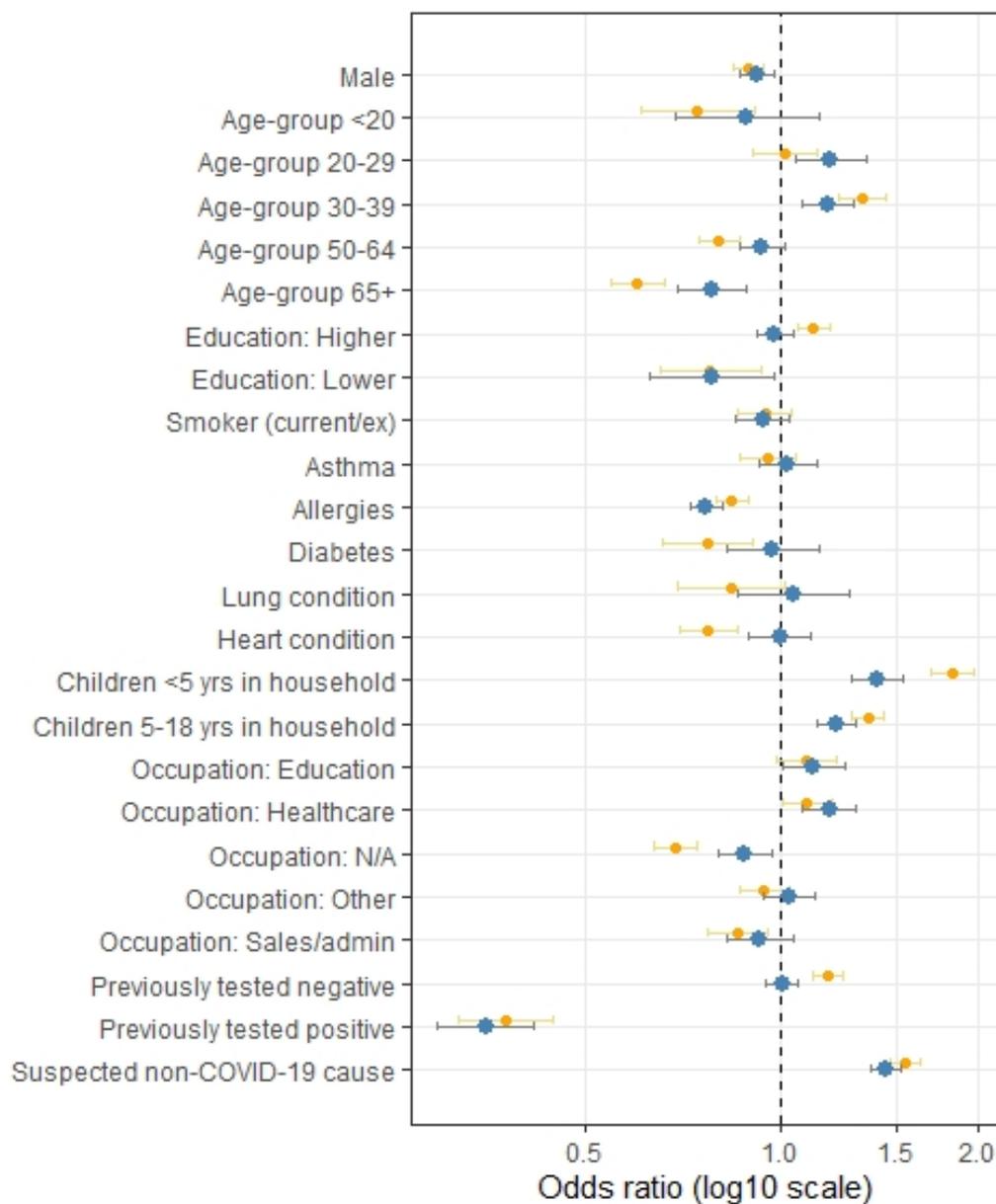


**Table 1.** Distribution over participant and other factors, and results of univariate and multivariable logistic regressions using generalised estimating equations for the outcome test propensity; study period 17 Nov 2020 through 18 April 2021 ( $n=50,946$  surveys with at least one reported symptom). The model-predicted proportion of surveys in which testing was reported is also shown.

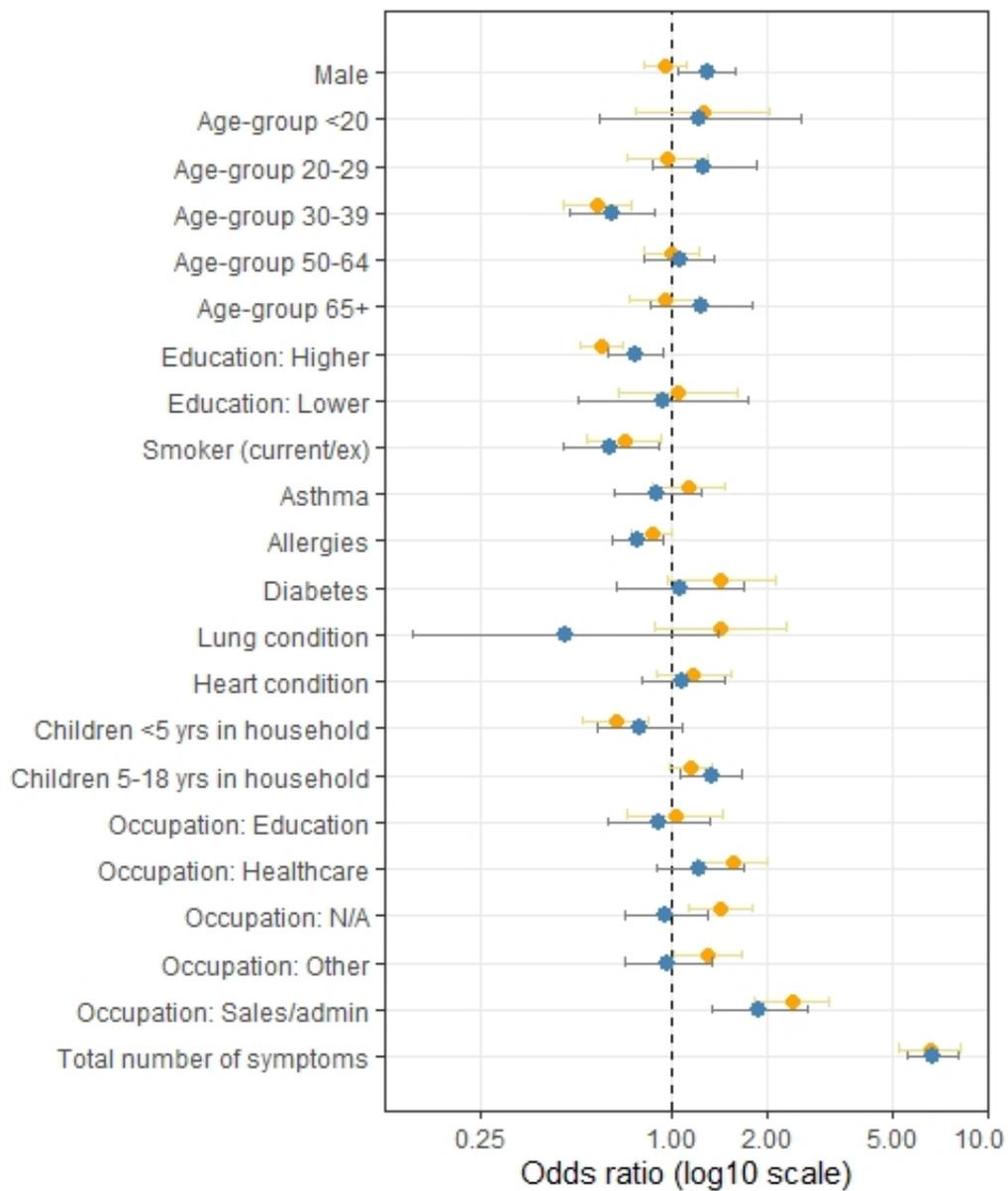
Variable	n reporting tested	N weekly surveys	Predicted proportion tested	Unadjusted Odds Ratio (95% CI)	Adjusted Odds Ratio (95% CI)
(All)	8997	50946	0.195	–	–
<i>Sex</i>					
Male	3118	18993	0.185	0.89 (0.84-0.94)	0.92 (0.86-0.98)
Female	5879	31953	0.199	Ref.	Ref.
<i>Age-group</i>					
< 20 years	161	1052	0.179	0.75 (0.61-0.91)	0.89 (0.69-1.14)
20-29 years	661	3338	0.225	1.01 (0.90-1.13)	1.19 (1.05-1.35)
30-39 years	1962	7974	0.224	1.33 (1.22-1.45)	1.18 (1.08-1.29)
40-49 years	2007	10312	0.197	Ref.	Ref.
50-64 years	3115	19393	0.187	0.80 (0.75-0.86)	0.94 (0.86-1.02)
65+ years	1091	8877	0.162	0.60 (0.55-0.66)	0.78 (0.69-0.89)
<i>Education level</i>					
None/lower or missing	203	1556	0.162	0.78 (0.65-0.93)	0.78 (0.63-0.98)
Middle	2840	17279	0.197	Ref.	Ref.
Higher	5954	32111	0.194	1.12 (1.06-1.19)	0.98 (0.92-1.04)
Never smoker	8192	46129	0.195	Ref.	Ref.
Current/ex smoker	805	4817	0.185	0.94 (0.86-1.04)	0.94 (0.85-1.03)
No asthma	8212	46281	0.194	Ref.	Ref.
Asthma	785	4665	0.197	0.95 (0.86-1.05)	1.02 (0.92-1.14)
No allergy(s)/hay fever	5198	27474	0.213	Ref.	Ref.
Allergy(s)/hay fever	3799	23472	0.173	0.84 (0.80-0.89)	0.77 (0.73-0.81)
No diabetes	8379	49116	0.194	Ref.	Ref.
Diabetes	258	1830	0.190	0.77 (0.66-0.90)	0.97 (0.83-1.14)
No chronic lung disease	8775	49489	0.194	Ref.	Ref.
Chronic lung disease	222	1457	0.201	0.84 (0.70-1.01)	1.05 (0.86-1.27)

No cardiovascular disease	8324	46264	0.194	Ref.	Ref.
Cardiovascular disease	673	4682	0.194	0.77 (0.70-0.86)	1.00 (0.89-1.11)
<i>1+ children &lt; 5 years in household</i>					
No	7498	45457	0.187	Ref.	Ref.
Yes	1499	5489	0.243	1.83 (1.70-1.97)	1.41 (1.29-1.54)
<i>1+ children 5-18 years in household</i>					
No	6195	37835	0.186	Ref.	Ref.
Yes	2802	13111	0.216	1.36 (1.28-1.44)	1.22 (1.14-1.31)
<i>Occupation category</i>					
Education	821	3916	0.213	1.09 (0.98-1.22)	1.12 (1.00-1.25)
Healthcare	1502	7077	0.222	1.10 (1.00-1.20)	1.19 (1.08-1.30)
Knowledge	1965	9964	0.195	Ref.	Ref.
Sales/admin	774	4587	0.183	0.86 (0.77-0.96)	0.93 (0.82-1.04)
Other	1528	8279	0.200	0.94 (0.86-1.03)	1.03 (0.94-1.13)
Not applicable	2407	17123	0.176	0.69 (0.64-0.74)	0.88 (0.80-0.97)
<i>Previous swab test</i>					
No	3917	25699	0.202	Ref.	Ref.
Yes, negative	4708	21399	0.203	1.18 (1.12-1.25)	1.00 (0.95-1.06)
Yes, positive	372	3848	0.084	0.38 (0.32-0.45)	0.35 (0.30-0.42)
<i>Suspected non-COVID-19 cause</i>					
No	2855	20128	0.160	Ref.	Ref.
Yes	6142	30818	0.215	1.55 (1.47-1.63)	1.45 (1.37-1.53)

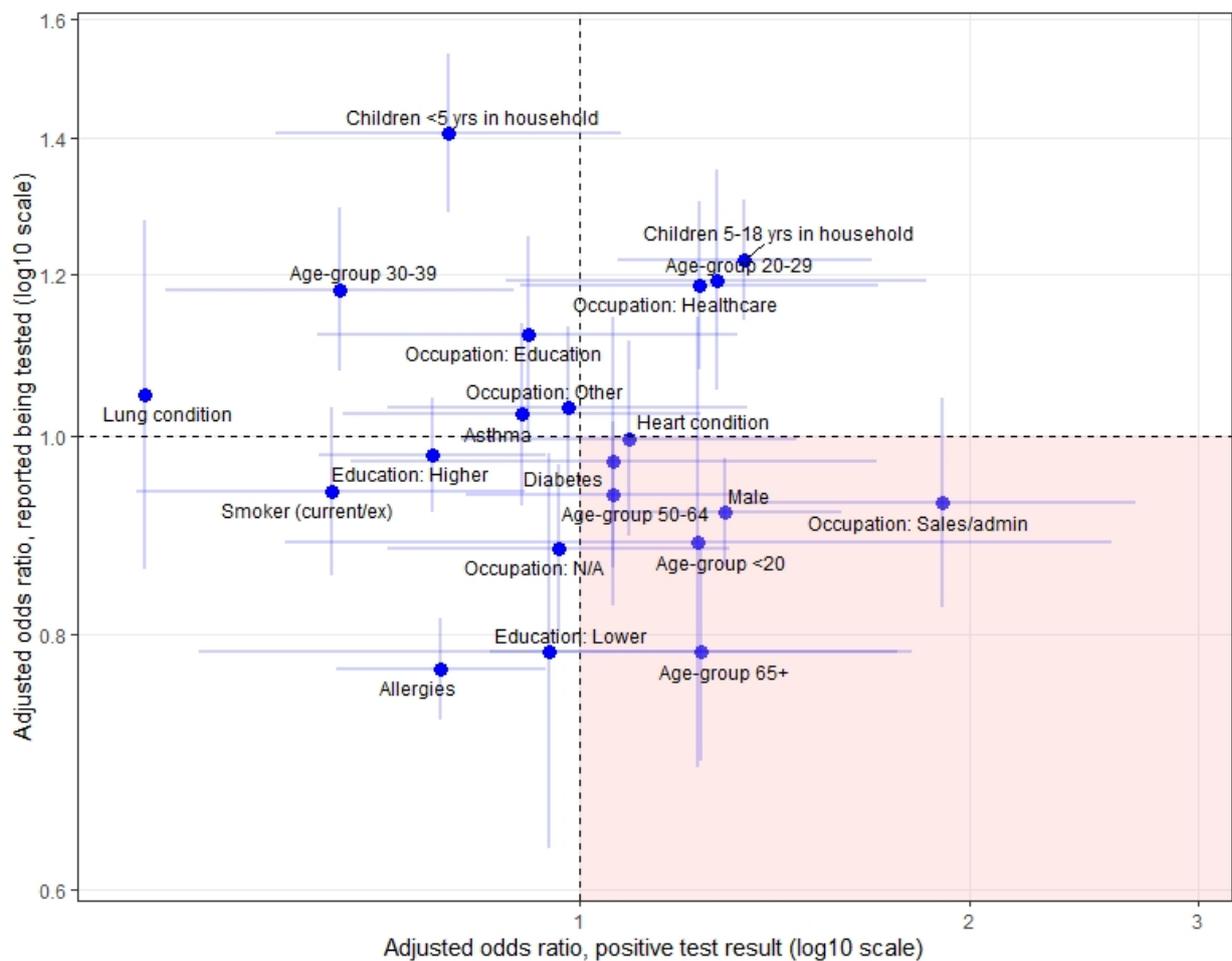
**Fig. 1.** Coefficients from generalized estimating equations regression analysis of test propensity, among all participants reporting at least one symptom (analysis period 17 November 2020 to 18 April 2021);  $n=50,946$  surveys). Orange and blue symbols represent unadjusted and adjusted coefficients, respectively. Reference categories for age-group was 40-49 years; for education level, middle; for occupation, knowledge worker; for previous swab test, no.



**Fig. 2.** Coefficients from generalized estimating equations regression analysis of positive test result among all participants who reported undergoing a swab test since their previous survey (analysis period 17 November 2020 to 18 April 2021;  $n=12,315$ ). Orange and blue symbols represent unadjusted and adjusted coefficients, respectively.



**Fig. 3.** Crossplot comparing adjusted odds ratios (ORs) for test propensity and for positivity, for the same participant factors (analysis period 17 November 2020 to 18 April 2021). Light blue crosses indicate the 95% CIs in each dimension. The shaded quadrant indicates the combination of interest: lower test propensity and higher positivity. Occupation category *N/A* (not applicable) consists of children, students, household, unemployed, and retired persons.



## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryfileInfectieradartestpropensitypositivity.docx](#)