medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

COVID-19 symptom frequency comparison: non-hospitalised positively and negatively tested persons with flu-like symptoms in Austria

Nicolas Musch[1], Stefanie Gruarin[2], Alistair Martin[1], Jama Nateqi[2,3], Thomas Lutz[4], Michael Binder[5], Judith H. Aberle[6], Bernhard Knapp[1*]

[1] Symptoma, Data Science Department, Vienna, Austria
[2] Symptoma, Medical Department, Attersee, Austria
[3] Department of Internal Medicine, Paracelsus Medical University, Salzburg, Austria
[4] Symptoma, Software Development Department, Salzburg, Austria
[5] Vienna Health Care Company, Vienna, Austria
[6] Center for Virology, Medical University of Vienna, Vienna, Austria

* Corresponding author: science@symptoma.com

Abstract

Background

The majority of clinical studies reporting on COVID-19 symptom frequencies focus on patients already hospitalized. Thus, reported symptom frequencies may not be applicable to the general population. Here we report COVID-19 symptom frequencies for the general population in a major European city.

Methods

During a scientific collaboration between the Vienna Social Fund (FSW, Vienna, Austria), the Public Health Services of the City of Vienna (MA15) and the AI-biotech company Symptoma we recorded symptom frequencies gathered by the COVID-19 chatbot of the city government of Vienna and corresponding SARS-CoV-2 nucleic acid amplification test (NAAT) results. Chatbot

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

users answered 13 yes/no questions about symptoms and provided information about age and sex. Subsequently a medically trained professional came to their address to take a sample and NAAT results were obtained.

Findings

Between November 2 and January 5, a total of 3011 persons experiencing flu-like symptoms had completed the chatbot-session and were subsequently tested by a NAAT. NAATs were performed by at home visitations of medical professionals. NAAT analysis was positive in 816 persons (27.1%). We compared the symptom frequencies between COVID-19 positive and negative users, and between male and female users. The symptoms (sorted by frequency) of users with positive NAATs were malaise (81.1%), fatigue (72.9%), headache (64.1%), cough (57.7%), fever (50.7%), sore throat (40.7%), rhinorrhea (31.0%), sneezing (28.4%), dysgeusia (27.1%), hyposmia (26.5%), dyspnea (11.4%) and diarrhea (10.9%) while 34.9% reported a close contact with a COVID-19 case. Among these the frequencies of cough, fever, hyposmia, dysgeusia, malaise, headache, close contact with COVID-19 case and fatigue were significantly ($P < 0.01$) increased in COVID-19 positive persons while the frequencies of dyspnea, diarrhea and sore throat were significantly ($P < 0.01$) decreased in COVID-19 positive persons. There was no significant difference for rhinorrhea and sneezing.

Keywords: COVID-19; symptom checker; symptom frequencies, chat bot

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

COVID-19 symptom frequencies deliver valuable information for health authorities during this pandemic (e.g. screening the population and rapidly identifying cases). Symptom frequencies are used in applications including triage recommendation [1] and COVID-19 diagnostics [2] [3]. However, the reported symptom frequencies vary considerably across studies [1]. An example is “fever” for which symptom frequencies were reported between 7% and 91% [1]. Causes for these discrepancies include the over-representation of patients in hospital settings [2] and the lack of clarity on how symptoms were collected [1].

This study describes the COVID-19 symptom frequencies observed in a non-hospitalized metropolitan population experiencing flu-like symptoms and compares them with COVID-19 negatively tested persons also experiencing flu-like symptoms. The data were collected systematically via a standardized web interface and the COVID-19 status was determined by standardized NAAT.

Methods

Data collection

From November 2020, the City of Vienna's online COVID-19 symptom checker provided inhabitants of Austria’s capital with an initial COVID-19 risk assessment, as well as possible options for further action, e.g. a NAAT [4] [5] test, taken by visitation at home.

The aim was to offer an additional service, complementing the medical telephone health service “1450”. The symptom checker lent itself as an easily executable digital channel, available via https://coronavirus.wien.gv.at/site/symptomchecker/ .

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The Vienna Social Fund (FSW), the Public Health Services of the City of Vienna (MA15) and the company Symptoma mutually developed the chatbot based on the database, artificial intelligence, and algorithms of Symptoma [3], [6], [7]. During the chatbot conversation, the user is answering a series of 13 yes/no questions about symptoms. These were fever (>38°C), cough, dyspnea, sneezing, rhinorrhea, sore throat, malaise, fatigue, diarrhea, headache, hyposmia, dysgeusia and close contact with COVID-19 case. Each user was asked the same set of questions and responses were recorded accordingly. If the chatbot AI assessed a user to have a medium/high risk for a COVID-19 infection, the user reported a positive self-test, the user returned from abroad or the user had a severe medical precondition and any type of symptom then the user was offered a NAAT [4] [5].

The statistics reported in this paper are based on the combined information of the chatbot conversations and the results of the NAATs. A total of 3011 users were screened this way between the 2nd of November 2020 and 5th of January 2021 .

Data analysis

All data were anonymised prior to this analysis. Only sex, age group, the answers to the questions and the result of the NAAT were used for this study. Users who did not provide a sex information (n=20) were exlcuded from the symptom frequencies comparison between female and male. All analyses were done in Python 3.8 using the libraries numpy (1.19.4), pandas (1.1.5), scikit-learn (0.24.0) and statsmodels (0.12.1). P-values were calculated by a two-tailed Fisher's exact test and corrected for multiple testing by the Benjamini-Hochberg method [8].

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Results

Symptom frequencies among COVID-19 positiv and COVID-19 negativ users

Summary statistics of participants and numerical details can be found in Table 1. Our study cohort consists of 3011 non-hospitalised persons experiencing flu-like symptoms of which 816 (27.1%) tested positive for COVID-19 (C19+) and 2142 (71.2%) tested negative for COVID-19 (C19-). The test was unclear for 53 persons (1.7%). While we did not record the exact age of users for data protection issues (see Table 1) the median group age is 35 years for C19+ as well as for C19-.

In Figure 1, we compared the symptom frequencies between C19+ and C19-. The symptoms most frequently reported by C19+ users were malaise (81.1%), fatigue (72.9%), headache (64.1%), cough (57.7%) and fever (50.7%). Users less frequently reported sore throat (40.7%), close contact with COVID-19 case (34.9%), rhinorrhea (31.0%), sneezing (28.4%), dysgeusia (27.1%) and hyposmia (26.5%). Dyspnea (11.4%) and diarrhea (10.9%) were rarely reported (Figure 1A).

C19+ users significantly ($P < 0.001$) more often experienced cough, fever, hyposmia, dysgeusia, malaise, headache and close contact with COVID-19 case. On the contrary, C19- users significantly ($P < 0.01$) more often experienced sore throat, diarrhea and dyspnea.

No significant difference between the C19+ and C19- group was found for rhinorrhea (P = 0.19) and sneezing (P = 0.11) (Figure 1A).

The largest relative increase in C19+ persons was found for hyposmia (+63% or +16 percent points), dysgeusia (+46% or +12 percent points), cough (+39% or +22 percent points) and fever (+39% or +20 percent points). The largest relative decrease in C19+ persons was found in dyspnea (-56% or -6 percent points) and diarrhea (-60% or -7 percent points) (Figure 1B).

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 1. Characteristics of users.

Users characteristics	**- Non-hospitalised - Experiencing flu-like symptoms - November 2020 until January 2021**				
		Total	**Positive NAAT**	**Negative NAAT**	**Unclear NAAT**
N users		3011	816	2142	53
Sex	**Female**	1646	409	1206	31
	Male	1345	399	925	21
	Unknown	20	8	11	1
Age group	**<16**	129	27	100	2
	16-20	173	42	127	4
	21-30	1044	259	768	17
	31-40	813	207	594	12
	41-50	372	113	253	6
	51-60	298	106	183	9
	61-70	114	35	76	3
	>71	63	25	38	0
	Unknown	5	2	3	0
Symptom	**Fatigue**	2077	595	1442	40
	Hyposmia	439	216	209	14
	Sore Throat	1373	332	1018	23
	Cough	1250	471	755	24
	Dyspnea	481	93	382	6
	Rhinorrhea	990	253	719	18
	Malaise	2181	662	1472	47
	Dysgeusia	546	221	311	14

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

	Sneezing	925	232	677	16
	Fever	1113	414	667	32
	Diarrhea	476	89	376	11
	Headache	1665	523	1110	32
	Close Contact COVID-19 Case	851	285	547	19

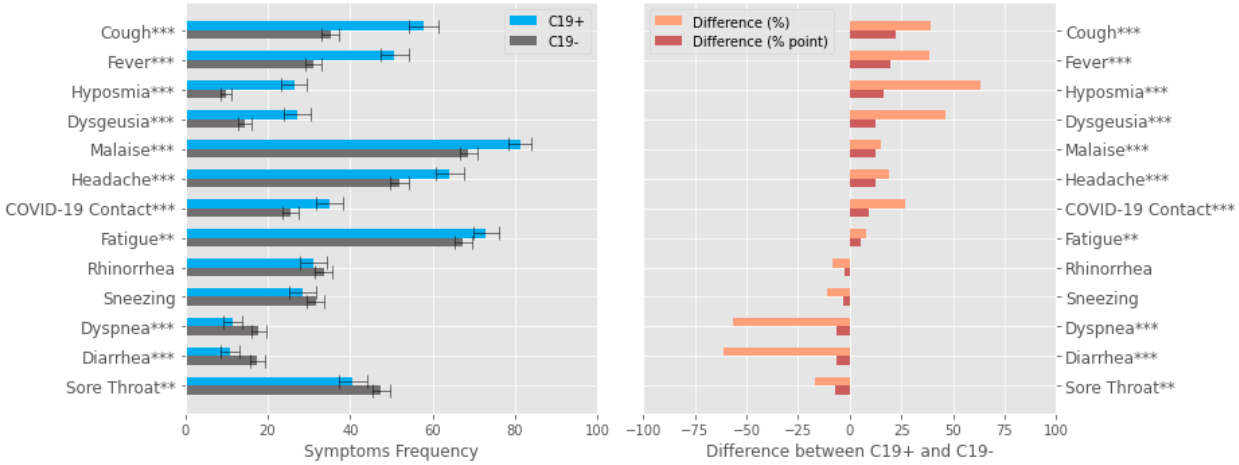


Figure 1. Symptom frequencies for the C19+ and C19- groups (left) and the difference between those groups in percentage and percentage point (right). The difference in percent is always expressed as the difference between C19+ and C19- divided by C19+. An alternative version using max(C19+,C19-) as denominator is shown in Supplementary Figure 1. Errors bars show the 95% confidence intervals. Symptoms with a p-value less than 0.05, 0.01 and 0.001 are indicated with one, two and three asterisks respectively.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Co-occurrence and correlation of symptoms

Co-occurrence and correlation matrices between the symptoms for C19+, C19- and for all users together are shown in Supplementary Figure 2. The three highest co-occurrences can be found for malaise and fatigue in C19+ (present in 67%) and all users (60%) as well as for malaise and headache in C19+ (59%).

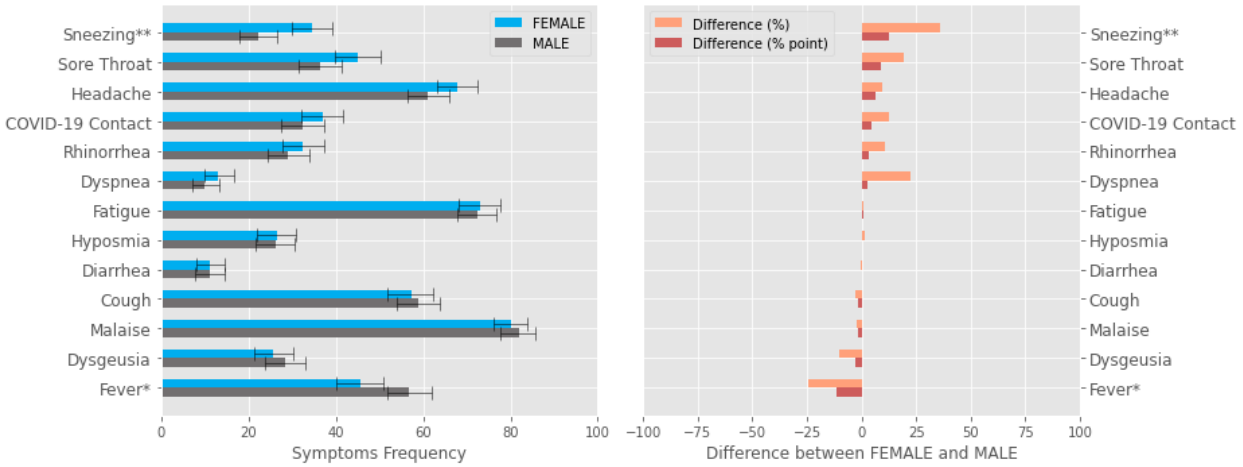
The three highest correlations can be found for dysgeusia and hyposmia in C19+ (r_s=0.70), all (r_s=0.64) and C19- (r_s=0.58). Other correlations for all users include expected pairs like malaise and fatigue (r_s=0.48), sneezing and rhinorrhea (r_s=0.45), malaise and headache (r_s=0.37), headache and fatigue (r_s=0.33) as well as fever and malaise (r_s=0.31). There were no notable correlations for age and sex. For details see Supplementary Figure 2.

Symptom frequency comparison between female and male users

In the C19+ group, fever is more frequently present for men than for women (P = 0.010) while sneezing is more frequently present for women than for men (P = 0.0012).
In the C19- group, the fever frequency is also higher for men than for women (P < 0.001) but for sneezing the difference is insignificant (P = 0.27). In the C19- group, women more frequently experience headache and sore throat (P < 0.001) (Figure 2).

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

A) C19+ group



B) C19- group

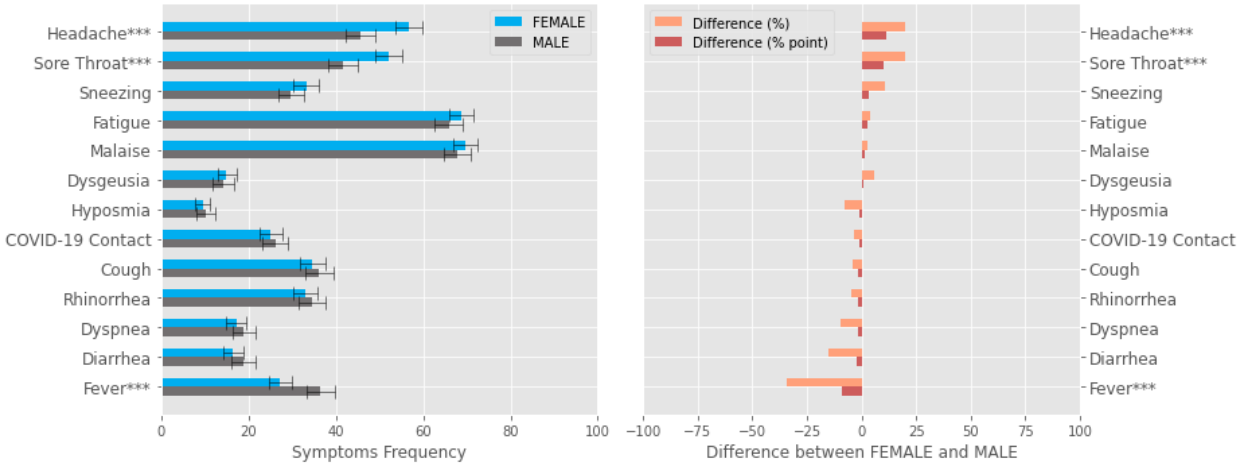


Fig 2. Symptom frequencies among the female and male (left) and the difference between those frequencies in percentage and percentage points (right). (A) For the C19+ group and (B) for the C19- group. Errors bars show the 95% confidence intervals. Symptoms with a p-value less than 0.05, 0.01 and 0.001 are indicated with one, two and three asterisks respectively.

Discussion

To the best of our knowledge, this is the first study reporting COVID-19 symptom frequencies gathered by an automated system directly associated with experimental NAATs.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Our results showed a cough frequency of 58% which falls within the previously reported symptom frequency range of the literature [9] (range [43%-71%]). Similar agreements were found for sore throat (41% vs [5%-71%]), fever (51% vs [7%-91%]), malaise (81% vs [19%-86%]) and headache (64% vs [3%-71%]). The relatively high upper limit of these literature symptom frequency ranges might be due to the majority of previous studies being based on hospitalized patients [1].

Only the frequency of fatigue (73%) does not fall within the range of [9] (10%-57%). This discrepancy might be explained by the non-specific nature of fatigue.

The results also show that all the symptom frequencies reported here (with the exception of rhinorrhea and sneezing) are different between C19+ from C19- users. This agrees with other studies [10] [11] that also reported hyposmia, dysgeusia and fever as significantly increased in C19+ persons. Considering the high relative difference of hyposmia and dysgeusia frequencies for C19+ users, our study suggests that hyposmia and dysgeusia are specific but not sensitive, i.e. their presence likely leads to C19+ but no conclusion can be drawn from their absence.

Finally, results show symptoms were reported in different frequencies by men and women, which could be caused by sex-specific differences in the clinical course [12], [13] or women being more likely to take a test at lighter symptoms (COVID-19 positiv rates were 24.8% for women and 29.7% for men while paticipants were 55.0% female and 45.0% male).

One strength of our study is the systematic symptom collection i.e. each user was asked the same questions by the chatbot which allows for higher data consistency than handwritten reports. Additionally, the result of the NAAT is automatically associated with the reported symptoms, instead of being communicated by the users themselves which would risk the inclusion of falsely reported NAAT results. Another strength is that most symptom frequencies reported in the literature are based on hospitalized patients [1] while this study reports symptom frequencies for the general population in central Europe.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

However, our study also has limitations: there is a selection bias for participants as old people are unlikely to use a chatbot (see Table 1 and [14]), there is an overrepresentation of female participants in this study (45.0% male and 55.0% female) and persons with a low risk of being C19+ were not offered a test (see methods). Also, the dyspnea frequency and difference reported here might be due to a sample bias as (a) dyspnea is often a late symptom of an infection while chatbot users might rather be at an earlier stage of an infection (b) dyspnea can be a distressing symptom and affected individuals might rather call an emergency hotline instead of using a chatbot [15] and/or (c) dyspnea might be underreported in patients with COVID-19 pneumonia which might not be the case for other (non-COVID-19-related) pulmonary diseases, such as chronic obstructive pulmonary disease, or acute hypercapnic respiratory failure [16].

Also the finding that C19- persons more often experienced sore throat and diarrhea than C19+ persons might be due to the potentially rather early time point at which the chatbot is used during an infection progression i.e. it is likely that persons use the chatbot once they start experiencing flu-like symptoms but seldomly after being a week or two into the infection. Nevertheless we believe that these differences are interesting insights for early onset diagnosis of COVID-19.

Data Availability

All relevant data is reported within the study.

Funding

This study has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 830017 and by the Austrian Research Promotion Agency under grant agreement No 880939 (supported by the Federal Ministries

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Republic of Austria for Digital and Economic Affairs and Climate Action, Environment, Energy, Mobility, Innovation and Technology).

Acknowledgement

We would like to thank the Vienna Social Fund (FSW), Public Health Services of the City of Vienna (MA15) and the Vienna Health Authority for providing data and making this publication possible.

Declaration of interests

NM, SG, AM, JN, TL and BK are employees of Symptoma GmbH. JN and TL hold shares of Symptoma.

Ethical considerations

This study was exempted from ethics review by the ethics commission of the city of Vienna (MA15-EK/21-037-VK_NZ). All individuals using the chatbot agreed that their data will be used in an anonymised way.

Author contributions

Study design: BK, JN, TL. Data compilation: NM, TL, MB. Data analysis and critical revision: NM, BK, AM, JA. Writing the manuscript: NM, BK. Revising the manuscript critically: BK, SG, AM, JN, JA, MB.

[1] T. Struyf *et al.*, “Signs and symptoms to determine if a patient presenting in primary care or hospital outpatient settings has COVID-19 disease,” *Cochrane Database Syst. Rev.*, no. 7, 2020, doi: 10.1002/14651858.CD013665.

[2] C. Menni, “Real-time tracking of self-reported symptoms to predict potential COVID-19,” *Nat. Med.*, vol. 26, p. 8, 2020, doi: 10.1038/s41591-020-0916-2.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

[3] A. Martin *et al.*, "An artificial intelligence-based first-line defence against COVID-19: digitally screening citizens for risks via a chatbot," *Sci. Rep.*, vol. 10, no. 1, Art. no. 1, Nov. 2020, doi: 10.1038/s41598-020-75912-x.

[4] V. Corman *et al.*, "Diagnostic detection of 2019-nCoV by real-time RT-PCR," *https://www.who.int/docs/default-source/coronaviruse/protocol-v2-1.pdf*, 2020.

[5] "cobas® SARS-CoV-2 Test," *Diagnostics*. https://diagnostics.roche.com/global/en/products/params/cobas-sars-cov-2-test.html (accessed Mar. 19, 2021).

[6] N. Munsch *et al.*, "Diagnostic Accuracy of Web-Based COVID-19 Symptom Checkers: Comparison Study," *J. Med. Internet Res.*, vol. 22, no. 10, p. e21299, 2020, doi: 10.2196/21299.

[7] J. Nateqi *et al.*, "[From symptom to diagnosis-symptom checkers re-evaluated : Are symptom checkers finally sufficient and accurate to use? An update from the ENT perspective]," *HNO*, vol. 67, no. 5, pp. 334–342, May 2019, doi: 10.1007/s00106-019-0666-y.

[8] Y. Benjamini and Y. Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *J. R. Stat. Soc. Ser. B Methodol.*, vol. 57, no. 1, pp. 289–300, 1995, doi: https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

[9] P. H. A. of Canada, "COVID-19 signs, symptoms and severity of disease: A clinician guide," *aem*, Sep. 17, 2020. https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/guidance-documents/signs-symptoms-severity.html#toc0 (accessed Jan. 20, 2021).

[10] G. Quer *et al.*, "Wearable sensor data and self-reported symptoms for COVID-19 detection," *Nat. Med.*, vol. 27, no. 1, pp. 73–77, Jan. 2021, doi: 10.1038/s41591-020-1123-x.

[11] N. A. Dreyer *et al.*, "Self-reported symptoms from exposure to Covid-19 provide support to clinical diagnosis, triage and prognosis: An exploratory analysis," *Travel Med. Infect. Dis.*, vol. 38, p. 101909, Nov. 2020, doi: 10.1016/j.tmaid.2020.101909.

[12] H. Peckham *et al.*, "Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ITU admission," *Nat. Commun.*, vol. 11, no. 1, Art. no. 1, Dec. 2020, doi: 10.1038/s41467-020-19741-6.

[13] A. Pradhan and P.-E. Olsson, "Sex differences in severity and mortality from COVID-19: are males more vulnerable?," *Biol. Sex Differ.*, vol. 11, Sep. 2020, doi: 10.1186/s13293-020-00330-7.

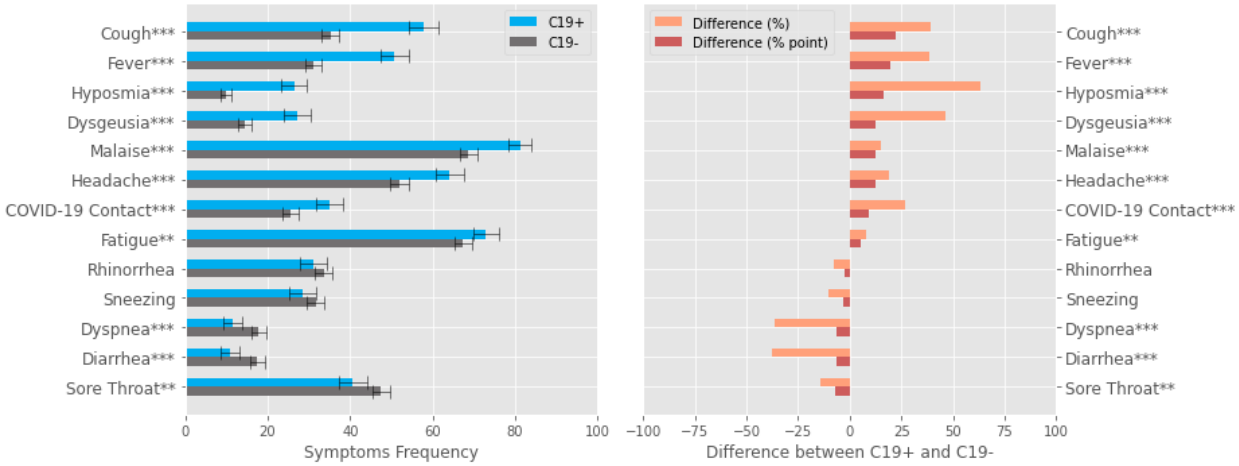
[14] G. Fagherazzi, C. Goetzinger, M. A. Rashid, G. A. Aguayo, and L. Huiart, "Digital Health Strategies to Fight COVID-19 Worldwide: Challenges, Recommendations, and a Call for Papers," *J. Med. Internet Res.*, vol. 22, no. 6, p. e19284, 2020, doi: 10.2196/19284.

[15] M. B. Parshall *et al.*, "An official American Thoracic Society statement: update on the mechanisms, assessment, and management of dyspnea," *Am. J. Respir. Crit. Care Med.*, vol. 185, no. 4, pp. 435–452, Feb. 2012, doi: 10.1164/rccm.201111-2042ST.

[16] G. Allali, C. Marti, O. Grosgurin, C. Morélot-Panzini, T. Similowski, and D. Adler, "Dyspnea: The vanished warning symptom of COVID-19 pneumonia," *J. Med. Virol.*, vol. 92, no. 11, pp. 2272–2273, 2020, doi: https://doi.org/10.1002/jmv.26172.

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Supplementary material



Supplementary Figure 1: Same as Figure 1 but the difference in percent is expressed as the difference between C19+ and C19- divided by max(C19+,C19-).

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC-ND 4.0 International license .

Normalized co-occurrence matrix (C19+)

	Fatigue	Hyposmia	Sore Throat	Cough	Dyspnea	Rhinorrhea	Malaise	Dysgeusia	Sneezing	Fever	Diarrhea	Headache	COVID-19 Contact
Fatigue	0.73	0.21	0.33	0.45	0.1	0.24	0.67	0.23	0.24	0.4	0.09	0.54	0.23
Hyposmia	0.21	0.26	0.1	0.14	0.03	0.12	0.22	0.21	0.12	0.11	0.05	0.17	0.09
Sore Throat	0.33	0.1	0.41	0.27	0.07	0.14	0.37	0.12	0.15	0.21	0.05	0.31	0.16
Cough	0.45	0.14	0.27	0.58	0.09	0.2	0.51	0.16	0.19	0.32	0.08	0.4	0.19
Dyspnea	0.1	0.03	0.07	0.09	0.11	0.04	0.11	0.04	0.04	0.07	0.02	0.08	0.04
Rhinorrhea	0.24	0.12	0.14	0.2	0.04	0.31	0.26	0.11	0.18	0.12	0.05	0.19	0.12
Malaise	0.67	0.22	0.37	0.51	0.11	0.26	0.81	0.24	0.25	0.46	0.1	0.59	0.25
Dysgeusia	0.23	0.21	0.12	0.16	0.04	0.11	0.24	0.27	0.11	0.11	0.06	0.19	0.09
Sneezing	0.24	0.12	0.15	0.19	0.04	0.18	0.25	0.11	0.28	0.11	0.04	0.2	0.12
Fever	0.4	0.11	0.21	0.32	0.07	0.12	0.46	0.11	0.11	0.51	0.06	0.37	0.12
Diarrhea	0.09	0.05	0.05	0.08	0.02	0.05	0.1	0.06	0.04	0.06	0.11	0.07	0.04
Headache	0.54	0.17	0.31	0.4	0.08	0.19	0.59	0.19	0.2	0.37	0.07	0.64	0.19
COVID-19 Contact	0.23	0.09	0.16	0.19	0.04	0.12	0.25	0.09	0.12	0.12	0.04	0.19	0.35

Correlation matrix (C19+)

	Fatigue	Hyposmia	Sore Throat	Cough	Dyspnea	Rhinorrhea	Malaise	Dysgeusia	Sneezing	Fever	Diarrhea	Headache	COVID-19 Contact	age	sex
Fatigue	0	0.11	0.15	0.15	0.13	0.09	0.47	0.17	0.15	0.14	0.09	0.32	-0.14	0.04	0.01
Hyposmia	0.11	0	-0.02	-0.06	0.01	0.18	0.06	0.7	0.2	-0.13	0.17	0.01	-0.01	-0.05	0
Sore Throat	0.15	-0.02	0	0.15	0.13	0.05	0.2	0.02	0.15	0.03	0.01	0.2	0.06	-0.04	0.09
Cough	0.15	-0.06	0.15	0	0.17	0.08	0.2	0.01	0.11	0.09	0.12	0.13	-0.06	0.03	-0.02
Dyspnea	0.13	0.01	0.13	0.17	0	0.05	0.11	0.06	0.06	0.06	0.07	0.06	-0.01	0.05	0.05
Rhinorrhea	0.09	0.18	0.05	0.08	0.05	0	0.06	0.12	0.43	-0.15	0.09	-0.02	0.07	-0.1	0.04
Malaise	0.47	0.06	0.2	0.2	0.11	0.06	0	0.11	0.14	0.25	0.07	0.36	-0.17	0.03	-0.02
Dysgeusia	0.17	0.7	0.02	0.01	0.06	0.12	0.11	0	0.16	-0.11	0.19	0.08	-0.04	-0.05	-0.03
Sneezing	0.15	0.2	0.15	0.11	0.06	0.43	0.14	0.16	0	-0.13	0.08	0.08	0.07	-0.09	0.14
Fever	0.14	-0.13	0.03	0.09	0.06	-0.15	0.25	-0.11	-0.13	0	0.05	0.21	-0.24	0.12	-0.11
Diarrhea	0.09	0.17	0.01	0.12	0.07	0.09	0.07	0.19	0.08	0.05	0	0.03	0.02	-0	-0
Headache	0.32	0.01	0.2	0.13	0.06	-0.02	0.36	0.08	0.08	0.21	0.03	0	-0.14	-0.03	0.07
COVID-19 Contact	-0.14	-0.01	0.06	-0.06	-0.01	0.07	-0.17	-0.04	0.07	-0.24	0.02	-0.14	0	-0.13	0.05
age	0.04	-0.05	-0.04	0.03	0.05	-0.1	0.03	-0.05	-0.09	0.12	-0	-0.03	-0.13	0	0.01
sex	0.01	0	0.09	-0.02	0.05	0.04	-0.02	-0.03	0.14	-0.11	-0	0.07	0.05	0.01	0

Normalized co-occurrence matrix (C19-)

	Fatigue	Hyposmia	Sore Throat	Cough	Dyspnea	Rhinorrhea	Malaise	Dysgeusia	Sneezing	Fever	Diarrhea	Headache	COVID-19 Contact
Fatigue	0.67	0.08	0.35	0.26	0.15	0.24	0.57	0.12	0.25	0.26	0.14	0.43	0.13
Hyposmia	0.08	0.1	0.05	0.05	0.03	0.05	0.08	0.08	0.05	0.04	0.02	0.06	0.02
Sore Throat	0.35	0.05	0.48	0.22	0.1	0.18	0.37	0.09	0.18	0.15	0.08	0.28	0.08
Cough	0.26	0.05	0.22	0.35	0.1	0.17	0.26	0.07	0.17	0.11	0.06	0.2	0.06
Dyspnea	0.15	0.03	0.1	0.1	0.18	0.07	0.14	0.04	0.07	0.07	0.04	0.1	0.03
Rhinorrhea	0.24	0.05	0.18	0.17	0.07	0.34	0.24	0.07	0.21	0.09	0.06	0.18	0.07
Malaise	0.57	0.08	0.37	0.26	0.14	0.24	0.69	0.12	0.24	0.28	0.14	0.44	0.1
Dysgeusia	0.12	0.08	0.09	0.07	0.04	0.07	0.12	0.15	0.07	0.06	0.03	0.1	0.03
Sneezing	0.25	0.05	0.18	0.17	0.07	0.21	0.24	0.07	0.32	0.08	0.07	0.19	0.08
Fever	0.26	0.04	0.15	0.11	0.07	0.09	0.28	0.06	0.08	0.31	0.07	0.2	0.03
Diarrhea	0.14	0.02	0.08	0.06	0.04	0.06	0.14	0.03	0.07	0.07	0.18	0.11	0.04
Headache	0.43	0.06	0.28	0.2	0.1	0.18	0.44	0.1	0.19	0.2	0.11	0.52	0.1
COVID-19 Contact	0.13	0.02	0.08	0.06	0.03	0.07	0.1	0.03	0.08	0.03	0.04	0.1	0.26

Correlation matrix (C19-)

	Fatigue	Hyposmia	Sore Throat	Cough	Dyspnea	Rhinorrhea	Malaise	Dysgeusia	Sneezing	Fever	Diarrhea	Headache	COVID-19 Contact	age	sex
Fatigue	0	0.09	0.14	0.1	0.15	0.05	0.48	0.15	0.16	0.23	0.13	0.33	-0.21	0.02	0.03
Hyposmia	0.09	0	0.05	0.09	0.12	0.11	0.07	0.58	0.14	0.03	0.03	0.07	-0.05	0.02	-0.01
Sore Throat	0.14	0.05	0	0.2	0.07	0.09	0.18	0.1	0.14	0.01	-0.04	0.15	-0.2	-0.05	0.1
Cough	0.1	0.09	0.2	0	0.18	0.22	0.1	0.11	0.25	-0.02	-0	0.07	-0.14	0.01	-0.02
Dyspnea	0.15	0.12	0.07	0.18	0	0.04	0.1	0.13	0.08	0.06	0.03	0.05	-0.09	0.04	-0.02
Rhinorrhea	0.05	0.11	0.09	0.22	0.04	0	0.06	0.12	0.46	-0.06	-0.02	0.02	-0.1	-0.06	-0.02
Malaise	0.48	0.07	0.18	0.1	0.1	0.06	0	0.1	0.11	0.31	0.11	0.36	-0.36	0.09	0.02
Dysgeusia	0.15	0.58	0.1	0.11	0.13	0.12	0.1	0	0.16	0.07	0.04	0.12	-0.06	0	0.01
Sneezing	0.16	0.14	0.14	0.25	0.08	0.46	0.11	0.16	0	-0.09	0.07	0.1	-0.01	-0.02	0.04
Fever	0.23	0.03	0.01	-0.02	0.06	-0.06	0.31	0.07	-0.09	0	0.07	0.18	-0.23	0.04	-0.1
Diarrhea	0.13	0.03	-0.04	-0	0.03	-0.02	0.11	0.04	0.07	0.07	0	0.08	-0.06	-0.01	-0.03
Headache	0.33	0.07	0.15	0.07	0.05	0.02	0.36	0.12	0.1	0.18	0.08	0	-0.15	0	0.11
COVID-19 Contact	-0.21	-0.05	-0.2	-0.14	-0.09	-0.1	-0.36	-0.06	-0.01	-0.23	-0.06	-0.15	0	-0.15	-0.01
age	0.02	0.02	-0.05	0.01	0.04	-0.06	0.09	0	-0.02	0.04	-0.01	0	-0.15	0	-0
sex	0.03	-0.01	0.1	-0.02	-0.02	-0.02	0.02	0.01	0.04	-0.1	-0.03	0.11	-0.01	-0	0

Normalized co-occurrence matrix (all)

	Fatigue	Hyposmia	Sore Throat	Cough	Dyspnea	Rhinorrhea	Malaise	Dysgeusia	Sneezing	Fever	Diarrhea	Headache	COVID-19 Contact
Fatigue	0.69	0.12	0.35	0.31	0.13	0.24	0.6	0.15	0.24	0.3	0.13	0.46	0.16
Hyposmia	0.12	0.15	0.07	0.07	0.03	0.07	0.12	0.11	0.07	0.06	0.03	0.09	0.04
Sore Throat	0.35	0.07	0.46	0.23	0.09	0.17	0.37	0.1	0.17	0.17	0.07	0.29	0.1
Cough	0.31	0.07	0.23	0.42	0.09	0.18	0.33	0.09	0.17	0.17	0.07	0.26	0.1
Dyspnea	0.13	0.03	0.09	0.09	0.16	0.06	0.13	0.04	0.06	0.07	0.03	0.1	0.03
Rhinorrhea	0.24	0.07	0.17	0.18	0.06	0.33	0.25	0.08	0.2	0.1	0.05	0.18	0.08
Malaise	0.6	0.12	0.37	0.33	0.13	0.25	0.72	0.15	0.25	0.34	0.13	0.48	0.15
Dysgeusia	0.15	0.11	0.1	0.09	0.04	0.08	0.15	0.18	0.08	0.07	0.04	0.12	0.05
Sneezing	0.24	0.07	0.17	0.17	0.06	0.2	0.25	0.08	0.31	0.09	0.06	0.19	0.09
Fever	0.3	0.06	0.17	0.17	0.07	0.1	0.34	0.07	0.09	0.37	0.07	0.25	0.06
Diarrhea	0.13	0.03	0.07	0.07	0.03	0.05	0.13	0.04	0.06	0.07	0.16	0.1	0.04
Headache	0.46	0.09	0.29	0.26	0.1	0.18	0.48	0.12	0.19	0.25	0.1	0.55	0.13
COVID-19 Contact	0.16	0.04	0.1	0.1	0.03	0.08	0.15	0.05	0.09	0.06	0.04	0.13	0.28

Correlation matrix (all)

	Fatigue	Hyposmia	Sore Throat	Cough	Dyspnea	Rhinorrhea	Malaise	Dysgeusia	Sneezing	Fever	Diarrhea	Headache	COVID-19 Contact	age	sex
Fatigue	0	0.1	0.13	0.12	0.14	0.06	0.48	0.15	0.15	0.22	0.12	0.33	-0.19	0.03	0.02
Hyposmia	0.1	0	0.01	0.07	0.06	0.13	0.09	0.64	0.15	0.01	0.05	0.07	-0.01	0.01	-0.01
Sore Throat	0.13	0.01	0	0.17	0.09	0.08	0.17	0.07	0.14	0	-0.03	0.15	-0.13	-0.05	0.1
Cough	0.12	0.07	0.17	0	0.16	0.17	0.15	0.1	0.2	0.05	0.01	0.11	-0.09	0.03	-0.03
Dyspnea	0.14	0.06	0.09	0.16	0	0.04	0.09	0.09	0.07	0.04	0.05	0.04	-0.07	0.04	0
Rhinorrhea	0.06	0.13	0.08	0.17	0.04	0	0.05	0.11	0.45	-0.09	0.02	0	-0.05	-0.08	0
Malaise	0.48	0.09	0.17	0.15	0.09	0.05	0	0.12	0.11	0.31	0.09	0.37	-0.29	0.09	0
Dysgeusia	0.15	0.64	0.07	0.1	0.09	0.11	0.12	0	0.15	0.03	0.07	0.12	-0.04	0	-0.01
Sneezing	0.15	0.15	0.14	0.2	0.07	0.45	0.11	0.15	0	-0.11	0.07	0.08	0.02	-0.05	0.07
Fever	0.22	0.01	0	0.05	0.04	-0.09	0.31	0.03	-0.11	0	0.05	0.2	-0.21	0.08	-0.11
Diarrhea	0.12	0.05	-0.03	0.01	0.05	0.02	0.09	0.07	0.07	0.05	0	0.06	-0.05	-0.02	-0.02
Headache	0.33	0.07	0.15	0.11	0.04	0	0.37	0.12	0.08	0.2	0.06	0	-0.13	0.01	0.09
COVID-19 Contact	-0.19	-0.01	-0.13	-0.09	-0.07	-0.05	-0.29	-0.04	0.02	-0.21	-0.05	-0.13	0	-0.14	-0
age	0.03	0.01	-0.05	0.03	0.04	-0.08	0.09	0	-0.05	0.08	-0.02	0.01	-0.14	0	-0
sex	0.02	-0.01	0.1	-0.03	0	0	0	-0.01	0.07	-0.11	-0.02	0.09	-0	-0	0

medRxiv preprint doi: https://doi.org/10.1101/2021.02.24.21252426; this version posted March 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Supplementary Figure 2: Normalized co-occurrence (how often is a specific symptom present among all persons of the respective group) and spearman correlation matrices between symptoms.