

Performance of Existing and Novel Surveillance Case Definitions for COVID-19 in the Community

Hannah E. Reses, MPH^{*,†}, Mark Fajans, MPH^{*,†}, Scott H. Lee, PhD[†], Charles M. Heilig, PhD[†], Victoria T. Chu, MD, MPH^{†,‡}, Natalie J. Thornburg, PhD[†], Kim Christensen, BS[§], Sanjib Bhattacharyya, PhD[□], Alicia Fry, MD, MPH[†], Aron J. Hall, DVM, MSPH[†], U.S. COVID-19 Household Investigation Team, Jacqueline E. Tate, PhD[†], Hannah L. Kirking, MD[†], Scott A. Nabity, MD, MPH[†]

*Joint first author

Author Affiliations

[†]COVID-19 Response Team, Centers for Disease Control and Prevention, Atlanta, USA

[‡]Epidemic Intelligence Service, Centers for Disease Control and Prevention, Atlanta, USA

[§]Utah Department of Health, Salt Lake City, USA

[□]City of Milwaukee Health Department, Milwaukee, USA

Corresponding author:

Hannah E. Reses, MPH

Centers for Disease Control and Prevention

1600 Clifton Rd NE

Roybal Building 21 Mailstop H21-3

Atlanta, GA 30329

Phone: 404-639-3900

Email: hreses@cdc.gov

Abstract

Background: Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the virus that causes coronavirus disease 2019 (COVID-19), presents with a broad range of symptoms. Existing COVID-19 case definitions were developed from early reports of severely ill, primarily hospitalized, patients. Symptom-based case definitions that guide public health surveillance and individual patient management in the community must be optimized for COVID-19 pandemic control.

Methods: We collected daily symptom diaries and performed RT-PCR on respiratory specimens over a 14-day period in 185 community members exposed to a household contact with COVID-19 in the Milwaukee, Wisconsin and Salt Lake City, Utah metropolitan areas. We interpreted the discriminatory performance (sensitivity, specificity, predictive values, F₁ score, Youden's index, and prevalence estimation) of individual symptoms and common case definitions according to two principal surveillance applications (i.e., individual screening and case counting). We also constructed novel case definitions using an exhaustive search with over 73 million symptom combinations and calculated bias-corrected and accelerated bootstrap confidence intervals stratified by children versus adults.

Findings: Common COVID-19 case definitions generally showed high sensitivity (86–96%) but low positive predictive value (PPV) (36–49%; F1 score 52–63) in this community cohort. The top performing novel symptom combinations included taste or smell dysfunction. They also improved the balance of sensitivity and PPV (F1 score 78–80) and reduced the number of false positive symptom screens. Performance indicators were generally lower for children (<18 years of age).

Interpretation: Existing COVID-19 case definitions appropriately screened in community members with COVID-19. However, they led to many false positive symptom screens and poorly estimated community prevalence. Absent unlimited, timely testing capacity, more accurate case definitions may help focus public health resources. Novel symptom combinations incorporating taste or smell dysfunction as a primary component better balanced sensitivity and specificity. Case definitions tailored specifically for children versus adults should be further explored.

Funding: This research was wholly supported by the U.S. Centers for Disease Control and Prevention.

Disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the U.S. Centers for Disease Control and Prevention/the Agency for Toxic Substances and Disease Registry.

Research in Context

Evidence before this study

Coronavirus disease 2019 (COVID-19) incidence has accelerated globally over the last several months. As the full spectrum of clinical presentations has come into clearer focus, symptom-based clinical screening and case surveillance has also evolved. Preliminary understanding of the clinical manifestation of COVID-19 was driven primarily by descriptions of hospitalized patients, as early testing algorithms prioritized more severely ill persons with classic lower respiratory symptoms and fever. Since then, more data from ambulatory settings have emerged. We searched PubMed from 1 December 2019 to 21 August 2020 for studies that assessed the diagnostic performance of case surveillance definitions for COVID-19. We found no studies examining the discriminatory performance of case surveillance definitions among contacts with mild to moderate symptoms with documented exposure to persons with COVID-19. Nonetheless, we found nine highly relevant studies: seven original reports and two review articles. Five original studies evaluated individual, self-reported symptoms (two among healthcare workers in the United States, one among healthcare workers in the Netherlands, and one online survey for the general public in Somalia) and concluded that using dysfunction of taste or smell for routine COVID-19 screening likely had utility. The fifth study had a similar conclusion based on self-reported symptoms and laboratory results collected via smartphone from the general public in the United States and the United Kingdom. Another original study modeled the substantial effect that multiple revisions to the COVID-19 case definition had on the reported disease burden in the Chinese population. Lastly, an original study illustrated the shift in discriminatory performance of established influenza surveillance case definitions for influenza between adults and children. Age-specific differences in case definition performance may also apply to COVID-19. Two articles reviewed predictive algorithms to define outpatient COVID-19 illness and risk of hospitalization. The reviewed studies were limited in that they were either restricted to individual signs or symptoms, or they incorporated blood tests or imaging that required in-person access to medical care.

Added value of this study

The discriminatory performance of case surveillance definitions for COVID-19 is important for implementing effective epidemic mitigation strategies. Our study illustrates the performance of case definitions in community members with household exposure to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) based

solely on symptom profiles. Prior work overrepresented healthcare workers or otherwise studied non-representative populations, and they did not examine across the age spectrum. Our study also provides a novel framework for refining definitions. Using 15 symptoms associated with COVID-19 for all contacts regardless of disease status, we systematically evaluated the discriminatory performance of individual symptoms and previously defined case surveillance definitions across ages and according to two core surveillance applications: 1) screening non-hospitalized individuals to prioritize public health interventions, and 2) estimating the number of non-hospitalized persons with COVID-19 (i.e., community-based syndromic surveillance). We also constructed novel symptom combinations that effectively performed both functions and improved upon widely used case surveillance definitions that may help to target interventions in the absence of unlimited laboratory diagnostic capacity. Our analyses highlight the importance of ongoing re-evaluation of symptom-based surveillance definitions to suit the intended purpose and population under surveillance. Based on our results, which were derived from household members of all ages, case surveillance definition performance may improve if developed separately for adults and children.

Implications of all the available evidence

Case definitions for COVID-19 should be tailored to maximize the discriminatory performance dependent upon its intended use. Existing COVID-19 case definitions screened in most community members with COVID-19, but also yielded a high number of false positive results. When unlimited, timely diagnostic testing is not available symptom combinations with improved accuracy (i.e., more balanced sensitivity and specificity) may help focus resources, such as recommending self-isolation among community contacts.

Background

Coronavirus disease 2019 (COVID-19) is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The virus was first identified in a cluster of patients with atypical pneumonia in Wuhan, China, in December 2019.¹ Since its emergence, the virus has spread globally, causing widespread infection and death. Following evidence of person-to-person transmission and a broader clinical spectrum of infections, case definitions for COVID-19 have been revised.² In the initial weeks of the pandemic, COVID-19 was labeled a pneumonia of unknown etiology, and many who presented to medical care had classic pneumonia-like symptoms such as fever, cough, and dyspnea. An exceptional variety of symptoms has since been reported for COVID-19, ranging from none or mild indistinct symptoms to invasive neurological disease and fulminant respiratory failure.³⁻⁷ As is common in the early response phases to novel emerging pathogens, there is ongoing need to reassess and refine surveillance case definitions for COVID-19 based on new information. Changes to case definitions affect interpretation of surveillance data, as was demonstrated by substantially different prevalence estimates when China broadened the COVID-19 case definition early in its epidemic response.²

A few studies have demonstrated the predictive value of symptom profiles in healthcare workers^{4,8,9} and other populations potentially not necessarily representative of the general public.^{5,10} These studies are subject to other limitations, too. Some applied predictive models that included serum biomarkers and imaging.^{11,12} Obtaining this information may limit real-world capture of people with mild-to-moderate SARS-CoV-2 infection and may delay public health intervention. Further, few studies to date have examined symptom combinations exclusively.

Respiratory pathogens routinely behave differently in children and adults, and this appears to be true for COVID-19 as well.¹³ For example, an assessment of ambulatory case surveillance definitions for influenza demonstrated lower sensitivity among children less than 5 years of age.¹⁴ Similar analyses across age strata are lacking for COVID-19. Reliable, age-stratified syndromic surveillance definitions would likely aid public health officials to scale up community contact tracing and develop protocols to safely operate various congregate venues, such as schools and workplaces, should unlimited, timely diagnostic testing be unavailable.

Dedicated symptom-based surveillance systems have been developed to track COVID-19 cases. These include the U.S. Council of State and Territorial Epidemiologists (CSTE) original (CSTE combination 1; released 5 April

2020) and revised (CSTE combination 2; released 7 August 2020) clinical criteria for reporting SARS-CoV-2 infection, and the original CDC COVID-19-like illness (CLI) definition (Table 1). Similarly, the Centers for Disease Control and Prevention (CDC) maintains a list of symptoms for priority SARS-CoV-2 testing. These COVID-19 case definitions and the priority testing symptom list are intended to capture as many persons with COVID-19 as possible with confirmatory testing. Finally, longstanding respiratory virus surveillance networks established to monitor influenza-like illnesses (ILI) and acute respiratory infection (ARI), which is used for community-based syndromic surveillance of respiratory syncytial virus by the World Health Organization (WHO), may be plausibly adaptable platforms for monitoring COVID-19. The performance characteristics and utility of these syndromic surveillance platforms for COVID-19 have not been well defined.⁵

We aimed to describe the diagnostic performance of two existing case definitions for COVID-19, the CDC COVID-19 symptom list, and two longstanding viral respiratory disease surveillance definitions among persons with confirmed SARS-CoV-2 exposure, stratified between adults and children. We also aimed to derive novel, practical symptom combinations in the same population. We interpreted the results primarily within the framework of two core public health surveillance functions: 1) symptom-based screening of individuals to guide SARS-CoV-2 diagnostic testing, contact tracing, and community-based isolation and quarantine, and 2) estimating disease frequency in persons with documented SARS-CoV-2 exposure. For symptom screening, we considered the merits of novel combinations when unlimited, timely diagnostic testing is unavailable.

Methods

Study Design and Data Collection

CDC collaborated with state and local health departments in the Milwaukee, Wisconsin and Salt Lake City, Utah metropolitan areas in the United States to identify and enroll a convenience sample of people with laboratory-confirmed SARS-CoV-2 infection and their household contacts from 22 March to 22 April 2020. Ours are secondary analyses of this household transmission investigation whose methods were previously published in detail.¹⁵ This activity was reviewed by CDC and was conducted consistent with applicable federal law and CDC

policy. See e.g., 45 C.F.R. part 46, 21 C.F.R. part 56; 42 U.S.C. §241(d); 5 U.S.C. §552a; 44 U.S.C. §3501 et seq.

We administered questionnaires to household contacts to assess the presence of 15 symptoms during the 14 days prior to or at enrollment (day 0). Additionally, participants completed a daily symptom diary during days 1–14 after enrollment. We collected serum and upper respiratory specimens (i.e., both nasopharyngeal [NP] and anterior nares swabs) on day 0 and day 14. We additionally collected NP swabs at any interim date if any household contact newly developed or had worsening of any one of 15 symptoms consistent with COVID-19: nasal congestion or runny nose, sore throat, cough, chest pain, shortness of breath, discomfort while breathing, wheezing, headache, new loss of taste or smell, fever/chills, fatigue, muscle aches, diarrhea (≥ 3 loose stools per day), abdominal pain, or nausea/vomiting. The Milwaukee Health Department and Utah Public Health Laboratories tested the swabs using the CDC real-time Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) assay for SARS-CoV-2,¹⁶ and CDC tested sera using a CDC-developed SARS-CoV-2 enzyme-linked immunosorbent assay (ELISA).¹⁷

Definitions

We defined a household contact to be a COVID-19 case if they had at least one specimen test positive for SARS-CoV-2 by RT-PCR. We classified persons <18 years of age as children, and persons ≥ 18 years of age as adults. We combined all symptoms recorded at any time prior to enrollment (and after the index case's symptom onset date) through the end of the 14-day observation period. We assessed individual symptoms, existing symptom combinations, and newly constructed symptom combinations for their association with SARS-CoV-2 test result by RT-PCR (Table 1). We asked enrollees to state whether they experienced any loss of taste and, separately, smell during the specified time period. For enrollees who responded yes to this question, we then asked whether the loss was partial or complete. We defined loss and/or dysfunction of taste or smell to include any level of loss, whether partial or complete. For ARI, we interpreted coryza as runny nose or nasal congestion.

Analytic Methods

We excluded household contacts from the main analysis if not present at enrollment or not completing the study procedures. Our analysis of combinations predictive of COVID-19 included all 15 symptoms surveyed. We formally described the diagnostic performance of each individual symptom, existing COVID-19 case definitions, respiratory illness case definitions, and newly constructed symptom combinations (Table 1). The goal of assessing symptom combinations was to accurately divide the population into two groups: those who tested positive for SARS-CoV-2 and those who tested negative. For a given combination, we calculated the association of the combination with respect to laboratory-confirmed SARS-CoV-2 infection, yielding the number of contacts who were true positive (TP) (i.e., positive symptom profile and positive test), false negative (FN) (i.e., negative symptom profile and positive test), false positive (FP) (i.e., positive symptom profile and negative test), or true negative (TN) (i.e., negative symptom profile and negative test). From these values, we calculated the symptom combination's sensitivity (i.e., $TP / [TP + FN]$), specificity (i.e., $TN / [TN + FP]$), positive predictive value (PPV) (i.e., $TP / [TP + FP]$), negative predictive value (NPV) (i.e., $TN / [TN + FN]$), F1 score (the harmonic mean of sensitivity and PPV), and Youden's index ($[sensitivity + specificity] - 100$). To determine how well each definition would estimate prevalence in a syndromic surveillance system, we also calculated the difference in the number of positive symptom screens (i.e., $TP + FP$) from the actual number of contacts who tested positive by RT-PCR. We assessed combinations across all ages and in children and adults separately, and we reported all measures on the percentage scale. To assess variability in each performance measure, we constructed bias-corrected and accelerated bootstrap confidence intervals¹⁸ over 10,000 pseudosamples constructed by resampling households with replacement. We reported 95% confidence intervals with two exceptions. For measures estimated at 100% in the observed sample, we omitted confidence intervals, because the pseudosamples could not exhibit any variability. For the difference in specificity and sensitivity between adults and children, we reported 97.5% confidence intervals (a Bonferroni correction) to allow for a 95% joint confidence level regarding the differences in each pair.

We adapted innovative methods previously applied in the low-resource context to derive a parsimonious symptom combination to prioritize diagnostic testing for tuberculosis.¹⁹ We chose this approach to be as comprehensive as practical for COVID-19 in that it systematically assessed nearly every conceivable combination of symptoms.

First, we searched over 245,000 combinations of between one and 15 symptoms (i.e., simple combinations of the form “at least m symptoms present out of n symptoms considered”). We gave greater weight to combinations with high F-1 score or high Youden’s index. We then conducted an exhaustive search using pairs of these “ m -of- n ” combinations (i.e., compound combinations) to allow for more nuanced combinations. We limited this second search to single combinations of no more than five symptoms, such that the number of total symptoms evaluated for a compound combination was never more than ten. We allowed each pair of combinations to be joined by the logical operators [AND] and [OR], yielding approximately 73 million unique combination pairs. After the search, we selected four combination pairs to include in the primary analysis on the basis of diagnostic performance and parsimony. We measured diagnostic performance by F-1 score (higher being better). We measured parsimony by the total number of symptoms evaluated (fewer being better) (Table 1).

We performed all calculations in R 4.0.0 (R Core Team), Python 3.7 (Python Software Foundation), or both. To allow for parallel processing, the exhaustive combinatorial search and bootstrap confidence intervals were implemented on a scientific workstation with 24 logical cores and 64 GB of RAM. De-identified data and analytic scripts in R and Python are publicly available through a GitHub repository: <https://github.com/scottlee/covid-casedefs>.

Findings

Study Population

We enrolled 199 contacts of index patients with laboratory-confirmed SARS-CoV-2 infections within 62 households. We excluded one contact who was not living in the home on the day of enrollment, one who was hospitalized at enrollment, and two who did not consent to have specimens collected. Ten contacts had negative RT-PCR and positive serology test results; they were also excluded from the primary analyses. Therefore, our analyses included the remaining 185 household contacts. The median time interval from index patient’s symptom onset date to enrollment was 10 days (interquartile range [IQR]: 7–13). About half (95; 51%) were female. 108 (58%) were Caucasian/white, 32 (17%) Latinx/Hispanic, 23 (12%) African American/black, 14 (8%) Asian, four (2%) Native American, and four (2%) multiracial. The median age was 22 years ([IQR]: 14–47), with 122 (66%) adults and 63 (34%) children. Among children, nine (14%) were <5 years, 19 (30%) were 5–9 years, and 35

(56%) were 10–17 years of age. SARS-CoV-2 infection was detected by RT-PCR in 49 (27%) household contacts. Separated by age group, 35/122 (29%) adults and 14/63 (22%) children had laboratory-confirmed SARS-CoV-2 infection by RT-PCR. Among the 49 RT-PCR-positive contacts, most (45; 92%) also had a positive serology result, three had a negative serology result, and one was not tested by serology.

Performance Characteristics for Individual Symptoms (All Ages Pooled)

Individual symptoms with the highest sensitivity were nasal congestion or rhinorrhea, headache, and cough (Table 2, Figure 1). Many of the individual symptoms reported were highly specific, although generally resulting in lower sensitivity. The exception was loss or dysfunction of taste or smell (categorized as a single symptom), which had a moderate sensitivity of 63% (95% CI 47–77%), high specificity (96%; 95% CI 90–99%), high NPV (88%; 95% CI 80–93%), and the highest PPV (84%; 95% CI 64–94%), Youden's index (59%; 95% CI 42–73%), and F1 score (72%; 95% CI 57–83%).

Performance Characteristics for Existing COVID-19 Case Definitions, the CDC Symptom List, and Respiratory Syndromic Surveillance Networks (All Ages Pooled)

Among the existing case definitions, sensitivity was perfect (100%) for the CDC symptom list definition, and also high for ARI (96%; 95% CI 86–100%), CSTE combination 1 (original); 94%; 95% CI 82–98%), CSTE combination 2 (revised); 98%; 95% CI 88–100%), and CLI (86%; 95% CI 70–94%) (Table 2, Figure 1). While these definitions offered high sensitivity, they were poorly specific. Conversely, ILI was highly specific (90%; 95% CI 83–94%) but insensitive (51%; 95% CI 35–65%). All existing definitions demonstrated low PPV. Youden's indices and F1 scores were highest for CSTE combination 1 and CLI, though still suboptimal. None of the existing definitions predicted prevalence well; the difference from true prevalence ranged from –20 for ILI to 196 for the CDC symptom list. Compared to CSTE combination 1, CSTE combination 2 had slightly higher sensitivity and NPV, but performed more poorly on all other diagnostic performance indicators.

Performance Characteristics for Derived Compound Symptom Combinations (All Ages Pooled)

The four highest performing novel symptom combinations, based on F1 score and parsimony, were compound symptom combinations that included dysfunction of taste or smell. These four combinations performed similarly well on all performance measures (Table 2, Figure 1). We determined compound symptom combination 3 (i.e., loss or dysfunction of taste or smell, or at least two of the following: shortness of breath, wheezing, discomfort breathing, or fever/chills), to be simple to implement, have higher specificity (93%; 95% CI 85–97%), NPV (92%; 95% CI 86–96%), PPV (79%; 95% CI 60–91%), Youden’s index (70%; 95% CI 54–82%), and F1 score (78%; 95% CI 66–87%) than existing case definitions. The compound symptom combination 3 showed near-perfect prevalence prediction (92%; 95% CI 82–94), and sensitivity was moderately high (78%; 95% CI 63–88%).

Adult-Child Differences in Discriminatory Performance

The accuracy of symptom profiles for defining RT-PCR confirmed COVID-19 differed by age (Table 3, Figure 2). Overall, existing case definitions were less sensitive in children compared to adults. One exception, the CDC symptom list for priority testing (Table 1), captured all COVID-19 cases regardless of age. The existing case definitions were more specific in children, but the greater specificity was statistically significant for CSTE combination 1 only. Individual symptoms showed a similar pattern of lower sensitivity among children, notably taste/smell dysfunction. Sore throat was more sensitive in children, and fever/chills and nausea were similar regardless of age group. We observed a similar pattern of increased specificity for most derived symptom combinations in children (Table 3, Figure 2). Cough was the sole symptom where the difference in both sensitivity and specificity was statistically significant. For both children and adults, the CLI case definition provided the greatest balance between both sensitivity and specificity (Youden’s Index 53%; 95% CI 8–80% vs. 52%; 95% CI 36–66%, respectively) and harmonization of sensitivity with PPV (F1 61%; 95% CI 26–83% vs. 63%; 95% CI 49–76%, respectively) (Table 2). CLI also most accurately predicted overall prevalence amongst children (percent difference from true prevalence 36%; 95% CI 17–157) (Table 3).

Interpretation

Existing case surveillance definitions for COVID-19, as shown in Table 1, were generally sensitive in our study conducted among household contacts of infected persons, a population with proven SARS-CoV-2 exposure.

However, they tended to have low specificity and poorly estimated disease prevalence. By systematically screening novel definitions that optimized sensitivity, specificity, and PPV, we improved community prevalence estimation and overall accuracy of individual screening, which could be useful if diagnostic testing is limited. In particular, we affirmed loss or dysfunction of taste or smell as a uniquely discerning characteristic central to constructing an effective, concise case surveillance definition when applied across all age groups (i.e., derived compound combination 3).

An appropriate discriminatory balance between sensitivity and specificity for a newly emerging pathogen depends on the objectives of the surveillance activity.²⁰ Highly sensitive case definitions capture a larger proportion of true COVID-19 cases, which is ideal when diagnostic testing is widely available and results are timely. Highly sensitive definitions, however, generally rule in a larger number of non-cases (i.e., FP symptom screens).²¹ In addition to testing resources, the public health system's tolerance for false-positive screens is, of course, dependent on human resources. This is especially apparent when intensive interventions involve extensive contact tracing, isolation and quarantine. At high community COVID-19 prevalence, these intensive mitigation efforts may benefit from evidence-based prioritization. By example, CSTE combination 2 had a FP symptom screening rate (77/136; 57%) eight times that for derived compound combination 3 (10/136; 7%) in our cohort. At the population level, such differences could expose shortcomings in resources for core interventions, such as universal contact tracing. The COVID-19 response has repeatedly been strained in these requisite areas.²²⁻²⁴ Novel symptom screening criteria that more tightly couple sensitivity and specificity (i.e., diagnostic accuracy), such as the derived compound combination 3, could help to prioritize interventions when strategically deployed. This principle may also apply when evaluating novel vaccines or therapeutics in large clinical trials involving thousands of participants, where feasibility constraints often dictate the use of symptom-prioritized testing to confirm outcomes. Still, highly sensitive symptom rules, such as CSTE combination 2, are preferred for COVID-19 when resources are unlimited.

For using syndromic surveillance systems to estimate community burden, the highly sensitive existing case definitions overestimated true burden. Conversely, highly specific case definitions, such as ILI, may detect changes in disease trends over time but underestimate true burden.²⁰ ILI underestimated disease prevalence by

more than 80% in this study population. Current laboratory-based surveillance grossly under-ascertains incidence,²⁵ especially where diagnostic testing is not easily accessible or widespread. Retailoring community-based syndromic surveillance systems already in place²⁶ (i.e., altering the symptoms included or applying a correction factor based on results such as ours) would more accurately reflect true burden.

For most symptoms and their combinations, overall performance, most notably sensitivity, differed between children and adults. These findings are consistent with prior observations whereby children generally show fewer and milder symptoms of COVID-19 compared with adults,²⁷ and COVID-19 syndromes vary across ages.¹³ The small number of children with COVID-19 in this cohort limits the conclusion of specific recommendations, but further examination into the utility of age-specific case definitions is warranted in considering policies for schools and other child congregate settings, and for deriving accurate burden estimates from syndromic surveillance.

While the number of individuals in this study is relatively small, our investigation is among the largest and most well-characterized of its kind to date. We collected extensive symptom data, which yielded a comprehensive assessment of multiple symptom combinations. We also employed inclusion criteria that were not based on disease status or symptom status, and a reference category based on standardized laboratory testing. Nonetheless, we acknowledge this study's limitations. These analyses were not intended to produce definitive symptom combinations to be applied to the general public, however they may be used to guide the development of future candidate case definitions. One key consideration for future validation efforts is that enrollment started immediately after the precipitous decline in laboratory-confirmed influenza virus infections in the United States in mid-March 2020.²⁸ Thus, our estimates of diagnostic performance may differ during the viral respiratory season. In addition, COVID-19 prevalence was higher for our study population (i.e., contacts of laboratory-confirmed household members) compared to the entire community, thereby limiting the generalizability of predictive values (although sensitivity and specificity remain unaffected by disease prevalence). Screening criteria applied to persons seeking medical care may also perform differently, as those individuals probably tend to have more severe illness.

Additionally, we showed that existing COVID-19 case definitions are highly sensitive and do well to screen in persons for testing and individual-level public health interventions like community isolation. In the first such

endeavor for evaluating and deriving novel COVID-19 case surveillance definitions in a community setting among SARS-CoV-2—exposed individuals with largely mild illness, we evaluated novel symptom combinations for COVID-19 using methodology previously applied to tuberculosis in low resource settings.¹⁹ These derived combinations and CSTE definition 2 better estimated community disease burden and used taste and/or smell dysfunction as a primary component. The latter is supported by prior studies.^{5,8-10} Because most SARS-CoV-2 infections are mild²⁹ and core public health functions may need prioritization when testing and other resources are limited, case definitions that accurately determine COVID-19 status in the general public may assist continued interruption of community transmission.³⁰ When timely diagnostic testing is readily available, however, using less sensitive screening tools could inappropriately miss cases and lead to further community transmission.

Our study population, which includes persons across the age spectrum and enrolled participants independent of disease and symptom status, may better reflect the diagnostic performance in the general population than previously published research. Accurate clinical case definitions are likely to also apply to large clinical trials for candidate vaccines and therapeutics where serial confirmatory SARS-CoV-2 testing for any new symptom is impractical. It is important that our results be validated against the growing body of larger ambulatory surveillance databases in diverse communities and in other countries; in particular, our methodology should be assessed in the context of the annual influenza season, at varying community COVID-19 prevalence, and across the age spectrum. Such studies ideally can be accompanied by cost-effectiveness modeling of intervention strategies.

U.S. COVID-19 Household Investigation Team

Michelle Banks, MS*, Katherine A. Battey, MPH*, Allison Binder, MPH*, Sean Buono, PhD*[‡], Rebecca J. Chancey, MD*, Ann Christiansen, MPH*, Erin E. Conners, PhD*, Trivikram Dasu, PhD[¶], Patrick Dawson, PhD*[‡], Elizabeth Dietrich, PhD*, Lindsey M. Duca, PhD*[‡], Angela C. Dunn, MD[§], Victoria L. Fields, DVM*[‡], Garrett Fox, MPH*, Brandi D. Freeman, PhD*[‡], Radhika Gharpure, DVM*[‡], Christopher Gregory, MD*, Tair Kiphibane, BSN^{††}, Rebecca L. Laws, PhD*, Sandra Lester, PhD*, Nathaniel M. Lewis, PhD*^{‡,§}, Perrine Marcenac, PhD*[‡], Almea M. Matanock, MD*, Lisa Mills, PhD*, Henry Njuguna, MD*, Michelle O'Hegarty, PhD*, Daniel Owusu, DrPH*[‡], Lindsey Page, MPH[¶], Lucia Pawloski, PhD*, Eric Pevzner, PhD*, Mary Pomeroy, MSN*[‡], Ian W. Pray, PhD*^{‡,□}, Elizabeth M. Rabold, MD*[‡], Jared R. Rispens, MD*[‡], Phillip Salvatore, PhD*[‡], Amy Schumacher, PhD*, Cuc H. Tran, PhD*, Jeni Vuong, MS*, Ashutosh Wadhwa, PhD*[‡], Ryan P. Westergaard, MD, PhD*, Sarah Willardson, MPH^{‡‡}, Dongni Ye, PhD*, Sherry Yin, MPH*, Anna Yousaf, MD*[‡]

*COVID-19 Response Team, CDC

[†]Epidemic Intelligence Service, CDC

[‡]Laboratory Leadership Service, CDC

[§]Utah Department of Health

[□]Wisconsin Department of Health

[¶]City of Milwaukee Health Department

^{**}North Shore (WI) Health Department

^{††}Salt Lake City (UT) Health Department

^{‡‡}Davis County (UT) Health Department

Author Contributions

All authors contributed to study design, data collection, scientific interpretation, and review and approval of the final manuscript. In addition, the following authors made unique contributions:

Scott H. Lee and Charles M. Heilig conducted all data analysis, provided open analytic code for public dissemination, led data interpretation, wrote data-analytic methods, and assisted with integrating analysis and interpretation throughout the manuscript.

Robyn Atkinson-Dunn, Sanjib Bhattacharyya, Trivikram Dasu, Kim Christensen, Brandi D. Freeman, Sandra Lester, Lisa Mills, Lucia Pawloski, and Natalie J. Thornburg coordinated or performed diagnostic laboratory analyses.

Katherine A. Battey, Sean Buono, Rebecca J. Chancey, Ann Christiansen, Erin E. Connors, Patrick Dawson, Elizabeth Dietrich, Lindsey M. Duca, Victoria L. Fields, Radhika Gharpure, Christopher Gregory, Tair Kiphibane, Rebecca L. Laws, Nathaniel M. Lewis, Perrine Marcenac, Almea M. Matanock, Henry Njuguna, Michelle O’Hegarty, Daniel Owusu, Lindsey Page, Eric Pevzner, Mary Pomeroy, Ian W. Pray, Elizabeth M. Rabold, Jared R. Rispens, Phillip Salvatore, Cuc H. Tran, Jeni Vuong, Ashutosh Wadhwa, Sarah Willardson, Sherry Yin, Anna Yousaf implemented the study protocol, developed standards of practice, and collected data and specimens from the field.

Michelle Banks, Angela C. Dunn, Alicia Fry, Aron J. Hall, Hannah L. Kirking, Scott A. Nabity, Amy Schumacher, Jacqueline E. Tate, Ryan P. Westergaard, and Dongni Ye provided overall project coordination and leadership.

Hannah E. Reses, Mark Fajans, Garrett Fox, Allison Binder, and Victoria T. Chu curated the data.

Hannah E. Reses, Mark Fajans, Scott H. Lee, Charles M. Heilig, and Scott A. Nabity drafted the original manuscript.

Acknowledgments

Utah:

Davis County (UT) Health Department: Heather Gibb, Sara Hall

Summit County (UT) Health Department: Carolyn Rose

Salt Lake County (UT) Health Department: Dagmar Vitek, Ilene Risk, Lee Cherie Booth, Jeff Sanchez, Madison Clawson, Tara Scribellito, Ha Khong, Carlene Claflin, Theresa Beesley, Victoria Castaneda

Utah Department of Health: Nathan LaCross, Robyn Atkinson-Dunn, COVID-19 response team

Wisconsin:

North Shore Health Department: Kala Hardy, Christine Cordova, Kevin Rorabeck, Kathleen Platt

City of Milwaukee Health Department: Catherine Bowman, Nancy Burns, Barbara Coyle, Elizabeth Durkes,

Carol Johnsen, Jill LeStarge, Erica Luna-Vargas, Sholonda Morris, Mary Jo Gerlach, Jill Paradowski, Bill Rice,

Michele Robinson, Virginia Thomas, Keara Jones, Chelsea Watry, Richard Weidensee, Jeanette Kowalik, Heather Paradis, Julie Katrichis

Wauwatosa Health Department: Laura Conklin, Paige Bernau, Emily Tianen

Wisconsin Department of Health Services: COVID-19 response team

City of Milwaukee Laboratory SARS-CoV-2 Testing Team: Jordan Hilleshiem, Beth Pfothenauer, Manjeet Khubbar, Jennifer Lentz, Zorangel Amezcua-Montes, Kristin Schieble, Noah Leigh, Joshua Weiner, Tenysa Guzman, Kathy Windham, and Julie Plevak

Centers for Disease Control and Prevention:

CDC COVID-19 Response Team: Claire Midgley, Melissa Rolfes, Mayer Antoine, Adebowale Ojo, Fiona Havers

Laboratory Task Force: Brandi Limbago, Michelle Owens, Wendi Kuhnert-Tallman, Jeff Johnson, Collette Leumont, Gayle Langley

Field Investigations Team: Jenny Milucky, Emily Weston, Jessica Smith, Michelle Johnson Jones, Jennifer Huang, Erin Moritz, Aubrey Gilliland, Laura Calderwood, Jennifer Imaa, Kaytlin Renfro, Allison Miller, Katie Bantle, Margaret Williams, Stacy Thorne, Jana Manning, Micha Ghertner, Adriane Niare, Yoonjae Kang, Charlotte D. Kaboré

References

1. Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* 2020; **382**(8): 727-33.
2. Tsang TK, Wu P, Lin Y, Lau EHY, Leung GM, Cowling BJ. Effect of changing case definitions for COVID-19 on the epidemic curve and transmission parameters in mainland China: a modelling study. *Lancet Public Health* 2020; **5**(5): e289-e96.
3. Pascarella G, Strumia A, Piliago C, et al. COVID-19 diagnosis and management: a comprehensive review. *J Intern Med* 2020; **288**(2): 192-206.
4. Tostmann A, Bradley J, Bousema T, et al. Strong associations and moderate predictive value of early symptoms for SARS-CoV-2 test positivity among healthcare workers, the Netherlands, March 2020. *Euro Surveill* 2020; **25**(16).
5. Menni C, Valdes AM, Freidin MB, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. *Nat Med* 2020; **26**(7): 1037-40.
6. Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020; **395**(10223): 507-13.
7. Varatharaj A, Thomas N, Ellul MA, et al. Neurological and neuropsychiatric complications of COVID-19 in 153 patients: a UK-wide surveillance study. *Lancet Psychiatry* 2020.
8. Lan FY, Filler R, Mathew S, et al. COVID-19 symptoms predictive of healthcare workers' SARS-CoV-2 PCR results. *PLoS One* 2020; **15**(6): e0235460.
9. Clemency BM, Varughese R, Scheafer DK, et al. Symptom Criteria for COVID-19 Testing of Health Care Workers. *Acad Emerg Med* 2020; **27**(6): 469-74.
10. Ahmed MAM, Colebunders R, Siewe Fodjo JN. Evidence for important COVID-19 community transmission in Somalia using a clinical case definition. *Int J Infect Dis* 2020.
11. Struyf T, Deeks JJ, Dinnes J, 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* 2020; **7**: CD013665.
12. Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ* 2020; **369**: m1328.
13. Ludvigsson JF. Systematic review of COVID-19 in children shows milder cases and a better prognosis than adults. *Acta Paediatr* 2020; **109**(6): 1088-95.
14. Casalegno JS, Eibach D, Valette M, et al. Performance of influenza case definitions for influenza community surveillance: based on the French influenza surveillance network GROG, 2009-2014. *Euro Surveill* 2017; **22**(14).
15. Lewis N, Chu V. Household transmission of SARS-CoV-2 in the United States. *Clin Infect Dis* 2020.
16. CDC. 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel. 2020.
17. Freeman B, Lester S, Mills L, et al. Validation of a SARS-CoV-2 spike protein ELISA for use in contact investigations and serosurveillance. *bioRxiv* 2020.
18. Efron B. Better bootstrap confidence intervals. *J Am Stat Assoc* 1987; **82**(397): 171-85.
19. Cain KP, McCarthy KD, Heilig CM, et al. An algorithm for tuberculosis screening and diagnosis in people with HIV. *N Engl J Med* 2010; **362**(8): 707-16.
20. Giesecke J. Modern infectious disease epidemiology. New York: Oxford University Press; 2002.
21. Ngobeni H, Tempia S, Cohen AL, et al. The performance of different case definitions for severe influenza surveillance among HIV-infected and HIV-uninfected children aged <5 years in South Africa, 2011-2015. *PLoS One* 2019; **14**(9): e0222294.
22. Keeling MJ, Hollingsworth TD, Read JM. Efficacy of contact tracing for the containment of the 2019 novel coronavirus (COVID-19). *J Epidemiol Community Health* 2020.
23. Watson C, Cicero A, Blumenstock J, Fraser M. A national plan to enable comprehensive COVID-19 case finding and contact tracing in the U.S. Baltimore, MD: Johns Hopkins University, 2020.

24. Hellewell J, Abbott S, Gimma A, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob Health* 2020; **8**(4): e488-e96.
25. Havers FP, Reed C, Lim T, et al. Seroprevalence of antibodies to SARS-CoV-2 in 10 sites in the United States, March 23-May 12, 2020. *JAMA Intern Med* 2020; doi:**10.1001/jamainternmed.2020.4130**: E1-E11.
26. Chan AT, Brownstein JS. Putting the Public Back in Public Health - Surveying Symptoms of Covid-19. *N Engl J Med* 2020.
27. Davies NG, Klepac P, Liu Y, et al. Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat Med* 2020.
28. CDC. Influenza Surveillance. Available from [cdc.gov/flu/weekly/index.htm](https://www.cdc.gov/flu/weekly/index.htm) (Accessed 4 August 2020). 2020.
29. To KK, Tsang OT, Leung WS, et al. Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by SARS-CoV-2: an observational cohort study. *Lancet Infect Dis* 2020; **20**(5): 565-74.
30. Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* 2020; **368**(6490): 489-93.

Tables

Table 1. Existing COVID-19 case definitions, respiratory illness surveillance case definitions, and derived compound symptom combinations assessed for diagnostic performance in a community cohort of 185 individuals with household COVID-19 exposure in Utah and Wisconsin, United States, March–May 2020

Case definition category	Case definition	Criteria
Existing COVID-19 case definitions	U.S. Centers for Disease Control and Prevention (CDC) COVID-19-like illness (CLI)*	Fever, cough, or shortness of breath
	CDC symptom list [†]	Fever or chills, cough, shortness of breath or difficulty breathing, fatigue, muscle or body aches, headache, new loss of taste or smell, sore throat, congestion or runny nose, nausea or vomiting, diarrhea
	U.S. Council of State and Territorial Epidemiologists (CSTE) COVID-19 original clinical criteria (CSTE combination 1) [‡]	At least one of the following: cough, shortness of breath, or discomfort breathing, OR at least two of the following: fever, chills, rigors, myalgia, headache, sore throat, new olfactory disorder and taste disorders
	CSTE COVID-19 revised clinical criteria (CSTE combination 2) [§]	At least one of the following: cough, shortness of breath, discomfort breathing, new olfactory disorder, new taste disorder, or at least two of the following: fever, chills, rigors, myalgia, headache, sore throat, nausea or vomiting, diarrhea, fatigue, congestion or runny nose
Existing respiratory illness surveillance case definitions	Influenza-like illness (ILI) [□]	Fever AND cough and/or sore throat
	World Health Organization (WHO) acute respiratory infection (ARI) definition for community-based respiratory syncytial virus (RSV) surveillance [¶]	At least one of the following: shortness of breath or cough, sore throat, or coryza
Derived compound symptom combinations**	Combination 1	Taste and/or smell dysfunction, OR one of the following: shortness of breath, myalgia, or fever or chills
	Combination 2	Taste and/or smell dysfunction or discomfort breathing, OR at least two of the following: shortness of breath, wheezing, or fever or chills
	Combination 3	Taste and/or smell dysfunction, OR at least two of the following: shortness of breath, wheezing, discomfort breathing, or fever or chills
	Combination 4	Taste and/or smell dysfunction, OR shortness of breath and fever or chills

*U.S. Centers for Disease Control and Prevention (CDC) COVID-19-like illness (CLI) definition was used to guide early diagnostic testing strategies from 17 January 2020 – 08 March 2020

(<https://emergency.cdc.gov/han/han00426.asp>).[†]U.S. Centers for Disease Control and Prevention (CDC) list of symptoms that may indicate COVID-19 infection (<https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html>). This symptom list was last updated on 13 May 2020.

[‡]U.S. Council of State and Territorial Epidemiologists (CSTE) original clinical criteria for COVID-19 reporting (https://cdn.ymaws.com/www.cste.org/resource/resmgr/2020ps/interim-20-id-01_covid-19.pdf). This interim position statement (Interim-20-ID-01) was approved on 05 April 2020 and was replaced by Interim-20-ID-02 on 07 August 2020.

[§]U.S. Council of State and Territorial Epidemiologists (CSTE) revised clinical criteria for COVID-19 reporting (https://cdn.ymaws.com/www.cste.org/resource/resmgr/ps/positionstatement2020/interim-20-id-02_COVID-19.pdf). This interim position statement (Interim-20-ID-02) was approved on 07 August 2020 and replaced Interim-20-ID-01.

[□]Influenza-like illness (ILI) outpatient visit information collected through the U.S. Outpatient Influenza-like Illness Surveillance Network (ILINet) (https://www.cdc.gov/flu/weekly/overview.htm#anchor_1539281266932). This collaborative effort between CDC, state and local health departments, and healthcare providers has been tracking patients with ILI since the 1997-1998 influenza season.

[¶]World Health Organization (WHO) acute respiratory infection (ARI) definition for community-based respiratory syncytial virus (RSV) surveillance (https://www.who.int/influenza/rsv/rsv_case_definition/en/). Last updated 04 February 2020.

^{**}Compound symptom combinations were derived from all symptoms recorded at any time prior to enrollment (and after the index case's symptom onset date) through the end of the 14-day observation period.

Table 2. Diagnostic performance indicators for individual COVID-19 symptoms, existing case definitions, and derived compound symptom combinations for a community cohort of 185 people with household COVID-19 exposure in Utah and Wisconsin, United States, March–May 2020

	SARS-CoV-2-negative (n=136)		SARS-CoV-2-positive (n=49)		Specificity*		Sensitivity [†]		NPV [‡]		PPV [§]		Youden's index [¶]		F-1 score		Difference in prevalence**	
	TN ^{††}	FP ^{‡‡}	FN ^{§§}	TP ^{¶¶}	Value	(95% CI)	Value	(95% CI)	Value	(95% CI)	Value	(95% CI)	Value	(95% CI)	Value	(95% CI)	Value	(95% CI)
Individual symptoms																		
<i>Upper respiratory</i>																		
Nasal congestion or rhinorrhea	66	70	5	44	49	(38, 59)	90	(80, 96)	93	(85, 97)	39	(27, 50)	38	(26, 50)	54	(42, 65)	133	(78, 227)
Sore throat	96	40	22	27	71	(61, 79)	55	(40, 68)	81	(72, 88)	40	(26, 56)	26	(10, 40)	47	(33, 59)	37	(0, 98)
<i>Lower respiratory</i>																		
Cough	103	33	13	36	76	(65, 84)	74	(60, 85)	89	(81, 94)	52	(37, 67)	49	(36, 62)	61	(49, 72)	41	(8, 94)
Chest pain	122	14	35	14	90	(83, 94)	29	(15, 42)	78	(69, 85)	50	(24, 70)	18	(5, 33)	36	(21, 51)	-43	(-62, -15)
Shortness of breath	125	11	36	13	92	(86, 96)	27	(15, 40)	78	(67, 85)	54	(31, 75)	18	(6, 32)	36	(22, 49)	-51	(-68, -24)
Discomfort while breathing	132	4	37	12	97	(93, 99)	25	(14, 39)	78	(69, 85)	75	(40, 92)	22	(10, 36)	37	(22, 54)	-67	(-80, -49)
Wheezing	133	3	44	5	98	(94, 99)	10	(4, 21)	75	(65, 83)	63	(20, 100)	8	(1, 19)	18	(7, 33)	-84	(-92, -68)
<i>Neurological</i>																		
Headache	86	50	7	42	63	(52, 73)	86	(75, 93)	93	(85, 97)	46	(33, 59)	49	(35, 60)	60	(47, 70)	88	(44, 160)
Taste and/or smell dysfunction	130	6	18	31	96	(90, 99)	63	(47, 77)	88	(80, 93)	84	(64, 94)	59	(42, 73)	72	(57, 83)	-25	(-43, -5)
<i>Constitutional</i>																		
Fever or chills	113	23	18	31	83	(75, 89)	63	(49, 76)	86	(78, 92)	57	(40, 72)	46	(30, 60)	60	(46, 72)	10	(-15, 51)
Fatigue	94	42	20	29	69	(59, 79)	59	(41, 74)	83	(74, 89)	41	(25, 57)	28	(10, 44)	48	(33, 63)	45	(8, 106)
Myalgia	116	20	21	28	85	(77, 91)	57	(42, 70)	85	(77, 90)	58	(41, 74)	42	(28, 56)	58	(44, 69)	-2	(-28, 37)
<i>Gastrointestinal</i>																		
Diarrhea	108	28	31	18	79	(69, 88)	37	(20, 55)	78	(66, 86)	39	(22, 59)	16	(-3, 36)	38	(23, 54)	-6	(-40, 48)
Abdominal pain	120	16	34	15	88	(80, 94)	31	(19, 45)	78	(68, 85)	48	(27, 73)	19	(4, 35)	38	(23, 53)	-37	(-59, -2)
Nausea	124	12	40	9	91	(85, 95)	18	(10, 31)	76	(65, 83)	43	(20, 67)	10	(-1, 22)	26	(14, 40)	-57	(-74, -30)
Existing case definitions																		
<i>COVID-19 and respiratory illness surveillance case definitions</i>																		
CDC symptom list ^{†††}	40	96	0	49	29	(17, 45)	100	NA ^{□□□□}	100	NA	34	(24, 45)	29	(17, 45)	51	(38, 62)	196	(122, 322)
ARI ^{***}	52	84	2	47	38	(26, 51)	96	(86, 100)	96	(88, 100)	36	(25, 48)	34	(23, 45)	52	(40, 64)	167	(102, 279)
CSTE combination 1 ^{†††}	79	57	3	46	58	(47, 69)	94	(82, 98)	96	(91, 99)	45	(32, 58)	52	(39, 63)	61	(47, 72)	110	(64, 190)
CSTE combination 2 ^{†††}	59	77	1	48	43	(35, 52)	98	(88, 100)	98	(90, 100)	38	(30, 47)	41	(32, 50)	55	(46, 64)	155	(110, 225)
CLI ^{§§§}	93	43	7	42	68	(57, 79)	86	(70, 94)	93	(88, 97)	49	(35, 64)	54	(40, 65)	63	(48, 75)	74	(35, 141)
ILI ^{□□□}	122	14	24	25	90	(83, 94)	51	(35, 65)	84	(75, 90)	64	(45, 79)	41	(25, 55)	57	(42, 69)	-20	(-41, 9)
Derived compound symptom combinations^{†††}																		
Combination 1 ^{****}	129	7	12	37	95	(89, 98)	76	(60, 87)	92	(85, 96)	84	(67, 94)	70	(54, 82)	80	(67, 88)	-10	(-28, 9)
Combination 2 ^{††††}	125	11	10	39	92	(84, 97)	80	(65, 89)	93	(86, 96)	78	(59, 91)	72	(56, 83)	79	(66, 88)	2	(-17, 30)
Combination 3 ^{††††}	126	10	11	38	93	(85, 97)	78	(63, 88)	92	(86, 96)	79	(60, 91)	70	(54, 82)	78	(66, 87)	-2	(-20, 24)
Combination 4 ^{§§§§}	129	7	13	36	95	(89, 98)	74	(59, 84)	91	(84, 95)	84	(67, 94)	68	(54, 80)	78	(67, 86)	-12	(-29, 8)

*Specificity is the probability of testing negative when disease is absent.

[†]Sensitivity is the probability of testing positive when disease is present.

[‡]NPV = negative predictive value. NPV is the probability of a patient not having disease when test is negative.

[§]PPV = positive predictive value. PPV is the probability of a patient having disease when test is positive.

[□]Youden's index is defined as sensitivity plus specificity minus 100 (perfect score=100).

[¶]F-1 score is defined as the harmonic mean of sensitivity and PPV (perfect score=100).

**Difference in prevalence is defined as the difference in the number of positive symptom screens (i.e., true positive plus false positive) for each combination from the actual number of contacts who tested positive by RT-PCR.

^{††}TN = true negative.

^{‡‡}FP = false positive.

^{§§}FN = false negative.

^{□□}TP = true positive.

^{¶¶}U.S. Centers for Disease Control and Prevention (CDC) list of symptoms that may indicate COVID-19 infection (<https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html>). This symptom list was last updated on 13 May 2020.

^{***}World Health Organization (WHO) acute respiratory infection (ARI) definition for community-based respiratory syncytial virus (RSV) surveillance (https://www.who.int/influenza/rsv/rsv_case_definition/en/). Last updated 04 February 2020.

^{†††}U.S. Council of State and Territorial Epidemiologists (CSTE) original clinical criteria for COVID-19 reporting (https://cdn.ymaws.com/www.cste.org/resource/resmgr/2020ps/interim-20-id-01_covid-19.pdf). This interim position statement (Interim-20-ID-01) was approved on 05 April 2020 and was replaced by Interim-20-ID-02 on 07 August 2020.

^{‡‡‡}U.S. Council of State and Territorial Epidemiologists (CSTE) revised clinical criteria for COVID-19 reporting (https://cdn.ymaws.com/www.cste.org/resource/resmgr/ps/positionstatement2020/interim-20-id-02_COVID-19.pdf). This interim position statement (Interim-20-ID-02) was approved on 07 August 2020 and replaced Interim-20-ID-01.

^{§§§}U.S. Centers for Disease Control and Prevention (CDC) COVID-19-like illness (CLI) definition was used to guide early diagnostic testing strategies from 17 January 2020 – 08 March 2020 (<https://emergency.cdc.gov/han/han00426.asp>).

^{□□□}Influenza-like illness (ILI) outpatient visit information collected through the U.S. Outpatient Influenza-like Illness Surveillance Network (ILINet) (https://www.cdc.gov/flu/weekly/overview.htm#anchor_1539281266932). This collaborative effort between CDC, state and local health departments, and healthcare providers has been tracking patients with ILI since the 1997-1998 influenza season.

^{¶¶¶}Compound symptom combinations were derived from all symptoms recorded at any time prior to enrollment (and after the index case's symptom onset date) through the end of the 14-day observation period.

^{****}Combination 1: Taste and/or smell dysfunction, OR one of the following: shortness of breath, myalgia, or fever or chills.

^{††††}Combination 2: Taste and/or smell dysfunction or discomfort breathing, OR at least two of the following: shortness of breath, wheezing, or fever/chills.

^{‡‡‡‡}Combination 3: Taste and/or smell dysfunction, OR at least two of the following: shortness of breath, wheezing, discomfort breathing, or fever/chills.

^{§§§§}Combination 4: Taste and/or smell dysfunction, OR shortness of breath and fever/chills.

^{□□□□}NA = not applicable.

Table 3. Diagnostic performance indicators for individual COVID-19 symptoms, existing case definitions, and derived compound symptom combinations by age group for 122 adults and 63 children with household exposure to COVID-19 in Utah and Wisconsin, United States, March–May 2020

	Adults								Children								Difference [□]			
	Specificity*		Sensitivity [†]		Youden's index [‡]		F-1 score [§]		Specificity		Sensitivity		Youden's index		F-1 score		Specificity		Sensitivity	
	Value	(95% CI)	Value	(95% CI)	Value	(95% CI)	Value	(95% CI)	Value	(95% CI)	Value	(95% CI)	Value	(95% CI)	Value	(95% CI)	Value	(95% CI) [†]	Value	(95% CI) [†]
Individual symptoms																				
<i>Upper respiratory</i>																				
Nasal congestion or rhinorrhea	40	(30, 51)	91	(78, 97)	32	(17, 44)	54	(40, 66)	63	(44, 77)	86	(58, 100)	49	(25, 67)	55	(32, 71)	-23	(-44, 2)	6	(-13, 31)
Sore throat	61	(50, 72)	51	(33, 68)	12	(-9, 31)	41	(26, 56)	88	(74, 96)	64	(29, 88)	52	(21, 78)	62	(35, 82)	-27	(-43, -9)	-13	(-47, 28)
<i>Lower respiratory</i>																				
Cough	68	(56, 78)	86	(71, 94)	54	(37, 68)	65	(51, 76)	90	(77, 96)	43	(13, 77)	33	(3, 68)	48	(18, 75)	-22	(-39, -5)	43	(2, 78)
Chest pain	84	(74, 91)	34	(18, 52)	18	(0, 38)	39	(22, 57)	100	NA ^{***}	14	(0, 29)	14	(0, 29)	25	(11, 40)	-16	(-27, -8)	20	(-6, 45)
Shortness of breath	89	(80, 94)	29	(16, 44)	17	(2, 34)	36	(21, 52)	98	(87, 100)	21	(0, 56)	19	(-2, 56)	33	(11, 63)	-9	(-20, 0)	7	(-33, 38)
Discomfort while breathing	95	(90, 99)	29	(13, 48)	24	(8, 43)	41	(21, 61)	100	NA	14	(0, 29)	14	(0, 29)	25	(11, 40)	-5	(-10, -1)	14	(-10, 40)
Wheezing	97	(91, 99)	11	(3, 25)	8	(-1, 21)	19	(6, 37)	100	(100, 100)	7	(0, 33)	7	(0, 33)	13	(6, 18)	-3	(-9, 0)	4	(-18, 18)
<i>Neurological</i>																				
Headache	54	(43, 65)	86	(71, 94)	40	(24, 53)	57	(44, 68)	80	(63, 90)	86	(63, 100)	65	(47, 81)	67	(44, 81)	-26	(-44, -4)	0	(-21, 24)
Taste and/or smell dysfunction	95	(89, 99)	71	(51, 85)	67	(47, 82)	78	(63, 88)	96	(84, 100)	43	(18, 67)	39	(13, 65)	55	(29, 76)	-1	(-8, 10)	29	(-3, 58)
<i>Constitutional</i>																				
Fever or chills	82	(72, 89)	63	(46, 78)	45	(27, 61)	60	(46, 73)	86	(68, 95)	64	(33, 90)	50	(14, 78)	60	(30, 82)	-4	(-20, 16)	-1	(-36, 37)
Fatigue	59	(47, 70)	63	(43, 78)	22	(1, 40)	47	(32, 61)	88	(71, 96)	50	(9, 80)	38	(1, 72)	52	(19, 79)	-29	(-48, -8)	13	(-32, 60)
Myalgia	79	(69, 87)	63	(45, 77)	42	(24, 58)	59	(44, 71)	96	(85, 100)	43	(11, 69)	39	(10, 64)	55	(23, 75)	-17	(-29, -4)	20	(-19, 62)
<i>Gastrointestinal</i>																				
Diarrhea	82	(71, 90)	40	(19, 60)	22	(0, 43)	43	(23, 62)	76	(55, 89)	29	(8, 60)	4	(-24, 27)	27	(10, 47)	6	(-13, 23)	11	(-27, 46)

													37)				30)			
Abdominal pain	89	(80, 94)	37	(21, 54)	26	(8, 44)	45	(27, 61)	88	(73, 97)	14	(0, 57)	2	(-17, 39)	18	(7, 45)	1	(-13, 17)	23	(-20, 49)
Nausea	92	(85, 96)	17	(6, 32)	9	(-3, 25)	25	(10, 43)	90	(77, 96)	21	(4, 62)	11	(-11, 50)	27	(9, 57)	2	(-8, 15)	-4	(-48, 23)
Existing case definitions																				
<i>COVID-19 and respiratory illness surveillance case definitions</i>																				
CDC symptom list**	21	(13, 32)	100	NA	21	(13, 32)	50	(38, 63)	45	(22, 65)	100	NA	45	(22, 65)	51	(29, 68)	-24	(-50, 6)	NA	NA
ARI††	29	(19, 41)	100	NA	29	(19, 41)	53	(40, 65)	55	(34, 72)	86	(58, 100)	41	(17, 60)	50	(29, 68)	-26	(-50, 0)	14	(0, 38)
CSTE combination 1‡‡	46	(34, 58)	97	(81, 100)	43	(30, 55)	59	(45, 70)	80	(63, 91)	86	(43, 100)	65	(33, 83)	67	(38, 84)	-34	(-53, -12)	11	(-7, 45)
CSTE combination 2§§	36	(26, 46)	97	(84, 100)	33	(20, 44)	54	(44, 65)	57	(43, 70)	100	NA	57	(43, 70)	57	(39, 72)	-22	(-43, 0)	-3	(-14, 0)
CLI□□	61	(48, 73)	91	(78, 98)	52	(36, 66)	63	(49, 76)	82	(64, 92)	71	(20, 95)	53	(8, 80)	61	(26, 83)	-21	(-41, 2)	20	(-10, 71)
ILI¶¶	86	(76, 93)	54	(37, 70)	41	(23, 59)	58	(42, 71)	96	(85, 100)	43	(11, 80)	39	(7, 78)	55	(20, 85)	-10	(-22, 2)	11	(-36, 53)
Derived compound symptom combinations***																				
Combination 1†††	94	(88, 98)	83	(64, 94)	77	(58, 89)	84	(70, 92)	96	(84, 100)	57	(22, 83)	53	(17, 81)	67	(29, 86)	-2	(-10, 9)	26	(-10, 64)
Combination 2‡‡‡	90	(81, 96)	86	(67, 96)	75	(56, 87)	81	(67, 90)	96	(84, 100)	64	(38, 85)	60	(31, 83)	72	(43, 89)	-6	(-17, 5)	21	(-10, 51)
Combination 3§§§	91	(82, 96)	83	(64, 94)	74	(55, 86)	81	(66, 90)	96	(84, 100)	64	(38, 85)	60	(31, 83)	72	(43, 89)	-5	(-15, 5)	19	(-14, 49)
Combination 4□□□	94	(88, 98)	83	(64, 94)	77	(58, 89)	84	(70, 92)	96	(84, 100)	50	(17, 67)	46	(16, 69)	61	(30, 80)	-2	(-10, 9)	33	(3, 66)

*Specificity is the probability of testing negative when disease is absent.

†Sensitivity is the probability of testing positive when disease is present.

‡Youden's index is defined as sensitivity plus specificity minus 100 (perfect score=100).

§F-1 score is defined as the harmonic mean of sensitivity and positive predictive value (perfect score=100).

□ Difference in specificity and sensitivity between adults and children.

¶Confidence intervals are jointly 95%, or 97.5% with a Bonferroni correction.

**U.S. Centers for Disease Control and Prevention (CDC) list of symptoms that may indicate COVID-19 infection

(<https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html>). This symptom list was last updated on 13 May 2020.

††World Health Organization (WHO) acute respiratory infection (ARI) definition for community-based respiratory syncytial virus (RSV) surveillance

(https://www.who.int/influenza/rsv/rsv_case_definition/en/). Last updated 04 February 2020.

‡‡U.S. Council of State and Territorial Epidemiologists (CSTE) original clinical criteria for COVID-19 reporting

(https://cdn.ymaws.com/www.cste.org/resource/resmgr/2020ps/interim-20-id-01_covid-19.pdf). This interim position statement (Interim-20-ID-01) was approved on 05 April 2020 and was replaced by Interim-20-ID-02 on 07 August 2020.

§§U.S. Council of State and Territorial Epidemiologists (CSTE) revised clinical criteria for COVID-19 reporting

(https://cdn.ymaws.com/www.cste.org/resource/resmgr/ps/positionstatement2020/interim-20-id-02_COVID-19.pdf). This interim position statement (Interim-20-ID-02) was approved on 07 August 2020 and replaced Interim-20-ID-01.

□□ U.S. Centers for Disease Control and Prevention (CDC) COVID-19-like illness (CLI) definition was used to guide early diagnostic testing strategies from 17 January 2020 – 08 March 2020 (<https://emergency.cdc.gov/han/han00426.asp>).

¶¶ Influenza-like illness (ILI) outpatient visit information collected through the U.S. Outpatient Influenza-like Illness Surveillance Network (ILINet) (https://www.cdc.gov/flu/weekly/overview.htm#anchor_1539281266932). This collaborative effort between CDC, state and local health departments, and healthcare providers has been tracking patients with ILI since the 1997-1998 influenza season.

***Compound symptom combinations were derived from all symptoms recorded at any time prior to enrollment (and after the index case's symptom onset date) through the end of the two-week observation period.

††† Combination 1: Taste and/or smell dysfunction, OR one of the following: shortness of breath, myalgia, or fever or chills.

‡‡‡ Combination 2: Taste and/or smell dysfunction or discomfort breathing, OR at least two of the following: shortness of breath, wheezing, or fever/chills.

§§§ Combination 3: Taste and/or smell dysfunction, OR at least two of the following: shortness of breath, wheezing, discomfort breathing, or fever/chills.

□□□ Combination 4: Taste and/or smell dysfunction, OR shortness of breath and fever/chills.

¶¶¶ NA = not applicable.

1 **Supplemental Table 1. Individual COVID-19 symptoms and existing case definitions by 2019-nCoV Real-**
 2 **Time RT-PCR assay and a SARS-CoV-2 spike protein enzyme-linked immunosorbent assay (ELISA)**
 3 **results in Utah and Wisconsin, United States, March–May 2020**

	All participants (N=195)		RT-PCR*-positive (N=49)		RT-PCR-negative and ELISA [†] -positive (N=10)		RT-PCR-negative and ELISA-negative (N=128)	
	n	Percent	n	Percent	n	Percent	n	Percent
<i>Upper respiratory</i>								
Nasal congestion or rhinorrhea	118	61	44	90	4	40	67	52
Sore throat	70	36	27	55	3	30	38	30
<i>Lower respiratory</i>								
Cough	76	39	36	74	7	70	30	23
Chest pain	28	14	14	29	0	0	14	11
Shortness of breath	25	13	13	27	1	10	10	8
Discomfort while breathing	16	8	12	25	0	0	4	3
Wheezing	8	4	5	10	0	0	3	2
<i>Neurological</i>								
Headache	98	50	42	86	6	60	48	38
Taste and/or smell dysfunction	39	20	31	63	2	20	5	4
<i>Constitutional</i>								
Fatigue	72	37	29	59	1	10	40	31
Fever or chills	59	30	31	63	5	50	21	16
Myalgia	51	26	28	57	3	30	21	16
<i>Gastrointestinal</i>								
Diarrhea	48	25	18	37	2	20	24	19
Abdominal pain	32	16	15	31	1	10	14	11
Nausea	23	12	9	18	2	20	10	8
<i>COVID-19 and respiratory illness surveillance case definitions</i>								
CDC symptom list [‡]	155	80	49	100	10	100	91	71
ARI [§]	140	72	47	96	9	90	80	63
CSTE combination 1 [□]	112	57	46	94	9	90	54	42
CSTE combination 2 [¶]	134	69	48	98	9	90	72	56
CLI ^{**}	94	48	42	86	9	90	40	31
ILI ^{††}	42	22	25	51	3	30	13	10

4 *RT-PCR = Reverse Transcriptase Polymerase Chain Reaction.

5 [†]ELISA = enzyme-linked immunosorbent assay.

6 [‡]U.S. Centers for Disease Control and Prevention (CDC) list of symptoms that may indicate COVID-19 infection
 7 (<https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html>). This symptom list was last
 8 updated on 13 May 2020.

9 [§]World Health Organization (WHO) acute respiratory infection (ARI) definition for community-based respiratory
 10 syncytial virus (RSV) surveillance (https://www.who.int/influenza/rsv/rsv_case_definition/en/). Last updated 04
 11 February 2020.

12 [□]U.S. Council of State and Territorial Epidemiologists (CSTE) original clinical criteria for COVID-19 reporting
 13 (https://cdn.ymaws.com/www.cste.org/resource/resmgr/2020ps/interim-20-id-01_covid-19.pdf). This interim
 14 position statement (Interim-20-ID-01) was approved on 05 April 2020 and was replaced by Interim-20-ID-02 on
 15 07 August 2020.

16 [¶]U.S. Council of State and Territorial Epidemiologists (CSTE) revised clinical criteria for COVID-19 reporting
 17 ([https://cdn.ymaws.com/www.cste.org/resource/resmgr/ps/positionstatement2020/interim-20-id-02_COVID-](https://cdn.ymaws.com/www.cste.org/resource/resmgr/ps/positionstatement2020/interim-20-id-02_COVID-19.pdf)
 18 [19.pdf](https://cdn.ymaws.com/www.cste.org/resource/resmgr/ps/positionstatement2020/interim-20-id-02_COVID-19.pdf)). This interim position statement (Interim-20-ID-02) was approved on 07 August 2020 and replaced
 19 Interim-20-ID-01.

20 ^{**}U.S. Centers for Disease Control and Prevention (CDC) COVID-19-like illness (CLI) definition was used to
 21 guide early diagnostic testing strategies from 17 January 2020 – 08 March 2020
 22 (<https://emergency.cdc.gov/han/han00426.asp>).

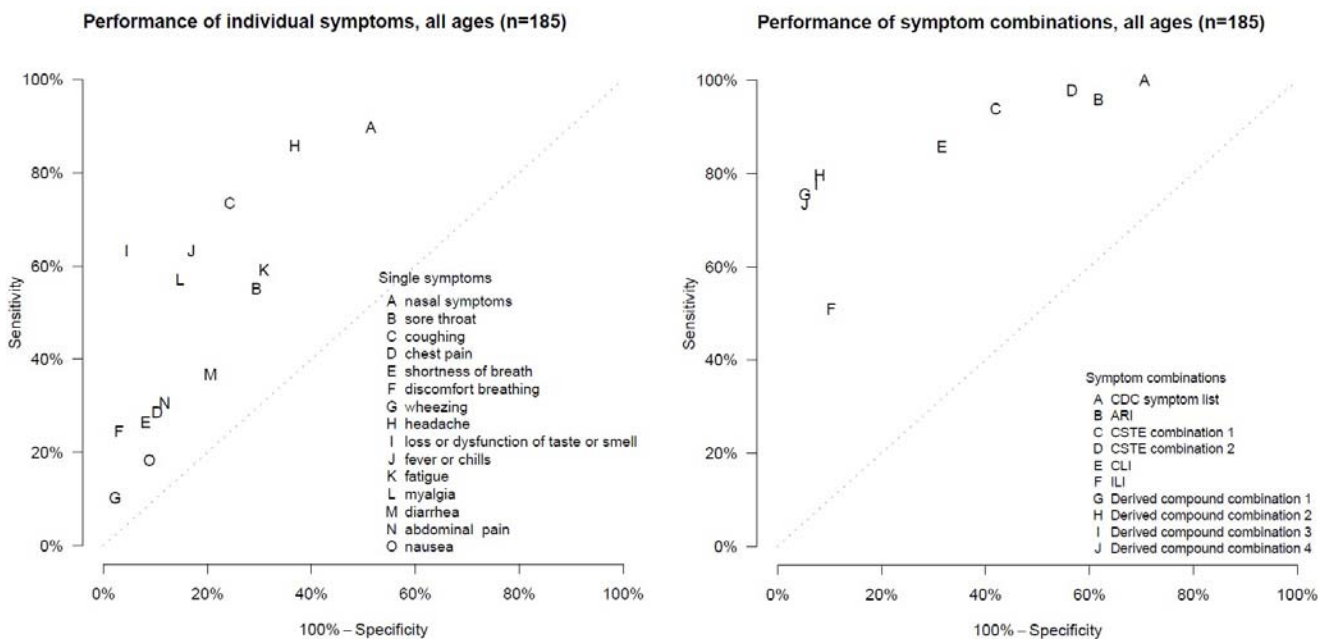
23 ^{††}Influenza-like illness (ILI) outpatient visit information collected through the U.S. Outpatient Influenza-like
24 Illness Surveillance Network (ILINet) (https://www.cdc.gov/flu/weekly/overview.htm#anchor_1539281266932).
25 This collaborative effort between CDC, state and local health departments, and healthcare providers has been
26 tracking patients with ILI since the 1997-1998 influenza season.
27 Of the ten RT-PCR-negative and serology-positive contacts, three (30%) seroconverted during the two-week
28 observation period. Median ages were the same as those who were RT-PCR-positive and serology-positive (22
29 years). The most common symptoms in these ten contacts were cough (n=7), headache (n=6), and fever or chills
30 (n=5). Compared to those who were RT-PCR-positive and serology-positive, the ten individuals with only
31 positive serology had a lower median number of symptoms (2 vs. 7) and longer time interval from the household
32 index case's symptom onset to first diagnostic specimen collection: for the ten participants with discordant
33 results, the median interval was 13 days (interquartile range [IQR] 10–17) whereas for the 185 remaining
34 participants it was 10 days (IQR 7–13). Taken together, these results suggest that mucosal viral shedding ceased
35 in these participants before study enrollment. These 10 enrollees were excluded from the main analyses.
36

37

38

39 Figures

40 **Figure 1. Sensitivity and 100%-specificity for individual COVID-19 symptoms, existing case definitions,**
 41 **and derived compound symptom combinations for a community cohort of 185 individuals with household**
 42 **exposure to COVID-19 in Utah and Wisconsin, United States, March–May 2020**



43
44

45 *Specificity is the probability of testing negative when disease is absent.

46 †Sensitivity is the probability of testing positive when disease is present.

47 ‡U.S. Centers for Disease Control and Prevention (CDC) list of symptoms that may indicate COVID-19 infection
 48 (<https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html>). This symptom list was last
 49 updated on 13 May 2020.

50 §World Health Organization (WHO) acute respiratory infection (ARI) definition for community-based respiratory
 51 syncytial virus (RSV) surveillance (https://www.who.int/influenza/rsv/rsv_case_definition/en/). Last updated 04
 52 February 2020.

53 □U.S. Council of State and Territorial Epidemiologists (CSTE) original clinical criteria for COVID-19 reporting
 54 (https://cdn.ymaws.com/www.cste.org/resource/resmgr/2020ps/interim-20-id-01_covid-19.pdf). This interim
 55 position statement (Interim-20-ID-01) was approved on 05 April 2020 and was replaced by Interim-20-ID-02 on
 56 07 August 2020.

57 ¶U.S. Council of State and Territorial Epidemiologists (CSTE) revised clinical criteria for COVID-19 reporting
 58 (https://cdn.ymaws.com/www.cste.org/resource/resmgr/ps/positionstatement2020/interim-20-id-02_COVID-19.pdf). This interim position statement (Interim-20-ID-02) was approved on 07 August 2020 and replaced
 60 Interim-20-ID-01.

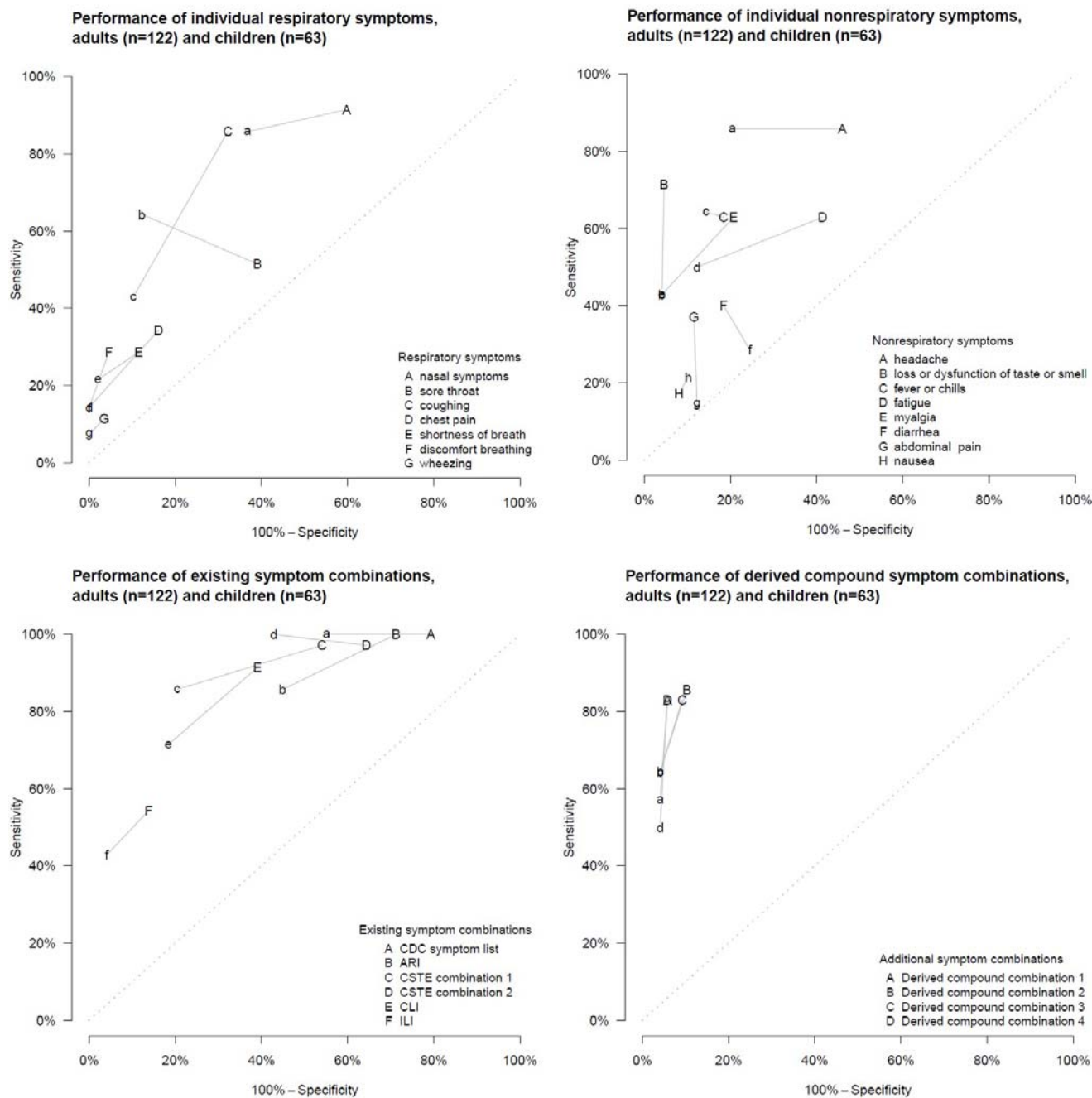
61 **U.S. Centers for Disease Control and Prevention (CDC) COVID-19-like illness (CLI) definition was used to
 62 guide early diagnostic testing strategies from 17 January 2020 – 08 March 2020
 63 (<https://emergency.cdc.gov/han/han00426.asp>).

64 ††Influenza-like illness (ILI) outpatient visit information collected through the U.S. Outpatient Influenza-like
 65 Illness Surveillance Network (ILINet) (https://www.cdc.gov/flu/weekly/overview.htm#anchor_1539281266932).
 66 This collaborative effort between CDC, state and local health departments, and healthcare providers has been
 67 tracking patients with ILI since the 1997-1998 influenza season.

68 ††Combination 1: Taste and/or smell dysfunction, OR one of the following: shortness of breath, myalgia, or fever
 69 or chills.

- 70 ^{§§}Combination 2: Taste and/or smell dysfunction or discomfort breathing, OR at least two of the following:
71 shortness of breath, wheezing, or fever/chills.
72 ^{□□}Combination 3: Taste and/or smell dysfunction, OR at least two of the following: shortness of breath,
73 wheezing, discomfort breathing, or fever/chills.
74 ^{¶¶}Combination 4: Taste and/or smell dysfunction, OR shortness of breath and fever/chills.
75 Points closest to the upper left corner represent those with the highest sensitivity and specificity values.

76 **Figure 2. Sensitivity and 100%-specificity for individual COVID-19 symptoms, existing case definitions,**
 77 **and derived compound symptom combinations for a community cohort of 122 adults (upper case letters)**
 78 **and 63 children (lower case letters) with household exposure to COVID-19 in Utah and Wisconsin, United**
 79 **States, March–May 2020**



80

81 ^{*}Specificity is the probability of testing negative when disease is absent.

82 [†]Sensitivity is the probability of testing positive when disease is present.

83 [‡]U.S. Centers for Disease Control and Prevention (CDC) list of symptoms that may indicate COVID-19 infection
 84 (<https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html>). This symptom list was last
 85 updated on 13 May 2020.

86 [§]World Health Organization (WHO) acute respiratory infection (ARI) definition for community-based respiratory
 87 syncytial virus (RSV) surveillance (https://www.who.int/influenza/rsv/rsv_case_definition/en/). Last updated 04
 88 February 2020.

89 [□]U.S. Council of State and Territorial Epidemiologists (CSTE) clinical criteria for COVID-19 reporting
90 (https://cdn.ymaws.com/www.cste.org/resource/resmgr/2020ps/interim-20-id-01_covid-19.pdf). This interim
91 position statement (Interim-20-ID-01) was approved on 05 April 2020 and was replaced by Interim-20-ID-02 on
92 07 August 2020.

93 [¶]U.S. Council of State and Territorial Epidemiologists (CSTE) clinical criteria for COVID-19 reporting
94 (https://cdn.ymaws.com/www.cste.org/resource/resmgr/ps/positionstatement2020/interim-20-id-02_COVID-19.pdf). This interim position statement (Interim-20-ID-02) was approved on 07 August 2020 and replaced
95 Interim-20-ID-01.

96 ^{**}U.S. Centers for Disease Control and Prevention (CDC) COVID-19-like illness (CLI) definition was used to
97 guide early diagnostic testing strategies from 17 January 2020 – 08 March 2020
98 (<https://emergency.cdc.gov/han/han00426.asp>).

99 ^{††}Influenza-like illness (ILI) outpatient visit information collected through the U.S. Outpatient Influenza-like
100 Illness Surveillance Network (ILINet) (https://www.cdc.gov/flu/weekly/overview.htm#anchor_1539281266932).
101 This collaborative effort between CDC, state and local health departments, and healthcare providers has been
102 tracking patients with ILI since the 1997-1998 influenza season.

103 ^{‡‡}Combination 1: Taste and/or smell dysfunction, OR one of the following: shortness of breath, myalgia, or fever
104 or chills.

105 ^{§§}Combination 2: Taste and/or smell dysfunction or discomfort breathing, OR at least two of the following:
106 shortness of breath, wheezing, or fever/chills.

107 ^{□□}Combination 3: Taste and/or smell dysfunction, OR at least two of the following: shortness of breath,
108 wheezing, discomfort breathing, or fever/chills.

109 ^{¶¶}Combination 4: Taste and/or smell dysfunction, OR shortness of breath and fever/chills.

110 Points closest to the upper left corner represent those with the highest sensitivity and specificity values.
111