Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Characterizing COVID-19: A chief complaint based approach

Rimma Perotte, PhD\textsuperscript{a,c}, Gregory Sugalski, MD\textsuperscript{b}, Joseph P. Underwood, MD\textsuperscript{b}, Michael Ullo, MD\textsuperscript{a,*}

\textsuperscript{a} Department of Emergency Medicine, Hackensack University Medical Center, Hackensack, NJ, United States of America
\textsuperscript{b} Department of Emergency Medicine, Hackensack Meridian School of Medicine, Hackensack, NJ, United States of America
\textsuperscript{c} Department of Biomedical Informatics, Columbia University Irving Medical Center, New York, NY, United States of America


titutututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututututu
imaging [13,14]. Due to widespread concern of infection in the community, an increasing number of patients present to the hospital for testing despite lack of overt symptomatology.

Given the wide array of manifestations and clinical presentations reported thus far, our aim was to evaluate the most commonly reported symptoms for patients presenting to the emergency department with presumed or confirmed COVID-19 infection. To date, there has been no large-scale attempt to classify and categorize the most commonly reported symptoms in emergency department patients. Our emergency department, located seven miles west of New York City in northern New Jersey, continues to be severely impacted by cases of COVID-19. This study will describe the most common presenting complaints of our patients and describe variations in presentations across different patient cohorts.

2. Methods

2.1. Patient flow model

At the onset of the pandemic, our department deployed a split-flow patient segregation model in an effort to streamline flow through the department. Patients arriving to the ED with COVID-19 like illness were sorted by a provider-in-triage based on clinical suspicion for need of hospitalization.

Patients with high suspicion for COVID-19 but low likelihood of hospitalization based on vital signs and comorbidities were directed to our bioccontainment unit (BCU) for evaluation. Due to the scarcity of testing and guidelines from the Center for Disease Control and Prevention (CDC), patients sent to the BCU were not routinely tested for COVID-19. Patients with a high likelihood of requiring admission were placed in the acute care section of the ED with an COVID-19 specific isolation order set including a nasopharyngeal assay. Of note, both patient care streams included patients that had a high pre-test probability of COVID-19 infection.

2.2. Data acquisition and processing

A retrospective analysis of all visits from March 1 to April 24, 2020 in our adult emergency department was initiated from the Epic electronic health record. The chief complaint and demographic data were extracted from Epic’s underlying database. The adult ED treats patients that are twenty-two years of age and older. The full list of ED visits were subset into two cohorts in accordance with IRB approval:

1. COVID-suspected ED visits: defined as visits where patients were triaged into the BCU or placed on a COVID-19 specific isolation order set.
2. COVID-positive ED visits: defined as all ED visits where a patient had a COVID-19 laboratory test ordered between ED arrival and hospital discharge, and the laboratory test came back positive for COVID-19. This is a subset of the COVID-suspected ED visits. Due to the patient flow model, this subset represents COVID-suspected patients who are more likely to be admitted.

For each of these visits, the full list of chief complaints and patient age were extracted from the electronic health record. To more meaningfully characterize the chief complaint list, a group of emergency medicine trained physicians aggregated semantically similar chief complaints. An example of aggregated complaints include “Abrasion” and “Abrasion (Complicated)”. In addition, a group of emergency medicine physicians reached consensus on the categorization of the aggregated chief complaints by organ systems. The organ system level categorization allowed for higher level understanding of presenting complaints.

2.3. Quantitative analyses

In our analysis we describe the prevalence of each chief complaint and its co-occurrence with all other chief complaints. In addition, we performed a clustering analysis on all COVID-19 positive-only patients to identify clusters in the set of chief complaints. Finally, we analyzed age-specific chief complaint differences.

Hierarchical clustering was implemented as a method for identifying commonly co-occurring chief complaints in an entirely data-driven way. Clustering is often applied as a technique to identify similar patients by examining sets of properties shared by these patients. However, our goal was not to computationally group patients, but to computationally group COVID-19 symptoms and patterns of co-occurrence. We therefore chose to perform clustering analysis on the symptoms themselves and use each patient’s full set of recorded chief complaints as input. This implementation of clustering allows for exploration of chief complaint sets that occur in COVID-positive patients.

Differences in chief complaint distributions across different age groups were assessed with Fisher’s exact test. All analyses were completed using Python v3.7.

3. Results

During the study period there were 5015 visits for COVID-suspected patients. 1483 (29%) of those visits were for COVID-positive patients (Table 1). Unsurprisingly, as a consequence of our split flow model, the COVID-positive patients were older with a higher prevalence of hospital admission.

For the study population, 209 unique chief complaints were identified and subsequently culled to 181 unique chief complaints post aggregation. To visualize the descriptive statistics of chief complaint prevalence among the COVID-positive group, a word cloud was generated (Fig. 1).

The word cloud demonstrates the relative prevalence of each of the top 40 individual chief complaints with a relative scaling of 0.3 to allow for legibility. The font size of each complaint is a direct representation of the frequency of the complaint in our dataset. In addition, to provide a graphical representation of the variety of systems appearing in chief complaints, the complaints were color-coded by the primary organ system affected. Two emergency medicine trained physicians categorized each chief complaint independently into an organ system. Disagreements in categorization were adjudicated by a third emergency medicine physician.

For the COVID-positive subset, the list of 113 was reduced to 101 unique chief complaints through the physician-led aggregation of semantically similar complaints (Table 2).

Among the entire study population, “fever” was the most common chief complaint, occurring in 41% of visits followed by “shortness of breath” which occurred in 38%. In the COVID-positive group, this order was reversed: “shortness of breath” occurred 49% of the time with “fever” occurring 45% of the time (Fig. 1). A tabular format of the chief complaint prevalences are included in the appendix.

Table 1

|                      | COVID suspected | COVID positive | COVID negative | COVID non-tested |
|----------------------|-----------------|----------------|----------------|------------------|
| Total number of visits| 5015            | 1483           | 896            | 2636             |
| Number of unique chief complaints | 209            | 113            | 133            | 147              |
| % of 65+ patients    | 26%             | 43%            | 36%            | 13%              |
| % Hospitalization    | 40%             | 78%            | 61%            | 11%              |

Table 1

Descriptive statistics for the COVID patients. The COVID-suspected population is the union of the COVID positive, COVID negative, and COVID non-tested groups.
distinct complaint clusters (Fig. 2). A complaint cluster is a group of complaints that are more likely to occur together than with any other complaint. The clusters below describe the most common chief complaint groupings in our studied encounters.

The identified clusters were:

1. Shortness of Breath, Fever, Cough, Weakness, Chest Pain, Headache
2. Abdominal Pain, Diarrhea, Vomiting, Nausea
3. Altered Mental Status, Fatigue
4. Fall, Back Pain
5. Flu Symptoms, Syncope
6. Generalized Body Aches, Chills, Sore Throat
7. Respiratory Distress

“Respiratory Distress” represented an outlier complaint that was unique in the patients who were affected by it and therefore it did not group with any other complaints.

3.2. Age-based chief complaint differences

COVID-positive patients over the age of 65 were significantly more likely to present with weakness, altered mental status, and diarrhea, while younger patients were significantly more likely to present with fever, cough, headache (Fig. 3).

4. Discussion

Our study reports on over 5000 adult patients with suspected COVID-19 infection presenting to our ED from March 1st to April 24, 2020. Our institution is a tertiary care teaching hospital with an adult ED annual census of over 94,000 patients. It is located seven miles west of New York City representing one of the most impacted regions in the nation for COVID-19 infections [15].

Of the 5015 patients evaluated, a subset analysis of the 1483 confirmed positive patients added to our understanding of the presentations of COVID-19 in the emergency setting. Our analysis confirmed the already understood typical symptoms of COVID-19 infection including fever and shortness of breath [6,16]. However, we were also able to identify additional symptomatology that may prove useful to front-line healthcare providers and public health surveillance systems in combating this pandemic.

Our data shows that patients actually presented with a wide range of complaints spanning a multitude of organ systems. In total, over two hundred chief complaints were recorded in the medical charts of suspected patients, adding complexity to the accurate identification of this disease.

In addition to looking at the encounters of all suspected COVID-19 patients, our study further investigated the subset of patients that ultimately tested positive for the virus. This cohort of patients represents the true positive population of patients that presented to our ED during the study period.

Our findings indicate that the most prevalent chief complaints of studied patients included shortness of breath, cough, and fever. These chief complaints were the most frequently found presentations in both the COVID-19 suspected cohort and those who ultimately tested positive. Despite this, when comparing the two groups, we found several distinct differences. Shortness of breath and weakness were more likely to be present in the confirmed positive group while chest pain was more commonly reported in the larger cohort.

The two groups differed significantly in terms of patient age and hospitalization rate. Our COVID-19 confirmed positive cohort had an almost 80% admission rate and nearly half were over 65 years of age. This cohort also had significantly fewer chief complaints as compared to the larger group. The difference in these two groups may be explained by a combination of our split flow process and the inherent nature of the disease.
When categorizing chief complaints based on organ system involvement, the majority of presentations were respiratory in nature which is not surprising given the widely accepted mode of virus transmission through respiratory droplets [17,18]. Upon further exploration, our hierarchical clustering analysis revealed that patients with confirmed infection separately presented with a myriad of symptoms not related to the respiratory system. Gastrointestinal and neurologic manifestations of illness were present in 11.4% and 8.7% of visits, respectively. These prevalences highlight the importance of maintaining a low threshold for consideration of COVID-19 infection in the absence of classically reported respiratory symptoms.

Respiratory distress represents a unique cluster of visits given the fact that these patients were unable to provide a chief complaint. These patients arrived in extremis and were unable to provide meaningful history regarding symptomatology. This was observed in 1.4% of confirmed positive cases presenting to our ED.

![Fig. 2. Bubble size represents the prevalence of each chief complaint in the set of COVID-positive visits. The raw count of patients with each presenting chief complaint is in parentheses below the complaint. The color and x-axis location represent the cluster that each complaint belongs to.](image)

![Fig. 3. Chief complaint prevalence by age group (65 and older vs. less than 65) of the top 10 complaints. Complaints are sorted in descending order by prevalence in the total population. Statistically significant differences (Fisher's exact test p-value <0.05) are denoted with an asterisk.](image)
Interestingly, in addition to the identification of unique symptom clusters within our patient cohorts, we were able to identify specific differences in the presenting complaints based on patient age. Patients aged 65 and above were significantly more likely to present with altered mental status, weakness, and diarrhea. This distinction is clinically important in the evaluation of geriatric patients because early detection is of utmost importance due age related physiologic changes that lead to increased morbidity and mortality [19]. Solely relying on symptoms such as fever and cough may lead to a delay in diagnosis of COVID-19.

Gastrointestinal complaints represented the second largest significant cluster group in our data set. This is of clinical importance as COVID-19 is largely viewed as a respiratory illness. Interestingly, the first known case of COVID-19 in the United States (Snohomish County, Washington) presented with a two-day course of nausea and vomiting. Evidence of viral nucleic acids were detected first in stool samples and then subsequently in respiratory specimens [20]. Front-line providers in the emergency department must remain hypervigilant in order to quickly identify and isolate potential COVID-19 patients based on complaints not classically associated with the virus. Our cluster analysis helps to highlight the significant occurrence of alternate complaint clusters outside of the frequently associated respiratory cluster.

4.1. Limitations

Our study has several limitations. Due to lack of widely available testing we were unable to test each patient presenting to the ED for COVID-19. As a result, we relied on the clinical gestalt of various providers to identify cases of suspected COVID-19 who were not ultimately tested for infection. The patient flow model deployed by our department created a selection bias for testing as we attempted to streamline patient throughput. As a result, patients who were more acutely ill received confirmatory testing and thus our data is skewed towards the presenting complaints of this patient population.

We acknowledge the statistical limitation that although the Fisher exact test assumes independent observations, our unit of measurement is visits and therefore we do have some instances of patients appearing in the dataset multiple times. However, this is a rare occurrence as we are only including visits to the ED when a positive COVID test was identified. We do not expect this limitation to change the findings.

We chose to explore the chief complaint data through hierarchical clustering as such a technique has no need for pre-specification of the size or number of clusters. However, we note that the interpretation of these results can be subjective as to where the clusters are defined, and which similarity metric is used. By determining the clusters solely based on a standard similarity metric, we limited the subjectivity of the analysis.

Additionally, the patients themselves may have been biased in their presenting chief complaints, choosing to describe symptoms commonly reported in various media outlets and downplaying other symptoms that they may have been experiencing. Furthermore, our electronic health record comes with pre-defined chief complaints that are not free text in nature. As a result, triage nurses may have miscategorized some chief complaints in order to make them fit into one of the predefined options. Of note, anosmia, which has been widely reported, is not found in our study, which speaks to this specific limitation with our electronic health record and within our dataset as anosmia is not in the predefined chief complaint list that our institution has.

5. Conclusions

Our quantitative analysis of COVID-19 presenting complaints gives providers and public health officials new information for the identification of potentially affected patients in the emergency department. The diversity of presentations suggests that the proportion of “non-classic” presentations is not insignificant. Therefore, casting a wide net and maintaining a high index of suspicion in atypical cases is of paramount importance to correctly make the diagnosis and prevent community and nosocomial spread. Future research on this topic should include determination of patient outcomes based on symptom presentation, in addition to further investigation into patient presentations in other health care settings such as primary care practices and surgical suites.

Presentations

This project has not been formally submitted nor presented at a scientific meeting.

Financial support

This project did not receive financial support.

Author contributions

MU developed the idea and concept of the study. JU, GS, MU, RP developed the methodology for data acquisition. All contributed equally to data interpretation as well as the drafting and revisions of the manuscript. RP provided data acquisition from the electronic health record and conducted statistical analyses.

Declaration of Competing Interest

The authors of this manuscript have no relevant conflicts of interest to report.

Appendix A. Appendix

Table A1

| Chief complaint                   | Prevalence | Organ system |
|----------------------------------|------------|--------------|
| Shortness of breath              | 49.0%      | Respiratory  |
| Fever                            | 44.7%      | Constitutional |
| Cough                            | 32.0%      | Respiratory  |
| Weakness                         | 7.8%       | Constitutional |
| Chest Pain                       | 4.4%       | Cardiovascular |
| Altered mental status            | 3.7%       | Neurologic   |
| Abdominal pain                   | 3.5%       | Gastrointestinal |
| Headache                         | 3.4%       | Neurologic   |
| Diarrhea                         | 3.1%       | Gastrointestinal |
| Vomiting                         | 2.5%       | Gastrointestinal |
| Flu symptoms                     | 2.0%       | Constitutional |
| Generalized body aches           | 1.9%       | Constitutional |
| Fall                             | 1.8%       | Trauma       |
| Nausea                           | 1.5%       | Gastrointestinal |
| Respiratory distress             | 1.4%       | Respiratory  |
| Chills                           | 1.1%       | Constitutional |
| Syncope                          | 1.0%       | Cardiovascular |
| Fatigue                          | 0.9%       | Constitutional |
| Back pain                        | 0.9%       | Musculoskeletal |
| Sore throat                      | 0.8%       | HEMT         |
| Hypotension                      | 0.7%       | Cardiovascular |
| Dizziness                        | 0.7%       | Neurologic   |
| Abnormal lab results             | 0.6%       | Other        |
| Loss of appetite                 | 0.5%       | Constitutional |
| High blood sugar                 | 0.5%       | Endocrine    |
| Dehydration                      | 0.5%       | Constitutional |
| Lethargy                         | 0.5%       | Constitutional |
| Leg Pain                         | 0.4%       | Musculoskeletal |
| Medical clearance                | 0.4%       | Other        |
| Palpitations                     | 0.3%       | Cardiovascular |
| Poor nutritional intake          | 0.3%       | Constitutional |
| Cancer                           | 0.3%       | Hematologic  |
| Cardiac arrest                   | 0.3%       | Cardiovascular |
Table A1 (continued)

| Chief complaint       | Prevalence | Organ system |
|-----------------------|------------|--------------|
| Near syncope          | 0.3%       | Cardiovascular |
| Rectal bleed          | 0.3%       | Gastrointestinal |
| Congestion            | 0.3%       | HEENT |
| Leg swelling           | 0.3%       | Musculoskeletal |
| Closed head injury    | 0.3%       | Trauma |
| Loss of consciousness | 0.2%       | Neurologic |
| Seizure               | 0.2%       | Neurologic |

References

[1] WHO Director-General’s Remarks at the Media Briefing on 2019-nCoV on 11 February 2020. Accessed Apr 26, 2020 https://www.who.int/dg/speeches/detail/who-director-general-s-remarks-at-the-media-briefing-on-2019-ncov-on-11-february-2020?.

[2] Geographic differences in COVID-19 cases, deaths, and incidence — united states, February 12-April 7, 2020, MMWR Morb Mortal Wkly Rep. 2020;69.

[3] Coronavirus Disease 2019 (COVID-19) – Transmission. Centers for Disease Control and Prevention; 2020 Accessed Apr 26, 2020.

[4] Wang D, Hu B, Hu C, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel Coronavirus–Infected pneumonia in Wuhan, China. JAMA. 2020;323:1061–9.

[5] Goyal P, Choi JJ, Pinheiro LC, et al. Clinical characteristics of covid-19 in New York City. N Engl J Med. 2020;382:2372–4.

[6] Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395:497–506.

[7] Giacomelli A, Pezzati L, Conti F, et al. Self-reported olfactory and taste disorders in patients with severe acute respiratory coronavirus 2 infection: a cross-sectional study. Clin Infect Dis. 1 August 2020;71(15):889–90.

[8] Lechien JR, Chiesa-Estomba C, De Siati DR, et al. Olfactory and gustatory dysfunctions as a clinical presentation of mild-to-moderate forms of the coronavirus disease (COVID-19): a multicenter european study. Eur Arch Otorhinolaryngol. 2020 (EPublished ahead of print).

[9] Jin X, Lian J, Hu J, et al. Epidemiological, clinical and virological characteristics of 74 cases of coronavirus-infected disease 2019 (COVID-19) with gastrointestinal symp- toms. Gut. 2020;69(6):1002–9.

[10] Recalcati S. Cutaneous manifestations in COVID-19: a first perspective. J Eur Acad Dermatol Venereol. May 2020;34(5):e212–3.

[11] Mao L, Jia H, Wang M, et al. Neurologic manifestations of hospitalized patients with coronavirus disease 2019 in Wuhan, China. JAMA Neurol. 2020;77(6):683–90.

[12] Inciardi RM, Lapi L, Zacone G, et al. Cardiac involvement in a patient with corona-virus disease 2019 (COVID-19). JAMA Cardiol. 2020;5(7):819–24.

[13] Mizumoto K, Kagaya K, Zarebski A, et al. Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the diamond princess cruise ship, Yokohama, Japan, 2020. Eurosurveillance. 2020;25:2000180.

[14] Wang Y, Liu Y, Liu L, et al. Clinical outcomes in 55 patients with severe acute respiratory syndrome coronavirus 2 who were asymptomatic at hospital admission in Shenzhen, China. J Infect Dis. 1 June 2020;221(11):1770–4.

[15] COVID-19 United States Cases by County, Johns Hopkins Coronavirus Resource Center; 2020 Accessed Apr 30, 2020.

[16] Shen 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;395:507–13.

[17] Liu J, Liao X, Qian S, et al. Community transmission of severe acute respiratory syndrome coronavirus 2, Shenzhen, China, 2020. Emerg Infect Dis. 2020;26.https://doi.org/10.3201/eid2606.200239.

[18] Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel Coronavirus–Infected pneumonia. N Engl J Med. 2020;382:1199–207.

[19] Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the chinese center for disease control and prevention. JAMA. 2020;323:1239–42.

[20] Holsue HL, DeBolt C, Lindquist S, et al. First case of 2019 novel coronavirus in the United States. N Engl J Med. 2020;382:929–36.