|  |  |  |
| --- | --- | --- |
| **eMERGE Network: Manuscript Concept Sheet** | | |
| **Reference Number**  *(to be assigned by CC)* | NT439 | |
| **Submission Date** | 12/16/2021 | |
| **Project Title** | Defining Acute SARS-CoV-2 Illness-Related Hospitalization in Electronic Health Records | |
| **Tentative Lead Investigator** *(first author)* | Gurmukteshwar Singh1 | |
| **Tentative Lead Investigator Email Address** | gsingh3@geisinger.edu | |
| **Tentative Senior Author**  *(last author)* | Marc Williams1, Laura J. Rasmussen-Torvik3 | |
| **All Other Authors** | Sarah B. Delozier2, Lubna Alhalabi4, Terrie Kitchner5, Shawn Murphy6, Benjamin Satterfield4, Mark Wurfel7, Gail Jarvik7, Pavan Hatraju7, David Crosslin7, Krystof Kiryluk8, Ning Shang8, Iftikhar Kullo4 | |
| **Sites Participating** | 1. Geisinger Health: 2 investigators 2. Vanderbilt University Medical Center: 1 investigator 3. Northwestern University: 1 investigator 4. Mayo Clinic: 3 investigators 5. Marshfield Clinic: 1 investigator 6. Massachusetts General Hospital: 1 investigator 7. University of Washington: 4 investigators 8. Columbia University: 2 investigators | |
| **Background / Significance** | Severe acute respiratory syndrome coronavirus 2 (SARS-COV-2), has been responsible for an ongoing global pandemic of coronavirus disease 2019 (COVID-19), resulting in significant morbidity and mortality 1. Recognizing the importance of population-based surveillance early, the COVID-19-Associated Hospitalization Surveillance Network (COVID-NET) was created for ongoing population-based surveillance, using the existing infrastructure of the Influenza Hospitalization Surveillance Network 2. Under this program, a SARS-COV-2 related hospitalization was defined as one occurring within 14 days of a positive SARS-COV-2 test result 3. Effective April 1, the Centers for Disease Control and Prevention’s National Center for Health Statistics implemented a new International Classification of Diseases, Tenth Revision (ICD-10) emergency code (U07.1) 4. Soon thereafter, early data from a single center registry was used to evaluate phenotyping algorithms to define COVID-19 cases. An algorithm containing the ICD-10 code or positive laboratory testing was found to have the best performance on recall 5.  With further experience with COVID-19 patients, our understanding of the disease spectrum has evolved. Time-sensitive phases in the disease course have been recognized, namely acute COVID-19, subacute or ongoing COVID-19 and chronic or post-COVID-19 6, 7. To allow precise identification of these phases in hospitalized patients, it is critical to have a validated protocol to accurately identify the sequence of events, starting with the initial acute hospitalization. However, studies describing hospitalization outcomes have used heterogenous definitions of COVID-19 related hospitalization. These include various approaches including, but not limited to the COVID-NET or institution-specific time-based definitions, ICD-10 code U07.1, use of laboratory test results and use of physician assessment questionnaires 3, 8-11. Up to 1 in 5 survivors of a COVID-19 hospitalization are readmitted within 60 days. Only 30% of these readmissions had a discharge diagnosis of COVID-19 9. Hence, an ideal database definition of acute SARS-CoV-2 hospitalization should have a high positive as well as negative predictive value. We will attempt to test addition of processes of care and COVID-19 treatments to laboratory test and ICD-10 code-based algorithms to identify the best performing definition. We will validate this by manually reviewing charts across 6 different health systems with heterogenous geographic and population distribution.  REFERENCES  1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis. 2020;20(5):533-534.  2. Ss C, R L, Ml L, J B, L F. The US Influenza Hospitalization Surveillance Network. Emerging infectious diseases. 2015;21(9).  3. Garg S, Kim L, Whitaker M, et al. Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019 - COVID-NET, 14 States, March 1-30, 2020. MMWR Morb Mortal Wkly Rep. 2020;69(15):458-464.  4. Centers for Disease Control and Prevention. New ICD-10-CM code for the 2019 Novel Coronavirus (COVID-19), April 1, 2020. https://www.cdc.gov/nchs/data/icd/Announcement-New-ICD-code-for-coronavirus-3-18-2020.pdf.  5. DeLozier S, Bland S, McPheeters M, et al. Phenotyping coronavirus disease 2019 during a global health pandemic: Lessons learned from the characterization of an early cohort. J Biomed Inform. 2021;117:103777.  6. Datta SD, Talwar A, Lee JT. A Proposed Framework and Timeline of the Spectrum of Disease Due to SARS-CoV-2 Infection: Illness Beyond Acute Infection and Public Health Implications. JAMA. 2020;324(22):2251-2252.  7. Nalbandian A, Sehgal K, Gupta A, et al. Post-acute COVID-19 syndrome. Nat Med. 2021;27(4):601-615.  8. V C, Sa F, M O, An M, Hc P. Sixty-Day Outcomes Among Patients Hospitalized With COVID-19. Annals of internal medicine. 2021;174(4).  9. Donnelly JP, Wang XQ, Iwashyna TJ, Prescott HC. Readmission and Death After Initial Hospital Discharge Among Patients With COVID-19 in a Large Multihospital System. JAMA. 2021;325(3):304-306.  10. Nguyen NT, Chinn J, Nahmias J, et al. Outcomes and Mortality Among Adults Hospitalized With COVID-19 at US Medical Centers. JAMA Netw Open. 2021;4(3):e210417.  11. Jd B, A B, N B, R B, J S. The Impact Of The COVID-19 Pandemic On Hospital Admissions In The United States. Health affairs (Project Hope). 2020;39(11). | |
| **Outline of Project** | **Proposed Methodology:**  Brief overview: Start with all authors doing a 100-chart reviews each to create a database of acute COVID vs non COVID vs subacute/reinfection/ chronic COVID hospitalizations. Then apply the definition algorithms from automated EHR data pulls to see which definitions perform best.  Step by step instructions:   1. **Data pull** 2. Search criterion: Patients with inpatient hospitalization April 1, 2020 through April 30, 2021 3. Run query: What % patient with U07.1 do not have a positive PCR during or within 14 days of admission? 4. Run query: What % of patients with a positive PCR within 14 days or during admission have a U07.1 code? 5. Interesting point to note: % differences between different centers and maybe month-over month 6. Data extraction: Pull 60 random patients with ICD-10 code U07.1 on active hospital problem list or billing for the hospitalization (possible true or false positives) 7. Data extraction: Pull 40 random patients with SARS-CoV-2 PCR positivity during or before admission but no U07.1 code (Possible true negatives, false negatives or true positives) 8. **Case definitions for chart review (look for overall impression, not all criteria need to be fulfilled):** 9. Acute COVID admission:  * 1st COVID hospitalization AND * Admitted within 2 weeks of first positive PCR/ 3 weeks of symptom onset with positive PCR or antigen testing documented in the clinical notes * If unknown duration from PCR/ symptoms, discharge summary or clinical notes attribute hospitalization directly to acute COVID-19 infection  1. Post-acute COVID admission or long/ chronic COVID:  * Prior COVID admission/ mention of prior COVID illness AND * > 2 weeks from positive PCR/ 3 weeks of symptom onset OR * Discharge summary or clinical notes mention this admission due to sequelae of prior COVID-19 infection  1. COVID reinfection  * Discharge summary/ post work-up clinical notes describe COVID reinfection or symptoms attributed to it OR * Prior COVID infection: either admission or mentioned in notes AND/OR * PCR positive again >= 90 days after prior positive PCR/ admission  1. Non COVID admission:  * No positive COVID PCR OR * PCR positive > 4 weeks ago and clinical notes identify non-COVID cause of hospitalization OR * Clinical notes identify non-COVID cause of hospitalization  1. **Chart Review:**   **Excel Spreadsheet with the following columns; use =DATE function for dates**   1. Patient identifier 2. Direct admission, transfer or unsure (0,1,2) 3. Date of admission (to outside hospital if transfer) 4. Date of symptom onset (if mentioned) 5. COVID testing  * Any PCR (+/-) present in system during/ before admission (1,0): Used any PCR to pick up old COVID, re-infection etc. This will help us determine the performance of a definition in recovered COVID patients as well as acute patients. We can use the date range to determine later if it was close to the admission or not * Were any PCR results positive? (1,0) * Date of first positive PCR * Mention of PCR/antigen test outside system in notes (1,0) * Date of positive antigen testing * Unclear/unsure about COVID testing (1,0)  1. Your impression  * Acute COVID (1,0) * Post-acute COVID (1,0) * COVID reinfection (1,0) * Post-acute COVID or COVID reinfection (hard to tell) (1,0) * Not COVID related (1,0) * Unsure (1,0)  1. **Explore data fidelity and missingness** 2. Evaluate for missing data/ alphanumeric data instead of 0,1, or dates 3. % of direct admissions and transfers 4. Are there differences in dates of admission by chart review and data pull (for direct admissions)? 5. % of transferred patients with outside admission dates available vs not 6. ? Difference between dates of admission and database extraction in transferred patients 7. Difference in % or dates of COVID testing in chart review and data pull? 8. % with COVID testing, PCR testing, antigen testing. Changes with direct admissions vs transfers? 9. % unsure responses in each impression 10. Validate the first 3 columns of the definitions to see if we can get excellent performance without complicating or adding processes of care. If that’s the case, stop. 11. Run query: Are airborne or COVID-19 precautions identifiable in the database? Can the duration of airborne precautions reliably be extracted? What % people with U07.1 had airborne precautions? What % of patients with a positive PCR within 14 days or during admission had airborne precautions? 12. % of patients with paO2, paCO2, LDH, procalcitonin 13. **Validate definitions**   Candidate definition rules:   1. Positive PCR during admission or within 2 weeks before admission 2. Positive PCR (ever) before/ during admission 3. U07.1 ICD code in hospital problem list, billing or discharge diagnosis 4. COVID airborne precautions +/- > 72 hours (TBD) 5. Concurrent labs: PaCO2, PaO2, LDH, Procalcitonin 6. Therapeutics use during admission: Remdesivir or hydroxychloroquine or baricitinib or convalescent plasma or steroids (dexamethasone, prednisone, methylprednisione) or ritonavir or lopinavir or tocilizumab or bamlanivimab or bamlanivimab/etesivimab or casirivimab/imdevimab or ivermectin or anakinra   Definition of acute case on chart review: Evidence of active infection (symptomatic/ asymptomatic) or within 2 weeks of positive PCR   |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | | To be tested | 1 | 1 | 2 | 2 | 2 | 2 | 2 | | *Qualifier* |  | OR | OR | OR | OR | OR | OR | | *To be tested* |  | 3 | 3 | 3 | 3 | 3 | 3 | | *Qualifier* |  |  |  | AND | AND | AND | AND | | *To be tested* |  |  |  | 4/5/6 | 4 | 5 | 6 | | *Definition* | 1 | 1/3 | 2/3 | 2/3+4/5/6 | 2/3+4 | 2/3+5 | 2/3+6 | | |
| **Desired Data - Common Variables\***  *(Available from the CC)* | Demographics  ICD9/10 codes  CPT codes  Phecodes  BMI | |
| **Other Desired Data *(Available from participating sites)*** | *Please specifically list out any data elements that participating sites would collect or extract from clinical or other sources for this project (i.e. not common variables above)* | Common Variable Labs  Common Variable Meds  ☐ Geocoding 2015 ACS variables  Other: Case/Control status |
| **Desired Genetic Data** | eMERGE I-III Merged set (HRC imputed, GWAS)  eMERGE PGx/PGRNseq data set  eMERGEseq data set (Phase III)  eMERGE Whole Genome sequencing data set  eMERGE Exome chip data set  eMERGE Whole Exome sequencing data set  Other (not listed above): | |
| **Does project pertain to an existing eMERGE Phenotype?** | Yes, if so please list  No | |
| **Planned Statistical Analyses** |  | |
| **Ethical Considerations** |  | |
| **Available Funding or Resources** |  | |
| **Target Journal** | JAMA/ JAMA Network Open | |
| **Milestones**  *(This section should include the key dates for completion of project, including approval, project duration, draft completion, and submission.)* | June 15, 2021: Determine site specific chart review plan/ initial data pull  June 20,2021: Discuss issues with initial chart review  July 15, 2021: Deadline to submit chart reviews  August 1, 2021: Data analysis  August 15, 2021: Develop consensus definition to use in phenotypes, temporal description, etc | |
|  |  | |

**\*Common Variables available across all datasets:**

* Demographics: sex, year of birth, decade of birth, race, ethnicity
* Codes: (repeated values & age at event): ICD, CPT, Phecodes
* BMI: (repeated value & age at event) height, weight, BMI
* Labs: (lab name, repeated lab value & age at event) Serum total cholesterol, LDL, HDL, Triglycerides, Glucose fasting/non-fasting/unknown, & White Blood Cell count
* Medications: (medication name, repeated, & age at event) Cerivastatin sodium, Rosuvastatin, Simvastatin, Fluvastatin, Pravastatin, Lovastatin, Atorvastatin, & Pitavastatin
* Other: Case/Control status on Phase I and Phase II phenotype: only on GWAS dataset participants