EDIITORS’ PICK

Christopher R. Carpenter, Philip A. Mudd, Colin P. West et al. 654

ORIGINAL CONTRIBUTIONS

Myocardial Infarction Can Be Safely Excluded by High-sensitivity Troponin I Testing 3 Hours After Emergency Department Presentation
W. Frank Peacock, Robert Christenson, Deborah B. Diercks et al. 671

Novel Use of Home Pulse Oximetry Monitoring in COVID-19 Patients Discharged From the Emergency Department Identifies Need for Hospitalization
Sonia Shah, Kaushal Majmudar, Amy Stein et al. 681

E. Brooke Lerner, Craig D. Newgard, N. Clay Mann et al. 693

Academic Emergency Medicine Physicians’ Anxiety Levels, Stressors, and Potential Stress Mitigation Measures During the Acceleration Phase of the COVID-19 Pandemic
Robert M. Rodriguez, Anthony J. Medak, Brigitte M. Baumann et al. 700

Relationship Between Pain, Opioid Treatment, and Delirium in Older Emergency Department Patients
Raoul Daoust, Jean Paquet, Valérie Boucher et al. 708

Randomized Clinical Trial of Intravenous (IV) Acetaminophen as an Adjunct to IV Hydromorphone for Acute Severe Pain in Emergency Department Patients
Polly E. Bijur, Benjamin W. Friedman, Deborah White et al. 717

Depression and Functional Outcomes in Patients Presenting to the Emergency Department With Low Back Pain
Joshua Ring, Sarah Peskoe, Congwen Zhao et al. 725

Feasibility Study of a Quasi-experimental Regional Opioid Safety Prescribing Program in Veterans Health Administration Emergency Departments
Nathalie Dieujuste, Rachel Johnson-Koenke, Melissa Christopher et al. 734

Contents continued inside.
SYSTEMATIC REVIEWS (WITH OR WITHOUT META-ANALYSES)
The Characteristics and Effectiveness of Interventions for Frequent Emergency Department Utilizing Patients With Chronic Noncancer Pain: A Systematic Review
Charles K. Wong, Connor M. O’Rielly, Braden D. Teitge et al. 742

ORIGINAL CONTRIBUTION
Interrupted Time Series of User-centered Clinical Decision Support Implementation for Emergency Department-initiated Buprenorphine for Opioid Use Disorder
Wesley C. Holland, Bidisha Nath, Fangyong Li et al. 753

RESEARCH LETTERS
Diagnostic Performance of a Rapid Point-of-care Test for SARS-CoV-2 in an Urban Emergency Department Setting
Samuel McDonald, D. Mark Courtney, Andrew E. Clark et al. 764
The Fine Balance: Adapting Clinical Research Into COVID-19 Response
Nicholas J. Giordano, Naomi Alanis, Jennifer A. Frey et al. 767
Testing Asymptomatic Emergency Department Patients for Coronavirus Disease 2019 (COVID-19) in a Low-prevalence Region
James S. Ford, Aman Parikh, Rupinder Sandhu et al. 771

HOT OFF THE PRESS
Hot Off the Press: Video for Acute Otitis Media Discharge Instructions
Christopher Bond, Justin Morgenstern, Corey Heitz et al. 775

THE BRASS TACKS: CONCISE REVIEWS OF PUBLISHED EVIDENCE
Anticonvulsants for the Treatment of Low Back Pain and Lumbar Radicular Pain
Michael Gottlieb, Alex Koyfman, Brit Long et al. 779
Ketamine Versus Opioids for Acute Pain in the Emergency Department
Michael J. Duhaime, Allan B. Wolfson 781

COMMENTARY - INVITED
Prognosis, Prognosis, Wherefore Art Thou Prognosis?
Phillip D. Levy 783
Pulse Oximetry as a Biomarker for Early Identification and Hospitalization of COVID-19 Pneumonia
Richard M. Levitan 785

COMMENTARY - UNSOLICITED
Guidance and Patient Instructions for Proning and Repositioning of Awake, Nonintubated COVID-19 Patients
Suzanne K. Bentley, Laura Iavicoli, David Cherkas et al. 787

CORRESPONDENCE – UNSOLICITED LETTERS TO THE EDITOR
In Response to “The Novel Use of Home Pulse Oximetry”: An Australian Offer of Support
Martin Dutch, Jonathan Knott 792
Intravenous Acetaminophen and Hydromorphone: The Bad and the Ugly of Emergency Department Pain Management
Maryann Mazer-Amirshahi, Sergey Motov, Lewis S. Nelson et al. 793
Response to Mazer-Amirshahi et al, Intravenous Acetaminophen and Hydromorphone: The Bad and the Ugly of Emergency Department Pain Management
Polly E. Bijur, Benjamin W. Friedman, Andrew K. Chang et al. 795

REFLECTIONS
Prehospital Care at the Epicenter of a Pandemic: The New York City EMS Response
Matt S. Friedman, Reuben J. Strayer 797
My Family Didn’t Sign Up for This
Liza Hartofilis 802
Nassim: The Boy Who Never Came Home and Other Reflections From War
CPT Anant Shukla 804

CME Editor: Corey Heitz, MD

Authors: Christopher R. Carpenter, MD, MSc, Philip A. Mudd, MD, PhD, Colin P. West, MD, PhD, Erin Wilber, and Scott T. Wilber, MD, MPH

If you wish to receive credit for this activity, please refer to the website: www.wileyhealthlearning.com/aem

Educational Objectives
After reading the article, participants should be able to discuss diagnostic accuracy of clinical findings and testing strategies for COVID-19.

Activity Disclosures
This activity received no commercial support.

CME Editor Corey Heitz discloses no relevant financial relationships.

This activity underwent peer review in line with standards of editorial integrity and publication ethics. Conflicts of interest have been identified and resolved in accordance with John Wiley and Sons, Inc.’s Policy on Activity Disclosure and Conflict of Interest.

Accreditation
John Wiley and Sons, Inc. designates this journal-based CME activity for a maximum of 1.0 AMA PRA Category 1 Credit™. Physicians should only claim credit commensurate with the extent of their participation in the activity.

For information on applicability and acceptance of continuing medical education credit for this activity, please consult your professional licensing board.

This activity is designed to be completed within 1 hour. To successfully earn credit, participants must complete the activity during the valid credit period, which is up to two years from initial publication. Additionally, up to 3 attempts and a score of 70% or better is needed to pass the post test.

Christopher R. Carpenter, MD, MSc, Philip A. Mudd, MD, PhD, Colin P. West, MD, PhD, Erin Wilber, and Scott T. Wilber, MD, MPH

ABSTRACT

Objective: Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged as a global pandemic in early 2020 with rapidly evolving approaches to diagnosing the clinical illness called coronavirus disease (COVID-19). The primary objective of this scoping review is to synthesize current research of the diagnostic accuracy of history, physical examination, routine laboratory tests, real-time reverse transcription–polymerase chain reaction (rRT-PCR), immunology tests, and computed tomography (CT) for the emergency department (ED) diagnosis of COVID-19. Secondary objectives included a synopsis of diagnostic biases likely with current COVID-19 research as well as corresponding implications of false-negative and false-positive results for clinicians and investigators.

Methods: A Preferred Reporting Items for Systematic Reviews and Meta-Analyses–Scoping Review (PRISMA-ScR)–adherent synthesis of COVID-19 diagnostic accuracy through May 5, 2020, was conducted. The search strategy was designed by a medical librarian and included studies indexed by PubMed and Embase since January 2020.

Results: A total of 1,907 citations were screened for relevance. Patients without COVID-19 are rarely reported, so specificity and likelihood ratios were generally unavailable. Fever is the most common finding, while hyposmia and hypogeusia appear useful to rule in COVID-19. Cough is not consistently present. Lymphopenia is the most commonly reported laboratory abnormality and occurs in over 50% of COVID-19 patients. rRT-PCR is currently considered the COVID-19 criterion standard for most diagnostic studies, but a single test sensitivity ranges from 60% to 78%. Multiple reasons for false-negatives rRT-PCR exist, including sample site tested and disease stage during which sample was obtained. CT may increase COVID-19 sensitivity in conjunction with rRT-PCR, but guidelines for imaging patients most likely to benefit are emerging. IgM and IgG serology levels are undetectable in the first week of COVID-19, but sensitivity (range = 82% to 100%) and specificity (range = 87% to 100%) are promising. Whether detectable COVID-19 antibodies correspond to immunity remains unanswered. Current studies do not adhere to accepted diagnostic accuracy reporting standards and likely report significantly biased results if the same tests were to be applied to general ED populations with suspected COVID-19.

Conclusions: With the exception of fever and disorders of smell/taste, history and physical examination findings are unhelpful to distinguish COVID-19 from other infectious conditions that mimic SARS-CoV-2 like influenza. Routine laboratory tests are also nondiagnostic, although lymphopenia is a common finding and other...
abnormalities may predict severe disease. Although rRT-PCR is the current criterion standard, more inclusive consensus-based criteria will likely emerge because of the high false-negative rate of PCR tests. The role of serology and CT in ED assessments remains undefined.

In December 2019 a novel viral respiratory pathogen emerged in China, ultimately named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) with the clinical illness dubbed coronavirus disease (COVID-19). COVID-19 became a global pandemic in early 2020 forcing governments worldwide to enact social isolation policies with dire economic ramifications. Emergency departments (ED) encountered decreased patient volumes before some in Seattle, New York City, New Orleans, and Detroit experienced waves of COVID-19 patients mixed with asymptomatic patients or those concerned about potential exposures. Diagnosing COVID-19 was hampered by inadequate supplies of reagents and kits, which was compounded by clinical and radiographic features that overlap with numerous seasonal viral respiratory infections.1

The U.S. Food and Drug Administration (FDA) issued an emergency use authorization (EUA) on February 4, 2020, to enable Centers for Disease Control (CDC)–qualified laboratories to perform COVID-19 testing. As of June 3, the FDA provided EUA for 85 commercial assays, including polymerase chain reaction (PCR) tests and immunoglobulin assays.2 Early real-time reverse transcription PCR (rRT-PCR) tests had false-negative (1 – sensitivity) rates as high as 40%.3 As waves of COVID-19 patients present to EDs in the coming months with symptoms or potential exposures, understanding the diagnostic accuracy and reliability of history, physical examination, routine laboratory tests, advanced imaging, and an evolving array of COVID-19 diagnostics will be essential knowledge to inform the timing of testing, optimal specimen and test selection, shared decision making, and ultimately derivation of clinical instruments to guide disposition, follow-up, and shared decision-making choices (Figure 1).4 This review provides a narrative overview of published research with the primary objective to describe the frequency, causes, and implications of false-negative rRT-PCR for diagnosis and surveillance. Secondary objectives include describing potential diagnostic biases in current rapid-cycle COVID-19 diagnostic research reports, while providing recommendations for clinicians for interpreting results with knowledge of these design and reporting limitations. A final objective is to add context to rRT-PCR ordering and interpretation by understanding the diagnostic accuracy and additive value of history, physical examination, routine hematology and chemistry tests, computed tomography (CT), and serology for COVID-19 immunoglobulins.

METHODS

Search Strategy

This is a scoping review that adheres to PRISMA-ScR reporting recommendations.5 The published literature was searched using strategies created by a medical librarian for COVID-19 and diagnostic accuracy. The search was implemented in PubMed 1946– and Embase 1947– with a date limit from January 1, 2020, until present with an English language limit. The search strategy used a combination of standardized terms and keywords, including but not limited to (Covid-19 or Novel Coronavirus or SARS-COV-2) and (diagnosis or PCR or serology or CRISPR-CAS or sensitivity/specificity; Data Supplement S1, available as supporting information in the online version of this paper, which is available at http://onlinelibrary.wiley.com/doi/10.1111/acem.14048/full). The testing search was based on Cheng et al.,6 adding to this prior publication by incorporating clinical examination, imaging, and serology into the synthesis of current diagnostic research. The searches were performed on April 23 and May 5, 2020.

Study Selection

One author (CRC) reviewed the title and abstracts for all identified citations. Other authors reviewed the manuscripts identified and added pertinent references. Original research studies describing the frequency of history/physical examination findings or diagnostic accuracy (sensitivity, specificity, likelihood ratios) of history/physical examination, laboratory tests, or imaging for COVID-19 were included. Exclusion criteria included non-English, animal research, study protocols, prevention, pathophysiology, laboratory processing, or policy manuscripts. Two authors (CRC, SW) abstracted diagnostic accuracy data and reported adherence to the Standards for Reporting of Diagnostic Accuracy (STARD) guidelines.7 Compliance with STARD was used as a measure of research
quality. The same two authors synthesized the results into summary conclusions.

RESULTS

A total of 1,907 citations were screened (Figure 2). None of the studies cite or adhere completely to either the STARD or the updated reporting framework for history and physical examination. Many of these early publications are letters or case reports with uncertain editorial rigor judging by the turnaround time from initial submission to publication. Many studies rely on rRT-PCR as the criterion standard for COVID-19, but few contemplate the possibility or likelihood of false-negative or false-positive rRT-PCR results. None of the studies discuss the possibility of various diagnostic biases (spectrum, incorporation, partial verification, differential verification, or imperfect criterion standard), nor the potential skew of these biases in observed estimates of sensitivity or specificity.

Clinical Examination

Fever is the most commonly reported finding in 84% to 87% of COVID-19 cases, but fever alone does not distinguish this virus from other infections. Therefore, absence of fever is inadequate for travel screening and likely for other decision thresholds such as whether ED staff can work shifts. Hyposmia (diminished sense of smell) and hypogeusia (diminished taste) have also emerged as COVID-19 symptoms. Both hyposmia (positive likelihood ratio \( LR^+ = 5.3 \), negative likelihood ratio \( LR^- = 0.61 \) ) and hypogeusia (\( LR^+ = 7.1 \), \( LR^- = 0.38 \)) are better to rule in than to rule out COVID-19, but neither may be fully adequate for either purpose. Although multiple COVID-19 studies report acute smell or taste disorders as a distinguishing symptom, no other studies report diagnostic...
accuracy or sufficient details to compute likelihood ratios for hyposmia or hypogeusia.\textsuperscript{16–21} Loss of smell is not necessarily associated with nasal obstruction or rhinorrhea.\textsuperscript{22} In one case–control study, new-onset smell and taste disorders are more common with COVID-19 than with influenza (39% vs. 13%).\textsuperscript{16} Consequently, influenza decision aids or diagnostic algorithms do not incorporate hyposmia or hypogeusia.\textsuperscript{23,24} Anosmia, which may be the only complaint in some COVID patients,\textsuperscript{25} is noted by 47% to 73% of COVID-19 patients and is the initial symptom in 27%.\textsuperscript{18–20} Additionally, 71% recall an acute onset of symptoms associated with taste and smell.\textsuperscript{16} Anosmia is more common in women and may persist for 2 weeks.\textsuperscript{17} Predictive models incorporating a change in taste or smell to distinguish COVID-19 from viral mimics appear most sensitive.\textsuperscript{26} Cough is only present in 58% patients.\textsuperscript{10,12} Neither cough, dyspnea, sore throat, nor fatigue distinguish COVID-19 from other illnesses,\textsuperscript{13} but current studies do not quantify accuracy.\textsuperscript{12,27}

**Routine Laboratory Tests**

Lymphopenia occurs in over 50% of COVID-19 patients.\textsuperscript{10,11,28} Neutrophil-to-lymphocyte and platelet-to-lymphocyte ratios do not distinguish COVID-19.\textsuperscript{29} Elevated lactate dehydrogenase (LDH) is also
frequently described.\textsuperscript{10,28,30,31} None of these laboratory findings are commonly utilized in the diagnosis of influenza, but their prevalence and accuracy to distinguish COVID-19 from other viral mimics merit further evaluation.\textsuperscript{23,24} Elevated prothrombin time (PT), ferritin, D-dimer, or IL-6 are associated with severe COVID-19.\textsuperscript{31–33} Existing studies do not report sensitivity or specificity of these laboratory tests.

**rRT-PCR**

Most studies use rRT-PCR as the criterion standard for diagnosing COVID-19. This test as used in current assays provides a qualitative detection of nucleic acid from the SARS-CoV-2 virus. Current research describes a variety of reagents and target RNA sequences with no accepted standard assay format. The U.S. CDC-developed rRT-PCR test detects two separate regions of the SARS-CoV-2 nucleocapsid gene (N1 and N2). Test results are considered positive when amplification of both the N1 and the N2 regions of the virus are detected and negative when both N1 and N2 are not detected. Detection of only one of the two amplified regions of the nucleocapsid gene results in an inconclusive test that must be repeated. This test can be performed on multiple diagnostic testing platforms and validated by laboratories under the U.S. FDA EUA.\textsuperscript{2} Inadequate supplies of reagents have restricted testing capacity and time to diagnosis in many settings,\textsuperscript{34,35} prompting laboratory researchers to explore the concept of specimen pooling in which multiple patients’ samples are tested simultaneously with further individual testing only if the pooled specimen is positive.\textsuperscript{36} The optimal pool specimen when COVID-19 community prevalence is less than 10% is four patients, which improves testing efficiency by 69%.\textsuperscript{37}

There is limited information on the diagnostic accuracy of the rRT-PCR test. Although an increasing number of studies provide head-to-head comparisons,\textsuperscript{38–40} systematic reviews provide little quantitative accuracy data and no meta-analysis or assessment of individual study quality.\textsuperscript{41} No rRT-PCR test is clearly superior to others in terms of diagnostic accuracy, but some provide faster results and commercial tests may be less sensitive than hospital-developed tests.\textsuperscript{40,42} It is known, however, that false negatives are frequent, so current recommendations advise incorporating patient’s exposure risk, clinical signs and symptoms, routine laboratory and imaging findings, serology, and (when available) CT results into real-time determination of COVID-19 status. Repeat or even serial rRT-PCR testing is required to confidently exclude COVID-19. Multiple studies report initially negative rRT-PCR results becoming positive with subsequent rRT-PCR tests in the following days or weeks.\textsuperscript{43,44} Others report hospitalized COVID-19 patients with initially positive rRT-PCR tests becoming negative prior to discharge with subsequent readmission for positive tests in the ensuing days.\textsuperscript{45} Ren et al.\textsuperscript{46} noted rRT-PCR sensitivity with one test was 78% and increased to 86% with a second test. A strategy of three negative rRT-PCR results is superior to two negative rRT-PCR followed by bronchoalveolar lavage.\textsuperscript{47} Repeating initially negative rRT-PCR up to five times increases sensitivity to 98%.\textsuperscript{48} COVID-19 patients identified on first rRT-PCR often have more severe disease associated with higher mortality, likely due to higher quantities of virus in those individuals.\textsuperscript{48} Older patients are more likely to remain rRT-PCR positive for an extended period, but whether this means they are contagious has yet to be determined.\textsuperscript{44}

Potential reasons for false-negative rRT-PCR results are summarized in Table 1.\textsuperscript{49–51} Emergency physicians will rely on the rRT-PCR assay selected by their hospital laboratory, which may balance the limit of detection and sensitivity against turnaround time, complexity, cost, workflow, availability of reagents and kits, specimen type, and laboratory personnel risk handling those specimens.\textsuperscript{52} Patients under investigation for COVID-19 who ultimately rule out are rarely reported in currently available studies, so specificity and false positives are generally not reported in the literature. However, false positives appear rare.\textsuperscript{53} In CDC testing, there was no significant cross-reactivity with other common respiratory viruses or seasonal coronaviruses.\textsuperscript{54} Contamination of the specimen or reagents used in the rRT-PCR is therefore the main mechanism for false-positive results. The CDC recommends protocols to prevent and detect potential contamination to minimize false-positive results.\textsuperscript{49,54}

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Common Causes of False-negative rRT-PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laboratory handling (heat inactivation)</td>
<td>Limit of detection (RNA particle detection)</td>
</tr>
<tr>
<td>Mutations in the probe target</td>
<td>Selective virus replication (patient variability, disease severity variability)</td>
</tr>
<tr>
<td>Sampling procedure (training, fidelity, patient cooperation)</td>
<td>Specimen sampled (NP, OP, saliva, sputum, BAL, stool)</td>
</tr>
<tr>
<td>Test kit quality</td>
<td>Timing of sampling in course of disease</td>
</tr>
</tbody>
</table>

BAL = bronchoalveolar lavage; NP = nasopharyngeal; OP = oropharyngeal; rRT-PCR = real-time reverse transcription–polymerase chain reaction.
Nasopharyngeal (NP) samples are most commonly obtained and studied, but oropharyngeal (OP), saliva, sputum, stool, blood, and/or urine specimens can also be evaluated. Obtaining NP samples requires time and appropriate training, increases exposure to staff secondary to coughing and gagging, and is uncomfortable for patients. Methodologically, few rRT-PCR accuracy studies describe how research or clinical staff were trained to collect NP specimens, so fidelity and reproducibility remain in question. Wang et al. provide videos describing NP and OP collection methods and note poor agreement between the two sampling methods (kappa = 0.31) and a higher yield with NP. Saliva can be collected outside the hospital without training, perhaps as part of a telemedicine evaluation for COVID-19. One small Italian study indicated that saliva specimens demonstrate detectable SARS-CoV-2 and the limit of detection is not affected by patient age. Sputum samples exhibit higher viral load than OP sites. However, as already noted many patients under investigation lack a cough and fewer still have sputum production. Expectoration of sputum may also expose health care workers collecting the sample to aerosols that would not have been generated without a sample collection attempt. Furthermore, a Bayesian analysis of prevalence-dependent positive and negative predictive values by Ghosal and Sinha demonstrates that even when the COVID-19 prevalence is high (54%), the positive predictive value (PPV) of sputum rRT-PCR is only 95.7%, and the negative predictive value (NPV) is 52%. PPV and NPV vary with disease prevalence. Specifically, PPV increases with higher disease prevalence and NPV increases with lower disease prevalence making extrapolation to clinical populations challenging if the study prevalence does not match the patients to whom the test is applied. For this reason, diagnosticians prefer likelihood ratios. Blood and urine are inadequate specimens for rRT-PCR because most patients do not exhibit virus in these body fluid compartments. In addition to the CDC-developed rRT-PCR test, manufacturers have developed molecular tests that target different portions of the SARS-CoV-2 viral genome and are performed on rapid testing platforms. For example, reverse transcription loop-mediated isothermal amplification can detect SARS-CoV-2 within 30 minutes. Other laboratories are exploring high-throughput sequencing for inconclusive fluorescence quantitative PCR specimens as a rapid mediator for the presence or absence of SARS-CoV-2. The diagnostic accuracy of these tests is similarly not reported, but these tests have not shown cross-reactivity to other respiratory viruses and bacteria.

**Antigen Tests**

On May 8, 2020, the U.S. FDA issued an EUA for a SARS-CoV-2 antigen test. This test detects SARS-CoV or SARS-CoV-2 nucleocapsid protein antigens in NP or nasal specimens using a lateral-flow immunofluorescent sandwich assay. This assay is performed on a point-of-care device in laboratories that are able to perform high, moderate, or waived complexity tests and can provide results within minutes. While the diagnostic accuracy of this test is not available at this time, the FDA reports high specificity but sensitivity that is less than that of rRT-PCR. The FDA and the manufacturer recommend that negative results “be treated as presumptive and confirmed with a molecular assay, if necessary for patient management.”

**Chest X-ray**

Chest X-ray is essential to evaluate for COVID-19 mimics such as pneumonia, pleural effusion, or pulmonary edema. Typical COVID-19 findings include hazy opacities that are often bilateral and peripheral. With the exception of one outlier, the reported sensitivity of single-view chest X-ray for COVID-19 ranges from 33% to 60%. Chen et al. represent the outlier reporting 100% accuracy of chest X-ray early in the COVID-19 pandemic. Currently available COVID-19 chest X-ray studies do not report specificity or reliability.

**CT of the Chest**

Computed tomography findings suggesting COVID-19 include ground glass opacity (often bilateral) and peripheral predominant lesions without mediastinal adenopathy or pleural effusions, although these findings represent nonspecific manifestations of acute lung injury associated with numerous infectious and noninfectious etiologies. Incidental findings consistent with COVID-19 are observed on CT of the chest in patients without respiratory symptoms. Multiple studies report COVID-19 identified by CT after one or more negative rRT-PCR tests. When the COVID-19 epidemic erupted in China, clinicians lacked access to rRT-PCR kits and then as rRT-PCR became available, low rRT-PCR sensitivity reinforced belief in the additive value of CT for many. These observations and scenarios prompted some to advocate for CT as a first-line supplement to the diagnostic
evaluation for COVID-19, combining rRT-PCR with CT.\textsuperscript{89,90} If CT alone or in combination with rRT-PCR reduced false-negative rates, the positive public health implications for case identification and control of disease transmission could be substantial. However, these benefits must be balanced against the cost of CT, medical radiation dangers, or practical limitations in busy hospitals with hourly trauma and stroke arrivals and potentially time-dependent emergencies juxtaposed against advised CT shutdowns for COVID-19 cleaning requiring 30 or more minutes.\textsuperscript{91,92} This cleaning time would also delay access to CT for every patient in the ED, thereby prolonging potential exposure to those in the ED to other patients with COVID-19.\textsuperscript{92} Some propose that COVID-19 patients wear N95 masks and plastic bags over their heads to eliminate or reduce these cleaning times.\textsuperscript{27} Pragmatically, among those detected by CT no defined benefit, such as reduced mortality or faster resolution of COVID-19 symptoms has been described.\textsuperscript{93} In addition, radiologists’ sensitivity for diagnosing COVID-19 by CT findings ranges from 72% to 94% with specificity from 24% to 100%.\textsuperscript{94} Preliminary artificial intelligence studies report radiologists’ sensitivity improves from 79% to 88% and specificity from 88% to 91% with this artificial intelligence image augmentation,\textsuperscript{95} while others hypothesize that the most valuable role for this technology may quantify the proportion of lung affected by COVID-19.\textsuperscript{96} Despite these issues, multiple studies highlight the fact that the sensitivity of CT is substantially higher than that of the first rRT-PCR, while combining CT and rRT-PCR provides maximal sensitivity (~97%).\textsuperscript{89,97,98} Theoretically, the sensitivity of CT would decline when testing populations outside of an epidemic (low prevalence rates), while specificity would be reduced when mimics like influenza are more common.\textsuperscript{9,27} The British Society of Thoracic Imaging recommends against CT when rRT-PCR is positive, but to consider CT when the initial rRT-PCR is negative to identify coexisting disease or potential COVID-19 complications such as pulmonary embolism.\textsuperscript{99} Tavare et al.\textsuperscript{100} developed a single-center protocol to prioritize inpatient CT decision making for initial COVID-19–negative rRT-PCR patients based on initial clinical suspicion and chest X-ray findings.

**Serology**

The U.S. FDA has also issued an EUA for the development of SARS-CoV-2 antibody tests. These antibody tests detect circulating IgM, IgG, or both that are reactive against SARS-CoV-2 virus using lateral-flow assays (LFAs) or enzyme-linked immunosorbent assay (ELISA).\textsuperscript{2} However, unlike rRT-PCR tests, there are data regarding the diagnostic accuracy of these serologic tests. Whitman and colleagues\textsuperscript{101} evaluated 10 LFA and two ELISA tests for SARS-CoV-2 antibodies. They used plasma or serum samples from patients with symptomatic, rRT-PCR confirmed-positive patients as the criterion standard for disease, and pre–COVID-19 specimens from the American Red Cross as negative controls. Sensitivity of both IgM and IgG varied by days since symptoms onset, with sensitivities for either IgM or IgG at >20 days ranging from 82% to 100%. Specificity for either IgM or IgG also varied by test, ranging from 87% to 100%.\textsuperscript{101} Similarly, Bendavid and colleagues\textsuperscript{102} used a commercially available LFA test to perform a seroprevalence study in Santa Clara county, California, and reported a sensitivity of 80% and a specificity of 99.5%.

True-positive serologic tests for SARS-CoV-2 antibodies indicate prior infection with SARS-CoV-2 and the development of an immune response. This may be helpful in identifying those who were asymptomatic or minimally symptomatic at the time of infection as well as those who were unable to receive a molecular test when symptomatic. While some experts believe that the presence of IgM or IgG reactive against SARS-CoV-2 will confer immunity,\textsuperscript{103} this has not yet been shown.\textsuperscript{104} If the presence of antibodies on a true-positive serologic test does confer immunity, the titer of antibodies required to confer immunity remains unknown, as does the duration of that immunity.

False-positive results may be due to cross-reactivity with other coronavirus strains that cause the common cold. The FDA recommends the following information be included in the instructions for use and patient test reports:

Positive results may be due to past or present infection with non–SARS-CoV-2 coronavirus strains, such as coronavirus HKU1, NL63, OC43, or 229E.\textsuperscript{105}

**DISCUSSION**

Knowledge of the diagnostic characteristics, including sensitivity, specificity, and likelihood ratios of tests for SARS-CoV-2, the virus that causes COVID-19, is important to understand how to best apply these tests for patient care and disease surveillance. Because this
novel virus emerged as a significant pathogen in humans only a few months ago, diagnostic tests have been developed rapidly under FDA EUAs in the United States. Consequently, we have less information about the diagnostic accuracy of these tests than we would under normal circumstances, but we do know that both false negatives and false positives may occur. An illustration of the false-positive and false-negative rate as a function of prevalence for two serologic tests for SARS-CoV-2 is provided in Figure 3.

**False-negative Test Implications**

False-negative tests commonly occur with rRT-PCR tests for several reasons (Table 1). There are a number of potential implications of a false-negative rRT-PCR test for SARS-CoV-2. From the patient’s standpoint, a patient with a negative test may lead to an assumption that they are not infected and subsequently diminished adherence to instructions to isolate and take other infection control measures, increasing the risk of infecting others. In the hospital setting, precautions may be relaxed in the presence of a negative test, increasing the risk of transmission to health care workers and other patients. In patients with moderate or high pretest probability of disease, a negative test may not reduce the posttest probability of disease below a level where precautions to prevent spread of disease become less necessary. In patients with a low pretest probability of disease, the likelihood of disease given a negative test will be low. However, even low individual likelihoods of disease can cumulatively contribute to substantial risk of outbreaks across larger groups for more contagious infectious diseases, such as COVID-19.

False-negative tests are also a consideration with serologic testing. However, since these tests should generally not be used to assess an active infection, the risks of a false negative are less significant for disease transmission. A false-negative serologic test would incorrectly classify a person as not having an immune response to SARS-CoV-2. If “immunity passports” became a reality, this could incorrectly and adversely affect a person’s ability to travel or work.104

**CT as Diagnostic Adjunct to Reduce False negatives**

The increased sensitivity of CT for COVID-19 might provide a net public health benefit if false-negative rRT-PCR patients with higher clinical suspicion were accurately identified during the initial ED evaluation. Mathematical models provide a theoretical basis for the concept that increasing diagnostic efficiencies (for example, by improving sensitivity with addition of CT) will decrease the risk of COVID-19 transmission.106 Pending the availability of rapid, reliable, and sufficiently accurate COVID-19 tests in ED settings, the identify—isolate–inform (3I) approach to decrease spread might

---

Figure 3. False-positive and false-negative rates as a function of pretest probability (or prevalence for surveillance studies) for serologic tests for SARS-CoV-2 antibodies. The left side of the graph illustrates the false-positive rate, and the right illustrates the false-negative rate.
be improved with more liberal CT use. One Italian hospital reported liberal CT screening for respiratory patients with possible false-negative chest X-ray results, but thus far has not reported on the positive or negative impact of that approach on individual patient care or public health. In the early stages of COVID-19, as many as 50% of patients with respiratory symptoms may have normal imaging. Radiologists have also noted that the quality of early CT accuracy studies is questionable because the rRT-PCR assay used as either the criterion standard is not described or the accuracy of that standard is undefined. In addition, CT findings are not pathognomonic for COVID-19 as influenza, cytomegalovirus, and atypical pneumonia have similar findings. As a consequence of these CT limitations in addition to the costs, radiation exposure, and downstream effect on other patients in terms of diagnostic delays and cross-contamination, multiple groups, including the Fleischner Society and the British Society of Thoracic Imaging, discourage CT as a routine screening approach. Nonetheless, CT likely plays a role when rRT-PCR tests are either too inaccurate or unavailable or suffer unacceptably slow turnaround times in patients with higher levels of COVID-19 concern based on exposure history or other clinical findings. The public health benefits of a more liberal CT screening approach from the ED merit additional research.

False-positive Test Implications
False-positive tests associated with rRT-PCR for SARS-CoV-2 are believed to be rare and would most commonly occur due to contamination. False-positive tests may occur more commonly with serologic tests, which have reported specificities ranging from 87% to 100%. The PPV is a function of both test sensitivity and specificity and the pretest probability of disease. This implies that positive test results are more likely to represent false-positive results when the pretest probability of disease is low. The instability of PPV is especially important as we apply imperfect diagnostic tests to low-risk patient populations, such as asymptomatic patients in low-prevalence communities. As an example, the antibody test used in a California community study has a reported sensitivity of 80.3% and a specificity of 99.5% (LR+ = 161, LR− = 0.20). Given the high specificity, clinicians expect a low number of false-positive tests. However, in a community with a low prevalence, for example, 1.25% (similar to the corrected raw prevalence in the California study), the posttest probability of COVID-19 given a positive rRT-PCR is only 67%, with 33% resulting in false positives.

For diagnostic tests, a manipulation of Bayes’ theorem can illustrate the effect of pretest probability of disease on false positives by determining the pretest probability at which false positives are equal to true positives. Above this probability, true positives exceed false positives, while below this probability, false positives exceed true positives. The equation is

$$Pretest\ Probability(\ TP = FP) = \frac{1}{\frac{Sn}{Sp}}$$

For the data in the study by Bendavid et al., the pretest probability at which true-positive and false-positive results are equally likely is 0.62%. However, for the Epitope ELISA results provided in the study by Whitman et al. with a sensitivity of 90.91% and a specificity of 81.82%, the probability at which true-positive and false-positive results are equally likely is 10%. Thus, if a patient with a <10% pretest probability of disease received a positive test result using this assay, it would be more likely to be a false positive than a true positive.

There are a number of potential adverse consequences of a false-positive rRT-PCR test for SARS-CoV-2. The CDC has published a Fact Sheet for Healthcare Providers that states that:

The CDC 2019-nCoV Real-Time RT-PCR Diagnostic Panel has been designed to minimize the likelihood of false positive test results. However, in the event of a false positive result, risks to patients could include the following: a recommendation for isolation of the patient, monitoring of household or other close contacts for symptoms, patient isolation that might limit contact with family or friends and may increase contact with other potentially COVID-19 patients, limits in the ability to work, the delayed diagnosis and treatment for the true infection causing the symptoms, unnecessary prescription of a treatment or therapy, or other unintended adverse effects.

For serology tests, the above risks of false positives could also exist if the test was thought to represent an active infection. Conversely, there are additional potential risks to patients and society with a false-positive serologic test for SARS-CoV-2 antibodies. Patients may
assume that they have developed immunity to COVID-19, leading to a reduction in risk-mitigating activities such as physical distancing. Health care workers with false-positive tests may similarly reduce their vigilance and use of precautions due to an incorrect assumption that they have immunity. This could place these individuals and their close contacts at increased risk of contracting COVID-19.

Potential Diagnostic Biases Skew Observed Accuracy

Multiple forms of diagnostic bias exist and each skew measured estimates of sensitivity and specificity in different directions. Incorporation bias is possible when the criterion standard includes the index test (for example, rRT-PCR) in ultimately determining whether the disease is present or absent. Incorporation bias increases measured sensitivity and specificity. This is pertinent to COVID-19 because most early studies incorporate rRT-PCR into the criterion standard. Differential verification bias is possible when patients with a positive or concerning index test (e.g., CT findings associated with COVID-19) are more likely to receive an immediate invasive criterion standard such as repeat rRT-PCR testing or bronchoalveolar lavage specimens. Differential verification bias raises specificity in diseases that resolve spontaneously or lowers specificity for diseases that only become detectable during follow-up. Imperfect criterion standard bias is possible when the standard used to classify the presence or absence of disease misclassifies some patients. Imperfect criterion standard bias raises observed specificity if errors on the index test and “copper standard” are correlated with true disease status and lowers observed specificity if errors on the index test and the copper standard are independent. This is pertinent to COVID-19 because no well-accepted criterion standard yet exists. A better criterion standard for COVID-19 will certainly emerge and we propose some ideas in Table 2. Temporal bias reflects variation in observed accuracy based on the period of time or stage of disease when index testing occurred. In COVID-19 viral shedding is highest in the early stages of disease with the highest positive rates noted within the first week. Spectrum bias is possible when the spectrum of disease severity differs between the study and clinical application (for example, critically ill COVID-19 patients in the intensive care unit vs. asymptomatic patients evaluated in ambulatory clinics). Spectrum bias skews observed sensitivity upward in sicker populations and skews specificity upward in healthier patients. Spectrum bias is worth considering when applying diagnostic accuracy results from patients with varying severity of illness to dissimilar populations. For example, among COVID-19 patients from cruise ships evaluated with CT, those with symptoms more commonly had COVID-19 CT findings than those without symptoms (80% vs. 40%).

Implications for Future Research

The rapidly expanding evidence base around COVID-19 diagnostic accuracy for clinical examination, routine laboratory tests, imaging, and advanced testing provides important lessons moving forward for clinicians, researchers, and journal reviewers. COVID-19 researchers need to contemplate myriad biases carefully in reporting observed diagnostic accuracy. If a bias is likely and the anticipated skew in observed sensitivity or specificity is upward and the observed accuracy is already too low, further studies of that diagnostic test may not be warranted. The STARD reporting guidelines provide manuscript protocols to ensure adequate description of methods and results so that diagnostic biases are easier to identify. Unfortunately, none of the early COVID-19 diagnostic research cites STARD or adheres to these reporting standards, which is not uncommon in emergency medicine.

Clinical decision aids consist of three or more findings on history, physical examination, routine laboratory tests, or imaging that, in combination, more accurately identify patients at lower or higher risk of a disease or outcome. Diagnostic and prognostic decision aids are commonly developed and employed in emergency medicine to reduce practice variability without compromising patient outcomes. Efforts to develop COVID-19 decision aids might include something like the Pulmonary Embolism Rule-Out Criteria (PERC) rule to identify subsets of ED patients at lower

### Table 2

Proposed COVID-19 Criterion Standard

<table>
<thead>
<tr>
<th>Expert consensus months after acute illness, including</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Exposure history</td>
</tr>
<tr>
<td>• Symptoms</td>
</tr>
<tr>
<td>• Laboratory tests</td>
</tr>
<tr>
<td>• rRT-PCR</td>
</tr>
<tr>
<td>• Imaging</td>
</tr>
<tr>
<td>• Serology</td>
</tr>
<tr>
<td>• Viral cultures</td>
</tr>
</tbody>
</table>

rRT-PCR = real-time reverse transcription-polymerase chain reaction.
risk of COVID-19 pending definitive testing.\textsuperscript{123} Alternatively, a decision aid might serve prognostic purposes to identify COVID-19 patients more likely to decompensate in response to the viral infection.\textsuperscript{124–127} When decision aid investigators attempt to derive and validate these instruments, higher-quality emergency medicine research quantifying accuracy and reliability (or the elements of history, physical examination, laboratory tests, and imaging that become predictor variables of the decision aid) will be required as the basis for selecting variables likely to improve model performance.

Most laboratory tests are quantitative rather than qualitative, including COVID-19 rRT-PCR and serologic assays. When sensitivity and specificity are reported, the quantitative laboratory tests have been dichotomized at some level. Another approach to evaluating diagnostic accuracy for quantitative data is interval likelihood ratios (iLR).\textsuperscript{128} One advantage of iLR’s is that indeterminant results are more readily interpreted. As COVID-19 diagnosticians identify the rRT-PCR and serologic tests that best balance availability, accuracy, reliability, and cost, reporting iLR’s could provide added value for clinicians.

Ultimately, ED physicians’ clinical impressions concerning the presence or absence of COVID-19 are communicated to patients and families—usually without access to definitive testing. Patient communication tools to convey the basics of COVID-19 personal protection and infection prevention exist,\textsuperscript{129,130} but shared decision-making instruments that communicate the uncertainties of clinical examination, imaging, and even rRT-PCR do not exist. Figure 4 provides one example of a Cates plot that could be used to communicate the accuracy limitations of rRT-PCR based on current evidence. Actual shared decision-making instruments will require scientific investigation using accepted methodology before widespread implementation.\textsuperscript{131}

**LIMITATIONS**

This scoping review has several limitations. The pace of publications around COVID-19 and diagnostics in the first half of 2020 has been astonishing. At best, this review will serve as a snapshot in time, although hopefully illuminating issues that require higher methodologic standards and peer-review attention moving forward. Due to time constraints, the search strategy was limited to English language and published research. More research undoubtedly exists in the gray literature. Since earlier systematic reviews exploring aspects of COVID-19 diagnostic testing did not identify or report additional measures of sensitivity, specificity, or likelihood ratios for hyposmia, hypogeusia, or rRT-PCR, we are confident that this search presents a complete scoping review of current knowledge.\textsuperscript{10,131} Others have also noted the absence of diagnostic accuracy reporting amidst the flurry of COVID-19 publications.\textsuperscript{133,134} Additionally, this scoping review does not focus on special emergency medicine populations such as pediatrics, geriatrics, or obstetrics because other reviews already exist for these patients.\textsuperscript{135,136} Most importantly, we report no formal assessment of study quality using accepted instruments such as the QUADAS-2,\textsuperscript{137} although informal assessment of published research to date suggests limited adherence to the full set of recommended methodologic standards for studies of diagnostic test performance.

**CONCLUSIONS**

Clinicians should be aware of the current limited knowledge around history, physical examination, laboratory tests, and imaging for COVID-19. Fever and acute-onset disorders of taste and/or smell are the most common findings on history and physical examination associated with COVID-19. Lymphopenia is associated with COVID-19 diagnosis, while elevated lactate dehydrogenase and prothrombin time are associated with severe disease. rRT-PCR has emerged as the primary diagnostic test for suspected COVID-19, but access has been limited, diagnostic accuracy is underreported, and between-assay comparative accuracy is rarely evaluated. However, typical testing algorithms and diagnostic accuracy studies rely heavily on rRT-PCR results with frequent false negatives. Chest computed tomography is indicated for equivocal cases or when considering diagnoses like pulmonary embolism but is not recommended as a general screening protocol. In cases with high clinical suspicions, repeat real-time reverse transcription–polymerase chain reaction testing with or without computed tomography scanning may be beneficial to reduce community spread. Antigen tests have only recently been approved, and diagnostic accuracy information is similarly limited. Serology may identify past COVID-19 exposure, but the role of antibody testing and implications for ED decision making remain undefined. Current clinical, imaging, and laboratory studies neglect diagnostic accuracy reporting standards and likely suffer from various biases.
The authors acknowledge the contributions of Washington University in St. Louis School of Medicine’s medical librarian Michelle Doering for providing electronic medical database search expertise. They also acknowledge illustrator Kai Choumanivong for producing Figure 1.
References


58. Sullivan PS, Sailey C, Guest JL, et al. Detection of SARS-CoV-2 RNA and antibodies in diverse samples: protocol to validate the sufficiency of provider-observed home-
collected blood, saliva and oropharyngeal samples. JMIIPub Health Surveill. 2020;6:e19054.
63. Ghosal S, Sinha B. Rapid sputum testing and not thermal screening alone should be the first-line screening test at airports: a Bayesian analysis. Diabetes Metab Syndr 2020;14:317–8.


Supporting Information
The following supporting information is available in the online version of this paper available at http://onlinelibrary.wiley.com/doi/10.1111/acem.14048/full
Data Supplement S1. Appendix.
Myocardial Infarction Can Be Safely Excluded by High-sensitivity Troponin I Testing 3 Hours After Emergency Department Presentation

W. Frank Peacock, MD1, Robert Christenson, PhD2, Deborah B. Diercks, MD3, Christian Fromm, MD4, Gary F. Headen, MD5, Christopher J. Hogan, MD6, Erik B. Kulstad, MD, MS3, Frank LoVecchio, MD7, Richard M. Nowak, MD, MBA8, Jon W. Schrock, MD9, Adam J. Singer, MD10, Alan B. Storrow, MD11, Joely Straseski, PhD12, Alan H. B. Wu, PhD13, and Daniel P. Zelinski, MD, PhD14

ABSTRACT

Background: The accuracy and speed by which acute myocardial infarction (AMI) is excluded are an important determinant of emergency department (ED) length of stay and resource utilization. While high-sensitivity troponin I (hsTnI) >99th percentile (upper reference level [URL]) represents a “rule-in” cutpoint, our purpose was to evaluate the ability of the Beckman Coulter hsTnI assay, using various level-of-quantification (LoQ) cutpoints, to rule out AMI within 3 hours of ED presentation in suspected acute coronary syndrome (ACS) patients.

Methods: This multicenter evaluation enrolled adults with >5 minutes of ACS symptoms and an electrocardiogram obtained per standard care obtained. Exclusions were ST-segment elevation or chronic hemodialysis. After informed consent was obtained, blood samples were collected in heparin at ED admission (baseline), ≥1 to 3, ≥3 to 6, and ≥6 to 9 hours postadmission. Samples were processed and stored at −20°C within 1 hour and were tested at three independent clinical laboratories on an immunoassay system (Dxi 800, Beckman Coulter). Analytic cutpoints were the URL of 17.9 ng/L and two LoQ cutpoints, defined as the 10 and 20% coefficient of variation (5.6 and 2.3 ng/L, respectively). A criterion standard MI diagnosis was adjudicated by an independent endpoint committee, blinded to hsTnI, and using the universal definition of MI.
Results: Of 1,049 patients meeting the entry criteria, and with baseline and 1- to 3-hour hsTnI results, 117 (11.2%) had an adjudicated final diagnosis of AMI. AMI patients were typically older, with more cardiovascular risk factors. Median (IQR) presentation time was 4 (1.6–16.0) hours after symptom onset, although AMI patients presented ~0.5 hour earlier than non-AMI. Enrollment and first blood draw occurred at a mean of ~1 hour after arrival. To evaluate the assay’s rule-out performance, patients with any hsTnI > URL were considered high risk and were excluded. The remaining population (n = 829) was divided into four LoQ relative categories: both hsTnI < LoQ (Lo-Lo cohort); first hsTnI < LoQ and 2nd > LoQ (Lo-Hi cohort); first > LoQ and second < LoQ (Hi-Lo cohort); or both > LoQ (Hi-Hi cohort). In patients with any hsTnI result <20% CV LoQ (Groups 1–3), n = 231 (23.9% ruled out), AMI negative predictive value (NPV) was 100% (95% confidence interval [CI] = 98.9% to 100%). In patients with any hsTnI below the 10% LoQ, n = 611 (58% rule out), AMI NPV was 100% (95% CI = 99.5% to 100%). Of the Hi-Hi cohort (i.e., no hsTnI below the 10% LoQ, but both < URL), there were four AMI patients, NPV was 98.2% (95% CI = 95.4% to 99.3%), and sensitivity was 96.6.

Conclusions: Patients presenting >3 hours after the onset of suspected ACS symptoms, with at least two Beckman Coulter Access hsTnI < URL and at least one of which is below either the 10 or the 20% LoQ, had a 100% NPV for AMI. Two hsTnI values 1 to 3 hours apart with both < URL, but also >LoQ had inadequate sensitivity and NPV.

Of the more than 145 million annual emergency department (ED) visits, it is estimated that 7.6 million patients present with a chief complaint of chest pain. An additional 3.4 million and 2.8 million will have complaints of the potential anginal equivalents of shortness of breath or vomiting, respectively. Because manifestations of coronary artery disease represent the number one cause of death in the United States, and the incidence rate of acute myocardial infarction (AMI) in these populations is often less than 20%, the initial evaluation in these nearly 14 million patients is intended to safely exclude the diagnosis of AMI. This is necessary because patients discharged from the ED with an undiagnosed AMI have worse outcomes that may include mortality. No contemporary estimates of the missed MI rate exist, but it is a generally unacceptable outcome, the concern for which results in extensive (and most commonly negative) evaluations. Although no MI recognition strategy has demonstrated infallibility, surveys of emergency physicians suggest an acceptable AMI and subsequent multiple adverse cardiac event miss rate not exceeding 1% is desirable. Historically, successful strategies demonstrating a negative predictive value (NPV) in excess of 99% required biomarker testing for at least 6 hours after presentation, which were then followed by further risk stratification techniques (e.g., myocardial perfusion evaluations).

High-sensitivity troponin assays have held the promise of a more accurate and rapid AMI rule outs. Unfortunately, due to a challenging regulatory environment, the United States is a decade behind the rest of the world in troponin research and clinical availability. It is only in the past 2 years that any high-sensitivity troponin assay was FDA cleared and thus available for U.S. physicians. This has resulted in the majority of the high-sensitivity troponin literature originating from European populations that have markedly higher AMI rule in rates than in the United States. Since the rates of MI are much lower in the United States, our different pretest odds create challenges in understanding the predictive values of high-sensitivity troponin when applied to our population.

The availability of high-sensitivity troponin assays has created an opportunity to shorten the time required for evaluation, as well as increase the number of patients who may be safely discharged based on the reported values, or changes in reported values, undetectable by contemporary assays. However, to minimize the probability of patient’s misclassification, the ability to measure low concentrations of an analyte must be considered in conjunction with the precision of that result, as it is not useful to have sensitive but imprecise troponin assays. The analytic imprecision of an assay, termed the coefficient of variation (CV), generally increases with decreasing analyte concentrations, and the level of quantification (LoQ) is the cutpoint at which a predefined CV is observed. The FDA requires a CV of ≤20% for the reporting of clinical results; however, the CV of ≤10% has been suggested as optimal at the decision point. While high-sensitivity troponin I (hsTnI) levels above the 99th percentile (defined as the upper reference level [URL]) are used to define high risk, the absence of high risk does not equate to a rule out, and lower levels are required to safely exclude a subsequent diagnosis of AMI. Few emergency medicine studies have presented data using different LoQs as a rule-out cutpoint, so the...
consequence on clinical impact is unclear.\textsuperscript{4,5} Since LoQ values are highly assay-specific, our objective was to evaluate the Beckman Coulter hsTnI assay using LoQ cutoffs of 10 and 20% CV as rule-out cutpoints, to exclude MI within 3 hours of ED presentation in patients with suspected acute coronary syndromes (ACSs).

**METHODS**

The database used for this analysis was prospectively collected during 2012 and 2013 as a part of an FDA submission for the Beckman Coulter hsTnI assay, with a patient cohort used in the evaluation of the AccuTnI+3 contemporary TnI assay and is described elsewhere.\textsuperscript{7} Local data are kept by the enrolling institution a minimum of 7 years after the FDA clearance and are kept in conglomerate by both the sponsor and the FDA. The analysis for this study was performed by Beckman Coulter.

The original study was a convenience sample investigation enrolling adult patients (age $\geq 21$ years) who presented to 14 geographically diverse, hospital-associated EDs. All participating hospitals obtained institutional review board approval and reflected urban, suburban, and rural patient populations. Eligible patients had an electrocardiogram (ECG) obtained as a part of their standard of care and reported at least 5 minutes of symptoms consistent with ACSs that included chest pain, shortness of breath, left arm pain, lightheadedness, dizziness, weakness, or syncope. Patients were excluded if the initial ECG demonstrated ST-segment elevation MI (STEMI), if they were on chronic hemodialysis, or if they were unable to provide informed consent. The primary endpoint was an adjudicated diagnosis of AMI.

After informed consent was obtained, serial blood samples were collected in lithium heparin tubes per the local standard of care and grouped into four different time frames: time of admission (baseline), $\geq 1$ to 3, $\geq 3$ to 6, and $\geq 6$ to 9 hours after baseline. Samples were centrifuged and then stored at $-20^\circ\text{C}$ within 1 hour of blood draw (the timing of which was reported in the FDA submission to not have analytic consequence). Plasma aliquots were tested at three independent clinical laboratories on the Access hsTnI DxI 800 immunoassay system (Beckman Coulter). The original study enrolled 1,929 patients of whom 75 were excluded due to insufficient sample volume and was powered for detection of 73 to 139 MIs with 90% to 95% diagnostic sensitivity.\textsuperscript{7} An additional 805 patients were excluded from the present retrospective analysis due to a missing baseline or 3-hour blood draw (these patients included those that developed a STEMI or were discharged/left the ED after a single sample).

The Beckman Coulter Access hsTnI assay is a two-site immunoenzymatic (“sandwich”) assay. For this assay, monoclonal anti–cTnI antibody conjugated to alkaline phosphatase is added to a reaction vessel along with a surfactant containing buffer and sample. After a short incubation period, paramagnetic particles coated with monoclonal anti–cTnI antibody are added. The sample cTnI then binds to the anti–cTnI antibody on the solid phase, while the anti–cTnI antibody–alkaline phosphatase conjugate reacts with a different antigenic site on the cTnI molecule. Materials bound to the solid phase are held in a magnetic field while unbound materials are washed away. Chemiluminescent substrate is added to the vessel, and light generated by the reaction is measured with a luminometer. The light production is directly proportional to the concentration of cTnI in the sample. The amount of analyte in the sample is determined from a stored, multipoint calibration curve, the analytical performance of which has been previously described.\textsuperscript{8-15}

**Data Analysis**

For this analysis, the URL for both sexes was defined as 17.9 ng/L and was derived from data that served as part of the FDA submission for this assay, as is documented in the package insert.\textsuperscript{7,16} The appropriate sample size was a priori determined for the FDA submission, but not for this secondary analysis. The LoQ was defined as the 20% CV cutpoint, occurring at 2.3 ng/L TnI, and a higher 10% CV cutpoint, which equaled 5.6 ng/L TnI. A clinically significant delta (that exceeding the larger of the CV’s used in this analysis) was subsequently defined as an hsTnI change $>25$%.

A clinical criterion standard index MI diagnosis was adjudicated by an independent clinical endpoints committee that consisted of four cardiologists, using criteria consistent with the Third Universal Definition of MI and the local troponin.\textsuperscript{4} Adjudicators were blinded to the Beckman Coulter assay results. Because the intent of this analysis was to identify low-risk cutpoints, patients adjudicated as MI were not differentiated into subtypes. All results presented here were based on the adjudicated diagnoses using the local
contemporary troponin assay. The overall patient population was characterized using descriptive statistics. Outcome rates are presented with sensitivity, specificity, and predictive values.

To evaluate the rule out capability of the assay, any patient with either a time 0 or a 1-to 3-hour hsTnI above the URL was considered higher risk and was excluded from the AMI rule-out category. The remaining population with TnI values below the URL at both time points could fall into one of four prospectively defined categories based on hsTnI levels relative to the two selected LoQ’s. They could have both measures below the LoQ (Lo-Lo cohort), the first measure below and the second above the LoQ (Lo-Hi cohort), the first above and the second below the LoQ (Hi-Lo cohort), or both concentrations could be above the LoQ (Hi-Hi cohort).

RESULTS

Overall, 1,049 patients met the entry criteria of serial hsTnI results available at baseline and between 1 and 3 hours. Demographic data using descriptive statistics and measures of dispersion are presented in Table 1. A total of 117 (11.2%) patients had an adjudicated diagnosis of MI. Those adjudicated as having a MI were older and more likely male with prior coronary artery disease and with a history of revascularization procedures as well as having known cardiovascular risk factors of hypertension, diabetes, tobacco use, heart failure, and renal disease. Patients presented a median (interquartile range [IQR]) of 4 (1.6–16.0) hours after symptom onset, although those diagnosed with MI presented almost 0.5 hour earlier. Enrollment and first blood draw occurred approximately 1 hour after arrival. Compared to the excluded patients (Table 1), the patients in this study were slightly younger and there were more females, with rates of prior coronary procedures, and they experienced a longer wait from presentation to first blood draw, with most of the significant differences occurring in non-MI group.

Of the total 1,049 patients, 220 had at least one hsTnI value above URL, of whom 113 were diagnosed with MI. Four additional AMI’s were adjudicated among 829 patients with both hsTnI values below the URL (see Figures 1 and 2). The diagnostic accuracy of using the URL cutpoint was as follows: sensitivity = 96.6% (95% CI = 91.3% to 99.0%), specificity = 88.5% (95% CI = 86.3% to 90.4%),

| Table 1 | Demographics and Comparison of Analytic (n = 1,049) and Excluded Subgroups (n = 805) |
|-----------------|---------------------------------|-----------------|-----------------|
| Category        | Subgroup (n = 1,049)       | Not in Subgroup (n = 805) | p-value           |
|                 | Non-MI, n = 932 (88.8%) | MI, n = 117 (11.2%) | Non-MI, n = 684 (85.0%) | MI, n = 121 (15.0%) | Non-MI | MI |
| Age (years), median (IQR) | 55 (48-64) | 62 (53-72) | 56 (49-68) | 62 (51-72) | 0.0002 | 0.7744 |
| Age ≥ 60 years | 36.1% | 55.6% | 41.5% | 57.0% | 0.0255 | 0.8197 |
| Male            | 48.5% | 66.7% | 57.9% | 73.6% | 0.0002 | 0.2466 |
| From symptom onset to presentation | 4.2 (1.6-16.8) | 3.7 (1.5-12.5) | 3.9 (1.7-16.0) | 2.9 (1.3-14.2) | 0.1124 | 0.1769 |
| From presentation to the first blood draw | 1.4 (1.0-1.9) | 1.2 (1.0-1.8) | 1.0 (0.5-1.5) | 0.7 (0.4-1.1) | <0.0001 | <0.0001 |
| Asian           | 1.0% | 3.4% | 4.1% | 2.831 | 0.4324 |
| African American | 39.1% | 34.2% | 23.8% | 23.1% | 0.2831 | 0.4324 |
| White           | 55.5% | 58.1% | 63.7% | 59.5% | 0.5765 | 0.9505 |
| Hypertension    | 70.3% | 76.9% | 69.0% | 77.7% | 0.0652 | 0.8105 |
| Hypercholesterolemia | 51.1% | 58.1% | 55.3% | 55.4% | 0.6907 | 0.7471 |
| Diabetes mellitus | 29.3% | 35.0% | 28.4% | 33.1% | 0.0110 | 0.1224 |
| Current smoker  | 29.8% | 24.8% | 23.7% | 33.1% | 0.0317 | 0.7220 |
| Past smoker     | 30.4% | 41.9% | 34.6% | 38.0% | 0.0494 | 0.9876 |
| Known >50% coronary stenosis | 27.5% | 46.2% | 30.4% | 43.8% | 0.0973 | 0.7428 |
| MI              | 21.2% | 35.9% | 24.1% | 32.2% | 0.0001 | 0.4027 |
| Coronary stent or angioplasty | 18.7% | 31.6% | 26.5% | 36.4% | 0.0001 | 0.0400 |
| CABG            | 9.0% | 14.5% | 13.2% | 14.9% | 0.1140 | 0.9188 |
| Heart failure   | 16.0% | 28.2% | 13.5% | 20.7% | 0.2960 | 0.2012 |
| CKD             | 9.3% | 13.7% | 6.9% | 13.2% | 0.9188 |

CABG = coronary artery bypass graft; CKD = chronic kidney disease; IQR = interquartile rank; MI = myocardial infarction.
positive predictive value (PPV) = 51.4% (95% CI = 44.8-57.9%), and NPV = 99.5% (95% CI = 98.7-99.9%). The high NPV obtained with this strategy did result in a poor PPV (51.4%) such that the clinical utility of a positive troponin in this setting is as likely as not to represent a MI.

The analysis of lower-risk cohorts is also shown in Figures 1 and 2. For either value of the LoQ, none of the patients with at least one hsTnI reading below the LoQ (Lo-Lo, Lo-Hi, Hi-Lo) were diagnosed with MI. Using the 20% CV LoQ, this strategy would identify 251 patients (24% of the total population) in whom the diagnosis of MI could be excluded with NPV = 100% (95% CI = 98.3-100%). The majority of patients in this cohort had both measures below the LoQ (n = 173, 69%), with fewer having a rising pattern with the second measure of hsTnI above the LoQ (n = 51, 20%). The fewest number had a falling hsTnI pattern (n = 27, 11%), with the first hsTnI above, and the second below, the LoQ. Of 51 subjects with a rising hsTnI pattern, the median (IQR) change was 0.7 (0.4–1.1 ng/L), with a maximum value of 16.7 ng/L. Performing the same analysis, with the 10% CV LoQ, resulted in an increase in the number of patients ultimately ruled out for MI to 611 (58% of the total population), while still maintaining a 100% NPV (95% CI = 99.6-100%) in the combined Lo-Lo, Lo-Hi, and Hi-Lo groups.

Four patients in Hi-Hi group with both hsTnI values between either LOQ (10% or 20%) and the URL, were adjudicated as MI (NPVs = 99.3 and 98.2%, respectively). Of these four, three had a local standard-of-care TnI above the local TnI URL within 3 hours, and the fourth had a standard of care TnI elevation identified 12 hours after presentation. In contrast, three high-risk (any hsTnI > URL) patients had an initial hsTnI below the LoQ and a second troponin above URL, but were ultimately adjudicated as non-MI. We also sought to determine if the magnitude of temporal fluctuations in troponin concentration in the Hi-Hi cohort (TnI < URL but >LoQ) could provide diagnostically useful information. Only four subjects in this subpopulation experienced an MI, making any conclusion from such limited data impractical.

An important question is if a single hsTnI < LoQ at baseline is sufficient for an AMI rule out. The
sensitivity, specificity, PPVs, and NPVs for the 20% LoQ (n = 224) were 100, 24, 14.2, and 100%, respectively, and 100, 63, 25.7, and 100%, for the 10% LoQ (n = 590), respectively. The fact that this approach was effective must be considered by the fact that the number of patients with early presentations is limited, and the clinician must precisely know the time of symptom onset (a challenging feat in some patients). The advantage of a two-blood-draw strategy is that it ensures that all patients have at least 3 hours of symptoms, or their symptoms have resolved for at least 3 hours.

Delineation of the cohort who presented with a baseline hsTnI < LoQ and subsequently ruled in for AMI (the penalty for not waiting for the second hsTnI) showed that, of the total population, 224 had a baseline hsTnI below the 20% LoQ, of whom zero subsequently ruled in for AMI. If the baseline hsTnI was below the 10% LoQ, then 593 patients were included, of whom three subsequently ruled in for AMI after 3 hours. Statistical performance of hsTnI in the early presenters provided a sensitivity, specificity, PPV, and NPV of 100, 26, 14.7, and 100% and 100, 65.6, 26.7, and 1000% for the 20% LoQ and the 10% LoQ, respectively. Specificity and PPV using the LoQ cutpoint alone are very low. However, a value below the LoQ is suggested as a ‘rule-out’ criteria, not a ‘rule-in’ criteria. Rule-in criteria should utilize the URL in line with current guidelines. This will provide much higher specificity for index MI. The distribution of the time from symptom onset to blood draw is presented in Figure 3. Note that ~30% of patients presented in the first 3 hours after symptom onset. A sensitivity analysis was performed in the 260 patients who presented to the ED within 3 hours of symptom onset. The incidence of MI in this “early presenter” population was 12.7% (33/260), and the NPV of both hsTnI values <URL was 99.5% (95% CI = 97.3% to 99.9%) for both LoQ values. This was similar to that of patients presenting after 3 hours of symptom onset, regardless of the LoQ cutpoint used, NPV = 99.5% (95% CI = 97.3% to 99.9%) for both LoQ values (see Table 2). While the sensitivity of the approach was 97.0% (95% CI = 83.0% to 100.0%), which is similar to that of the overall analysis, the low number of patients, and resulting wide CIs preclude reasonable extension of this finding to clinical practice.
In this prospective evaluation of an all comers ED population of patients presenting with potential acute coronary symptoms, we found that no patient with two hsTnI measurements below the URL cutpoint, and with at least one of which was below either the 10% or the 20% LoQ, had a diagnosis of AMI (Figures 1 and 2). It is important to note that being below the URL “rule-in cutpoint” does not equate to a rule out. Four patients with both levels below the URL were adjudicated as AMI and the sensitivity of this strategy was only 96.6%. However, considering the LoQ as a “rule-out” cutpoint, if one or both of the two values is below this LoQ, it represents a highly effective strategy for expediting management decisions in ED patients because it has a sensitivity and NPV of 100%. While other studies have utilized similar strategies, we are not aware of any that have reported NPVs of 100% for AMI in an “all-comers” ED patient population, while still ruling out up to 58% of the presenting population.

Our examination of two different LoQ values demonstrated that use of the higher 5.6 ng/L decision point, corresponding to a 10% CV LoQ, did increase the number of AMI rule outs, from 23.9% to 58% of all patients meeting the inclusion criteria, while still maintaining a 100% NPV. The high NPV obtained with LoQ-based strategy did result in a poor PPV (26.7%, 95% CI = 22.8% to 31.1%) requiring further

**DISCUSSION**

In this prospective evaluation of an all comers ED population of patients presenting with potential acute coronary symptoms, we found that no patient with two hsTnI measurements below the URL cutpoint, and with at least one of which was below either the 10% or the 20% LoQ, had a diagnosis of AMI (Figures 1 and 2). It is important to note that being below the URL “rule-in cutpoint” does not equate to a rule out. Four patients with both levels below the URL were adjudicated as AMI and the sensitivity of this strategy was only 96.6%. However, considering the LoQ as a “rule-out” cutpoint, if one or both of the two values is below this LoQ, it represents a highly effective strategy for expediting management decisions in ED patients because it has a sensitivity and NPV of 100%. While other studies have utilized similar strategies, we are not aware of any that have reported NPVs of 100% for AMI in an “all-comers” ED patient population, while still ruling out up to 58% of the presenting population.

Our examination of two different LoQ values demonstrated that use of the higher 5.6 ng/L decision point, corresponding to a 10% CV LoQ, did increase the number of AMI rule outs, from 23.9% to 58% of all patients meeting the inclusion criteria, while still maintaining a 100% NPV. The high NPV obtained with LoQ-based strategy did result in a poor PPV (26.7%, 95% CI = 22.8% to 31.1%) requiring further
risk stratification through clinical gestalt or risk-scoring approaches,
particularly in patients with both hsTnI values falling between LoQ and URL. Clinicians should still consider that hsTnI values on the upper end of the normal range have been associated with adverse long-term patient outcomes. In fact, in this study four patients within Hi-Hi cohort were diagnosed with MI. While earlier international publications on using Access hsTnI for rapid rule-out algorithms allowed only the consideration of the baseline hsTnI, our suggested approach is fully consistent with the fourth universal definition of MI and current U.S. guidelines for NSTEMI, because it relies on the sequential draws and considers patients with at least one value above URL as high risk.

Additionally, higher hsTnI values above the LoQ and below the URL were clinically useful, but only if the lower cutpoint was applied (LoQ = 20%, hsTnI = 2.3 ng/L). In this scenario, the lack of a detectable hsTnI increase (defined as delta <25% from baseline) had a NPV of 99.4% (95% CI = 98.3% to 99.9%) and was adequate for clinical decision making. However, if the physician used the higher LoQ (e.g., 5.6 ng/L), or the patient’s hsTnI was above 25%, the acceptable clinical threshold of an NPV above 99% is not met.

We addressed the challenge of defining a relevant troponin change as that exceeding 25%. We chose a 25% delta as the smallest possible relevant change to exceed the 20% CV, because a delta below the CV may simply represent laboratory error without clinical consequence. A percent delta is a somewhat arbitrary decision as we could have chosen absolute changes. The challenge of defining a percent delta is that when troponin levels are low, a percent change may represent clinically irrelevant changes (e.g., a change from 3 to 4 ng/L represents a 25% delta, but is only 1 ng/L). Conversely, we could have selected an absolute troponin change. This strategy is superior for the identification of low-level changes, but when applied to elevated troponins may not reflect a clinically relevant change. For example, a change of 5 ng/L in a patient with an initial hsTnI of 5 ng/L represents a doubling of the result and is likely to be clinically relevant, while a change of 5 ng/L in a patient with a troponin of 60 ng/L is likely not correlated with differences in outcomes or therapy.

It should be pointed out that certain cohorts of patients were excluded from this study and thus should be excluded from clinical adoption of an early rule-out strategy with this assay. This included patients with STEMI on initial ECG and those on dialysis. Further, because of the mechanics of performing a study (i.e., obtaining informed consent), the mean time for blood draw was 1 hour after ED arrival. While this represents earlier sampling than other previously published high-sensitivity troponin studies evaluating AMI rule out and is consistent with most ED patients presenting with the onset of ACS symptoms relative to timing of hsTnI measurements, the consideration of symptom onset is critical in diagnostic decision making. Patients presenting in less than 4 hours after symptom onset should not be considered to have ruled out with the strategy described herein until they have been able to have hsTnI measurements obtained at baseline and 3 hours afterward.

Many studies have been published using an accelerated diagnostic protocol strategy in the management of ED suspected ACS patients. These generally use risks scores (e.g., EDACS, HEART), ECG, and serial troponin obtained at baseline and 1, 2, or 3 hours. What is clear is that, as the sensitivity of the clinically available assays have improved, so have safe ED discharge rates as evaluated by 30-day outcomes. We cannot comment on safe discharge rates, since 30-day outcomes were not part of this study. Nonetheless, our findings strongly support the utilization of a highsensitivity troponin assay for excluding AMI diagnosis in an ED setting.

**LIMITATIONS**

This analysis evaluates the diagnostic value for AMI of the Beckman Coulter high-sensitivity troponin assay in a large prospectively obtained sample set. Post-ED discharge prognostic claims cannot be considered, because only the incident visit was evaluated by the physicians performing the adjudication. Further, since we only evaluated for the diagnosis of AMI and did not consider clinical risk scoring, myocardial perfusion evaluation (e.g., stress testing), or MI type, ED disposition should be considered with the possible necessity of further downstream testing. Additionally, this study was based on the gathering of samples for diagnosis but did not alter the standard of care in real time, such that future prospective studies may be needed to evaluate 30-day outcomes as a result of acting on this assay’s information in real time. Also, the present analysis and observations were based on a retrospective data exploration of a prospectively collected population. Consequently, some cohorts were of limited
size and had larger CIs that should be considered if applying these results to patients with matching characteristics. Moreover, the percentage of patients with an index MI is small (typical for U.S. populations) and this makes the NPV particularly high compared to sensitivity. These factors further underscore the need for prospective studies with prespecified endpoints and adequate statistical power to confirm our findings. Also, while some have supported the use of sex-specific cutpoints for the diagnosis of AMI, in regard to this assay, proof of the consequence of applying sex-specific cutpoints awaits the completion of additional investigations. Finally, patients presented a median of 3 hours after the onset of symptoms and had blood collected for this study an hour later, so that outcomes in individuals presenting earlier warrant further investigation.

CONCLUSIONS

Patients presenting >3 hours after the onset of suspected acute coronary syndrome symptoms, with at least two Beckman Coulter Access high-sensitivity troponin I < upper reference level, and at least one of which is below either the 10% or 20% level of quantification, had a 100% negative predictive value for acute myocardial infarction. Two high-sensitivity troponin I values 1 to 3 hours apart with both < upper reference level, but also > level of quantification had inadequate sensitivity and negative predictive value.

References


Novel Use of Home Pulse Oximetry Monitoring in COVID-19 Patients Discharged From the Emergency Department Identifies Need for Hospitalization

Sonia Shah, DO1, Kaushal Majmudar, DO2, Amy Stein, PhD3, Nita Gupta, MD1, Spencer Suppes, DO1, Marina Karamanis, DO1, Joseph Capannari, DO1, Sanjay Sethi, MD4, and Christine Patte, DO1

ABSTRACT

Objectives: Our objective was to evaluate patient-reported oxygen saturation ($\text{SpO}_2$) using pulse oximetry as a home monitoring tool for patients with initially nonsevere COVID-19 to identify need for hospitalization.

Methods: Patients were enrolled at the emergency department (ED) and outpatient testing centers. Each patient was given a home pulse oximeter and instructed to record their $\text{SpO}_2$ every 8 hours. Patients were instructed to return to the ED for sustained home $\text{SpO}_2 < 92\%$ or if they felt they needed emergent medical attention. Relative risk was used to assess the relation between hospitalization and home $\text{SpO}_2 < 92\%$ in COVID-19–positive patients.

Results: We enrolled 209 patients with suspected COVID-19, of whom 77 patients tested positive for COVID-19 and were included. Subsequent hospitalization occurred in 22 of 77 (29%) patients. Resting home $\text{SpO}_2 < 92\%$ was associated with an increased likelihood of hospitalization compared to $\text{SpO}_2 \geq 92\%$ (relative risk $= 7.0$, $95\%$ confidence interval $= 3.4$ to $14.5$, $p < 0.0001$). Home $\text{SpO}_2 < 92\%$ was also associated with increased risk of intensive care unit admission, acute respiratory distress syndrome, and septic shock. In our cohort, 50% of patients who ended up hospitalized only returned to the ED for incidental finding of low home $\text{SpO}_2$ without worsening of symptoms. One-third (33%) of nonhospitalized patients stated that they would have returned to the ED if they did not have a pulse oximeter to reassure them at home.

Conclusions: This study found that home pulse oximetry monitoring identifies need for hospitalization in initially nonsevere COVID-19 patients when a cutoff of $\text{SpO}_2 < 92\%$ is used. Half of patients who ended up hospitalized had $\text{SpO}_2 < 92\%$ without worsening symptoms. Home $\text{SpO}_2$ monitoring also reduces unnecessary ED revisits.

From the 1Department of Emergency Medicine, Swedish Hospital Part of NorthShore University HealthSystem, Chicago, IL; the 2Department of Gastroenterology, Advocate Lutheran General Hospital, Park Ridge, IL; the Office of Research and Sponsored Programs, Midwestern University, Glendale, AZ; and the 3Department of Medicine, University at Buffalo Jacobs School of Medicine and Biomedical Sciences, Buffalo, NY. Received May 12, 2020; revision received May 19, 2020; accepted May 21, 2020.

The authors have no relevant financial information or potential conflicts of interest to disclose.

NCT identifier: 04373161

Author contributions: SS, KM, and CP conceived the study. SS and KM designed the trial; obtained institutional review board approval; registered the trial with clinicaltrials.gov; collected the data; interpreted results; and drafted, revised, and wrote the final manuscript. Please note co-first authorship of SS and KM given the equal contributions of these authors. AS designed and performed statistical analysis and wrote the statistics methodology. NG, SS, MK, and JC collected data, contributed to writing the manuscript, and helped with revising the manuscript. SS contributed to interpreting results and writing and revising the manuscript. CP contributed to designing the trial, obtained funding for the project, interpreting results, and writing and revising the manuscript.

Supervising Editor: Jeffrey A. Kline, MD.

Address for correspondence and reprints: Kaushal Majmudar, DO; e-mail: ksmajmudar@gmail.com.

ACADEMIC EMERGENCY MEDICINE 2020;27:681–692

© 2020 by the Society for Academic Emergency Medicine

doi: 10.1111/acem.14053

ISSN 1553-2712
Background

In December 2019, a novel coronavirus called SARS-CoV-2 appeared in Wuhan city, Hubei Province, China, and rapidly spread across the rest of the world. This virus causes a disease known as COVID-19. Most patients with this infection recover after experiencing mild flu-like symptoms, but 20% of patients clinically deteriorate, requiring hospitalization and critical care. This deterioration can be quite rapid at times, resulting in patients requiring intubation and other advanced life support measures before or at arrival to the hospital.

One of the challenges of the COVID-19 pandemic in the United States is the strain it is placing on health care resources. Drastic measures have been taken to rapidly increase health care resources and reallocate health care workers to meet the needs during the pandemic. Given the severity of the ongoing global pandemic, the ability to remotely monitor patients who do not require hospitalization is essential for optimal utilization of health care resources.

Importance

A reasonable concern brought forward by emergency medicine physicians discharging initially nonsevere patients with COVID-19 is that these patients could potentially decompensate at home after discharge. Home pulse oximetry has been proposed as a way to monitor disease progression in such patients. However, there are currently no data to guide the use of home pulse oximetry in COVID-19 patients or its validity in identifying disease progression. Additionally, while it is generally known that patients with advanced age, comorbidities, or certain laboratory findings are at increased risk for worse clinical outcomes, specific predictors for who will require hospitalization are not known at this time.

Goal of Investigation

Our objective was to evaluate patient-reported oxygen saturation (SpO\textsubscript{2}) using pulse oximetry as a home monitoring tool for patients with initially nonsevere COVID-19 to identify need for hospitalization.

METHODS

Study Design and Setting

This prospective, uncontrolled open-label study took place at Swedish Hospital, part of NorthShore University Health System in Chicago, Illinois, between March 20 and April 22, 2020. The institutional review board approved the study and all patients consented to participate in the study. This study was registered with ClinicalTrials.gov (NCT04373161).

Study Population

All patients were older than 18 years of age. Patients were enrolled if they had suspected COVID-19 as defined by the World Health Organization (WHO). Testing for COVID-19 was performed using reverse transcriptase–polymerase chain reaction (RT-PCR) of an oropharyngeal or nasopharyngeal swab. Patient testing locations included the emergency department (ED) or Swedish Hospital–affiliated testing centers, including outpatient and employee testing sites for symptomatic individuals. For patients seen in the ED, only those being discharged to home were included. All patients had resting SpO\textsubscript{2} ≥ 92% on discharge from the ED. Patients being admitted to the hospital or discharged to a nursing facility were excluded. Other exclusion criteria included pregnancy and home oxygen use. Patients were not included if they were unable to be reached after enrollment.

Not all patients with suspected COVID-19 were tested due to ongoing test kit shortages during the time of this study. Only patients with positive COVID-19 testing were included in our outcome measures and analysis. Patients with suspected COVID-19 who did not undergo initial testing were still enrolled in case they were tested at a later time. ED physicians were not blinded to potential patient enrollment, but they were not specifically made aware of which patients were being enrolled into the study or if patients were already enrolled upon return to the ED.

Study Protocol

Upon discharge to home from the ED or testing site, patients were provided with an FDA-approved finger-tip pulse oximeter (EAD, Concord Health Supply, Skokie, IL) at no cost to the patient. Patients had their resting SpO\textsubscript{2} checked using this pulse oximeter at time of enrollment and this measurement was recorded as day 0. For 7 days, patients checked their SpO\textsubscript{2} using this pulse oximeter at 6:00 AM, 2:00 PM, and 10:00 PM. Seven-day follow-up was selected given the duration from symptom onset to hospitalization has been reported as 4 days (interquartile ratio [IQR] = 2–7 days). Investigators on the research study team called patients daily to collect data in real time.
In the study protocol provided to patients, they were instructed to return to the ED if: 1) their resting \( \text{SpO}_2 \) dropped below 92% and was confirmed with a separate reading 10 minutes later or 2) they felt they needed emergent medical attention. During these calls, patients were also surveyed on whether use of home pulse oximetry prevented further ED visits. The standardized script used for patient calls is available in Data Supplement S1, Method S1, available as supporting information in the online version of this paper, which is available at http://onlinelibrary.wiley.com/doi/10.1111/acem.14053/full. After the home pulse oximeter monitoring period, patients returned the pulse oximeter along with a standardized form detailing their measurements. The decision to hospitalize on subsequent return to the ED was left to the discretion of the ED physician evaluating the patient, independent of this study.

**Measurements**

Patients’ charts were reviewed to identify prior medical problems, laboratory values on preliminary ED visit, laboratory values on subsequent return to the ED or hospitalization, and outcomes of hospitalization. Obesity was defined as body mass index (BMI) \( \geq 30 \, \text{kg/m}^2 \) and lymphopenia was defined as lymphocyte count \(< 1.5 \times 10^9 \, \text{cells/L.} \)

**Outcomes**

The primary outcome was hospitalization in patients with resting home \( \text{SpO}_2 \) below 92%. Other outcomes measured included trend in resting home pulse oximetry readings, timing of \( \text{SpO}_2 < 92\% \), whether home pulse oximeter use decreased subsequent ED visits, and outcomes of hospitalization such as length of stay and transfer to the intensive care unit (ICU). We also measured time to drop (TTD), defined as time from symptom onset to \( \text{SpO}_2 < 92\% \), to see whether this predicted admission to the ICU, development of acute respiratory distress syndrome (ARDS), septic shock, or mortality. Finally, we collected data on demographics, past medical history, and laboratory values.

**Data Analysis**

The relative risk (RR) of hospitalization for COVID-19–positive patients with resting home \( \text{SpO}_2 \) below 92% was calculated, with \( p \)-value and associated 95% confidence interval (CI) determined using the Wald method. An a priori power analysis indicated a sample size of 76 to provide 80% power to detect a relative risk of 2.75 between hospitalizations and resting home \( \text{SpO}_2 \) below 92%. Differences in \( \text{SpO}_2 \) trends by time of day were compared with a linear mixed-effects model with an unstructured covariance matrix. The covariates considered were time of day and hospitalizations with a patient-specific intercept specified as a random effect. Differences between laboratory values for patients with both initial visit measurements and measurements at hospitalization were analyzed with a Wilcoxon signed rank test. We ran univariate logistic regression to identify predictors of ICU admission, development of ARDS, septic shock, or mortality. We considered running multivariate analysis but given the small sample size of our study, this was not considered to be statistically relevant and was not included. Statistical significance was set at the 0.05 level and analysis was performed using R version 3.6.2.

**RESULTS**

**Characteristics of Study Subjects**

A total of 209 patients with suspected COVID-19 were enrolled in our study. Of patients enrolled, 119 (57%) underwent RT-PCR testing and 79 (38%) tested positive for COVID-19. Patients who tested negative, withdrew consent, or were unable to be contacted after enrollment were excluded. A total of 77 COVID-19–positive patients were ultimately included and analyzed in our study (Figure 1). Of these 77 patients, nine patients were not initially tested on enrollment but tested positive at a subsequent ED visit. Enrollment locations included 61 (79%) patients enrolled from the emergency department, nine (12%) from employee testing, and seven (9%) from the outpatient testing center.

Demographic and baseline characteristics in COVID-19–positive patients are summarized in Table 1. Median (IQR) age was 44 (25-63) days, 43 (56%) were male, and median (IQR) BMI was 29.7 (21.8-37.6) \( \text{mg/kg}^2 \). Patients were Hispanic (57%), Asian (27%), African American (8%), and Caucasian (8%). In our cohort, 20 (26%) were health care workers. There were 32 (42%) patients with no medical problems, 20 (25%) with one medical comorbidity, 11 (14%) with two comorbidities, and 14 (18%) with three or more comorbidities. The most common medical comorbidities were obesity (27%), hypertension (26%), diabetes (16%), hyperlipidemia (13%), and asthma (9%). There were 10 (13%) patients on ACE inhibitor or angiotensin II receptor blockers.
Baseline laboratory values in patients at time of enrollment and subsequent laboratory values for hospitalized patients are summarized in Table 2. Patients had lymphopenia and elevated lactate dehydrogenase, C-reactive protein, liver enzymes, ferritin, and D-dimer on initial visit to the ED and upon hospitalization. Not all patients had laboratory studies drawn on enrollment as the decision to do so was left to the evaluating provider independent of this study. Laboratory values on day of admission to the hospital were not available for six patients because they were hospitalized at other institutions.

**Main Results**

There were 19 of 77 patients (25%) with home SpO₂ < 92%. Of these, 17 came back to the ED and 16 were hospitalized. Remarkably, eight of these 16 patients (50%) only returned to the ED for incidental finding of low home SpO₂ without worsening symptoms. The single patient with SpO₂ < 92% who returned to the ED and was not hospitalized had an SpO₂ of 94% in the ED and was discharged to home. Of the 58 patients who maintained SpO₂ ≥ 92%, 11 (19%) returned to the ED, where five patients were discharged and six patients were hospitalized.
Resting home SpO₂ < 92% was strongly associated with hospitalization compared to home SpO₂ ≥ 92% (RR = 7.0, 95% CI = 3.4 to 14.5, p < 0.0001; Figure 3). Symptoms were present for a median (IQR) of 5 (1-9) days prior to enrollment and 6 (4-8) days prior to hospitalization. The median (IQR) length of stay for hospitalized patients was 8 (2-14) days. Of hospitalized patients, eight (36%) were transferred to the ICU. Within the ICU cohort, six of eight (75%) patients had home SpO₂ < 92% and two of eight (25%) had home SpO₂ ≥ 92%. Of this ICU cohort, four of eight (50%) only came to the ED for incidental finding of low home pulse oximetry readings. Both patients within the ICU cohort with home SpO₂ ≥ 92% had downtrending SpO₂ with last reported reading of 93% prior to hospitalization. While in the ICU, seven patients developed ARDS requiring mechanical ventilation and six patients developed septic shock requiring vasopressors. There were two patients who died in the ICU. Resting home SpO₂ < 92% was associated with increased risk of ICU admission (RR = 9.8, 95% CI = 2.2 to 44.6, p < 0.002), ARDS (RR = 8.2, 95% CI = 1.7 to 38.7, p < 0.007), and septic shock (RR = 6.6, 95% CI = 1.3 to 32.9, p = 0.02). Resting home SpO₂ < 92% was not associated with increased mortality (p = 0.5). There were five (23%) patients still hospitalized at the time of data censoring.

There was no specific time of day that had higher likelihood of SpO₂ < 92% (p = 0.09). Presented in Figure 4 are longitudinal home pulse oximetry readings in patients who ended up hospitalized and patients who were not hospitalized. All hospitalizations occurred within 5 days of enrollment. The median

---

### Table 1

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Patients (N = 77)</th>
<th>Hospitalized Patients (n = 22)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Median (IQR) age (yr)</td>
<td>44 (19)</td>
<td>49 (19)</td>
</tr>
<tr>
<td>Male sex</td>
<td>43 (56)</td>
<td>16 (73)</td>
</tr>
<tr>
<td>Ethnicity†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>44 (57)</td>
<td>16 (73)</td>
</tr>
<tr>
<td>Asian</td>
<td>21 (27)</td>
<td>5 (23)</td>
</tr>
<tr>
<td>Caucasian</td>
<td>6 (8)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>African American</td>
<td>6 (8)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Health-care worker‡</td>
<td>20 (26)</td>
<td>4 (18)</td>
</tr>
<tr>
<td>Median (IQR) BMI (%)</td>
<td>29.7 (7.9)</td>
<td>30.1 (7.8)</td>
</tr>
<tr>
<td>Obesity§</td>
<td>21 (27)</td>
<td>9 (43)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>20 (26)</td>
<td>6 (27)</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>12 (16)</td>
<td>5 (23)</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>10 (13)</td>
<td>4 (18)</td>
</tr>
<tr>
<td>ACEI or ARB use</td>
<td>10 (13)</td>
<td>4 (18)</td>
</tr>
<tr>
<td>Asthma</td>
<td>7 (9)</td>
<td>3 (14)</td>
</tr>
<tr>
<td>Deep venous thromboembolism/</td>
<td>3 (4)</td>
<td>2 (9)</td>
</tr>
<tr>
<td>pulmonary embolism</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>3 (4)</td>
<td>2 (9)</td>
</tr>
<tr>
<td>Human immunodeficiency virus</td>
<td>3 (4)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Chronic kidney disease</td>
<td>2 (7)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Chronic obstructive pulmonary</td>
<td>2 (7)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>disease</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heart failure</td>
<td>2 (7)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Autoimmune disease</td>
<td>1 (1)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>History of malignancy</td>
<td>1 (3)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Hepatitis B virus</td>
<td>1 (1)</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Other†</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

---

Data are reported as n (%) unless otherwise reported. ARB = angiotensin II receptor blocker; ACEI = angiotensin-converting enzyme inhibitor; BMI = body mass index; IQR = interquartile range.

†The above characteristics are based on self-reported information and chart review of all patients who underwent confirmatory testing for COVID-19 represented by either IQR or nominal value.

‡Ethnicity determined by patient or family member report.

§Health care worker status determined by patient report.

¶Other comorbidities include cerebrovascular accident, cirrhosis, active malignancy, and hepatitis C virus.

---

### Table 2

<table>
<thead>
<tr>
<th>Laboratory Variable</th>
<th>Initial Visit (n = 28)</th>
<th>Day of Hospital Admission (n = 16)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hemoglobin (g/dL)</td>
<td>14.1 ± 1.8</td>
<td>13.9 ± 1.8</td>
</tr>
<tr>
<td>White cell count (×10⁹/L)</td>
<td>6.6 ± 2.8</td>
<td>6.4 ± 2.1</td>
</tr>
<tr>
<td>Lymphocyte count (×10⁹ cells/L)</td>
<td>1,226 ± 562</td>
<td>1,206 ± 764</td>
</tr>
<tr>
<td>Neutrophil count (×10⁹ cells/L)</td>
<td>4,754 ± 2,844</td>
<td>4,860 ± 2,004</td>
</tr>
<tr>
<td>Platelet Count (×10⁹/L)</td>
<td>226 ± 61</td>
<td>229 ± 77</td>
</tr>
<tr>
<td>Blood urea nitrogen (mg/dL)</td>
<td>14 ± 11</td>
<td>18 ± 21</td>
</tr>
<tr>
<td>Creatinine (mg/dL)</td>
<td>1.3 ± 2.0</td>
<td>2.0 ± 3.0</td>
</tr>
<tr>
<td>Albumin (g/dL)</td>
<td>4.5 ± 0.3</td>
<td>4.3 ± 0.3</td>
</tr>
<tr>
<td>AST (U/L)</td>
<td>40 ± 21</td>
<td>73 ± 79</td>
</tr>
<tr>
<td>ALT (U/L)</td>
<td>47 ± 31</td>
<td>75 ± 56</td>
</tr>
<tr>
<td>Total bilirubin (mg/dL)</td>
<td>0.6 ± 0.4</td>
<td>0.6 ± 0.2</td>
</tr>
<tr>
<td>C-reactive protein (mg/L)</td>
<td>70 ± 78</td>
<td>103 ± 81</td>
</tr>
<tr>
<td>Lactate dehydrogenase (U/L)</td>
<td>267 ± 68</td>
<td>430 ± 200</td>
</tr>
<tr>
<td>Ferritin (ng/mL)</td>
<td>516 ± 323</td>
<td>1,097 ± 1,273</td>
</tr>
<tr>
<td>Creatine kinase (U/L)</td>
<td>164 ± 135</td>
<td>174 ± 134</td>
</tr>
<tr>
<td>Troponin (ng/dl)</td>
<td>&lt;0.03 ± 0</td>
<td>0.04 ± 0.1</td>
</tr>
<tr>
<td>D-dimer (µg/mL)</td>
<td>0.3 ± 0.3</td>
<td>1.0 ± 0.9</td>
</tr>
<tr>
<td>Procalcitonin (ng/mL)</td>
<td>0.2 ± 0.6</td>
<td>0.4 ± 1.0</td>
</tr>
</tbody>
</table>

*Data are reported as mean ± SD. Laboratory values not available on all patients on initial visit due to enrollment in non-ED locations or due to no laboratory studies ordered by ED provider. Laboratory values not available on all patients on day of admission to the hospital if they were hospitalized at another institution.
(IQR) TTD was 6 (4-8) days. TTD was not associated with ICU admission (p = 0.3), ARDS (p = 0.5), septic shock (p = 0.7), or mortality (p= 0.7).

Trending laboratory values in patients who ended up hospitalized demonstrated significant increase in lactate dehydrogenase (p = 0.03) from initial ED visit to return ED visit for hospitalization (see Table 3). Of COVID-19–positive patients who did not return to the ED, 16 of 49 (33%) stated that they would have returned to the ED if they did not have the pulse oximeter to reassure them at home.

Univariate logistic regression found that lower initial pulse oximetry reading was associated with increased odds of hospitalization (odds ratio [OR] = 1.7, 95% CI 1.2 to 2.4, p < 0.004; see Table 4). Although lower platelet count (OR = 0.98, 95% CI = 0.96 to 0.99, p = 0.03) and lower albumin levels (OR = 0.5, 95% CI = 0.26 to 0.83, p = 0.03) were associated with hospitalization, the median levels were within the normal range. Asthma (OR = 9.5, 95% CI = 1.53 to 56.8, p = 0.01) and albumin (OR = 0.6, 95% CI = 0.35 to 0.91, p = 0.03) were associated with a composite outcome of ICU admission, ARDS, and septic shock (Table 5).

Demographic data and prior medical history in patients with suspected COVID-19 who did not undergo testing are summarized in Data Supplement S1, Table S1. Initial laboratory values on enrollment in this cohort are summarized in Data Supplement S1, Table S2. Longitudinal home pulse oximetry readings in these patients are presented in Figure S1.

**DISCUSSION**

In this study, we assessed the utility of home pulse oximetry monitoring in patients with initially
nonsevere COVID-19. Our study was designed to be a practical approach to monitor suspected and confirmed COVID-19 patients remotely and reduce in-person health care utilization. Our results found that pulse oximetry as a home monitoring tool identifies need for hospitalization in initially nonsevere COVID-19 patients when a cutoff of $\text{SpO}_2$ 92% is used.

We selected $\text{SpO}_2 < 92\%$, a measure of peripheral $\text{SpO}_2$, because this indicates the presence of hypoxemia, a measure of oxygen pressure in arterial blood ($\text{PaO}_2$). A

---

**Figure 4.** Longitudinal home pulse oximeter readings. (A) Home $\text{SpO}_2$ readings plotted over time at 6:00 AM, 2:00 PM, and 10:00 PM in COVID-19–positive patients who ended up hospitalized. Most patients had sudden drop below 92% in $\text{SpO}_2$ readings rather than a gradual decline. (B) Home $\text{SpO}_2$ readings plotted over time at 6:00 AM, 2:00 PM, and 10:00 PM in COVID-19–positive patients who were not hospitalized. $\text{SpO}_2$ = oxygen saturation. [Color figure can be viewed at wileyonlinelibrary.com]
recent multicenter, prospective study found SpO₂ < 92% had 95% sensitivity and 90% specificity for detecting PaO₂ < 60 mm Hg.³ PaO₂ < 60 mm Hg defines hypoxic respiratory failure.⁴ On the oxygen-dissociation curve, there is a steep drop in PaO₂ as PaO₂ approaches 60 mm Hg known as the “slippery slope.” Below this level, small reductions in PaO₂ correlate with disproportionately large reductions in SpO₂ and thereby oxygen delivery.⁵ In a cohort study of 2,923 patients seen in the ED with pneumonia, hospitalizing patients for SpO₂ < 92% was associated with improved mortality compared to hospitalizing patients with SpO₂ < 90%.⁶ These data support an intervention using SpO₂ < 92% as the cutoff to identify patients who may clinically deteriorate.

Over half of hospitalized patients in our cohort presented to the ED due to an incidental finding of low home SpO₂ without change in symptoms. A similar pattern has emerged recently whereby hypoxemia precedes severe symptoms in some patients with COVID-19, termed “silent hypoxemia.”⁷ Pathophysiology to explain this phenomenon is still being debated. Histologic evaluation on autopsy in a COVID-19–positive patient demonstrated diffuse alveolar damage, pulmonary edema, lymphocytic inflammatory infiltrate, and hyaline membrane formation, consistent with ARDS.⁸ A recent publication suggests that while ARDS is present in COVID-19, there appears to be heterogeneity in clinical presentation suggesting two disease phenotypes. They propose a varying combination of increasing inflammation and edema from patient self-inflicted lung injury related to increased negative intrathoracic pressure against the otherwise compliant lung.⁹ The use of supplemental oxygen improves hypoxemia and decreases work of breathing, which may reduce the risk of lung injury. It is plausible that outcomes could be improved with early intervention. Based on our findings, home pulse oximetry may identify these silent hypoxemia patients in the outpatient setting prior to onset of severe symptoms and respiratory failure. A randomized controlled trial of pulse oximetry in the patient population that we studied will be required to test that hypothesis.

In our cohort, most patients who had SpO₂ < 92% experienced an abrupt drop in SpO₂ rather than a gradual decline. This is consistent with emerging findings of certain patients rapidly deteriorating within a matter of hours.¹⁰ The underlying physiology for this sudden change in clinical status is attributed to a surge in proinflammatory molecules including IL-1β, IL-6, CCL-2, CCL-3, CCL-5, and TNF and has been termed the “cytokine storm” phase of COVID-19.¹¹ It is plausible that cytokine storm contributes to this drop in SpO₂.

Lactate dehydrogenase increased in patients who had labs drawn on enrollment and then were subsequently

---

Table 3

<table>
<thead>
<tr>
<th>Laboratory Variable</th>
<th>On Enrollment (n = 11)</th>
<th>Hospitalization (n = 11)</th>
<th>Paired p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>White cell count (x 10⁶/L)</td>
<td>5.8 (3.3–8.3)</td>
<td>5.7 (3.4–8.0)</td>
<td>0.742</td>
</tr>
<tr>
<td>Lymphocyte count (x 10⁶/L)</td>
<td>875 (579–1,171)</td>
<td>718 (230–1,206)</td>
<td>0.547</td>
</tr>
<tr>
<td>Neutrophil count (x 10⁶/L)</td>
<td>4,333 (1,876–6,790)</td>
<td>4,387 (2,423–6,351)</td>
<td>0.641</td>
</tr>
<tr>
<td>Hemoglobin (g/dl)</td>
<td>13.9 (10.7–17.1)</td>
<td>13.7 (9.7–17.7)</td>
<td>0.310</td>
</tr>
<tr>
<td>Platelet count (x 10⁹/L)</td>
<td>284,000 (131,000–437,000)</td>
<td>196,000 (118,000–274,000)</td>
<td>0.233</td>
</tr>
<tr>
<td>Blood urea nitrogen (mg/dl)</td>
<td>12 (3–21)</td>
<td>12 (5–19)</td>
<td>0.999</td>
</tr>
<tr>
<td>Creatinine (mg/dl)</td>
<td>0.8 (0.5–1.1)</td>
<td>0.8 (0.5–1.1)</td>
<td>0.999</td>
</tr>
<tr>
<td>AST (U/L)</td>
<td>46 (22–70)</td>
<td>41 (19–63)</td>
<td>0.400</td>
</tr>
<tr>
<td>ALT (U/L)</td>
<td>53 (14–92)</td>
<td>57 (11–103)</td>
<td>0.674</td>
</tr>
<tr>
<td>Total bilirubin (mg/dl)</td>
<td>0.7 (0.4–1.0)</td>
<td>0.8 (0.5–1.1)</td>
<td>0.462</td>
</tr>
<tr>
<td>Albumin (g/dl)</td>
<td>4.4 (4.2–4.6)</td>
<td>4.1 (3.7–4.5)</td>
<td>0.075</td>
</tr>
<tr>
<td>C-reactive protein (mg/L)</td>
<td>30 (16–44)</td>
<td>63 (24–102)</td>
<td>0.125</td>
</tr>
<tr>
<td>Lactate dehydrogenase (U/L)</td>
<td>291 (194–388)</td>
<td>379 (307–451)</td>
<td>0.031</td>
</tr>
<tr>
<td>Creatine kinase (U/L)</td>
<td>103 (63–143)</td>
<td>117 (68–166)</td>
<td>0.313</td>
</tr>
<tr>
<td>D-dimer (µg/ml)</td>
<td>0.3 (0.1–0.5)</td>
<td>0.3 (0.2–0.4)</td>
<td>0.999</td>
</tr>
<tr>
<td>Procalcitonin (ng/ml)</td>
<td>0.2 (0.0–1.0)</td>
<td>0.2 (0.0–1.3)</td>
<td>0.625</td>
</tr>
</tbody>
</table>

The values are bolded as they are the statistically significant values (p < 0.05) and hence are made to be easier for readers to find.

*Plus-minus values are median (IQR). Laboratory values not available on all patients on initial visit due to enrollment in non-emergency department locations, or due to no laboratory studies ordered by emergency department provider. Laboratory values not available on all patients on day of admission to the hospital if they were hospitalized at another institution. Hence data is available for 11 out of 22 patients who ended up hospitalized.
asthma to be associated with ICU admission, ARDS, and septic shock in our cohort.\textsuperscript{12,13} There are proposed mechanisms to account for a potential increased risk of severe disease in some patients with asthma including increased expression of angiotensin-converting enzyme 2 and transmembrane protease serine 2. Further investigation into the outcomes of asthma patients with COVID-19 will be needed to better risk stratify these patients.

Our patient cohort differs in several characteristics compared to other published studies. Most studies evaluate the hospitalized COVID-19 population, which is composed of patients who are more likely to be older and have more comorbid disease.\textsuperscript{14,15} In contrast, our patient population was younger and almost half had no chronic medical problems. Additionally, while our hospital serves a community that is 72% non-Hispanic, our hospitalized cohort was predominantly Hispanic. Despite this, Hispanic ethnicity did not emerge as a factor associated with hospitalization in our univariate analysis. It is unknown if our findings will translate similar to other patient populations.

This intervention was also successful in reassuring patients who may not require hospitalization, which in turn reduces ED utilization. This finding has two important benefits. Reducing ED utilization may reduce the risk of exposure to COVID-19 in healthcare workers in the ED. Additionally, this intervention may reduce unnecessary personal protective equipment (PPE) use. Globally, there is a PPE shortage including medical masks, respirators, gloves, gowns, and eye protection. The WHO has released guidelines that call for minimizing the need for PPE in healthcare settings given the global shortage.\textsuperscript{16} Our study found that providing home pulse oximeters to those with suspected or confirmed COVID-19 made patients feel more comfortable not returning to the ED as long as their SpO\textsubscript{2} remained appropriate.

Home pulse oximetry is made less accurate by nail polish, severe anemia, hyperbilirubinemia, hemoglobinopathies, or poor peripheral perfusion from severe vasoconstriction or poor cardiac output.\textsuperscript{17} While none of these conditions were present in our patients, it is important to note if applying to a larger patient population.

**LIMITATIONS**

Given that one-quarter of our COVID-19–positive patients were health care workers, it is possible that

### Table 4
Univariate Logic Regression of Factors Associated With Hospitalization in COVID-19 Patients

<table>
<thead>
<tr>
<th>Variable</th>
<th>OR (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.0 (0.99–1.08)</td>
<td>0.084</td>
</tr>
<tr>
<td>Male sex</td>
<td>2.8 (0.98–8.68)</td>
<td>0.064</td>
</tr>
<tr>
<td>BMI</td>
<td>1.1 (0.95–1.23)</td>
<td>0.2420</td>
</tr>
<tr>
<td>Lower SpO\textsubscript{2} at enrollment</td>
<td>1.7 (1.20–2.40)</td>
<td>0.004</td>
</tr>
<tr>
<td>Ethnicity*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>2.6 (0.91–8.07)</td>
<td>0.086</td>
</tr>
<tr>
<td>Asian</td>
<td>0.7 (0.21–2.18)</td>
<td>0.572</td>
</tr>
<tr>
<td>Health care worker†</td>
<td>0.5 (0.14–1.73)</td>
<td>0.328</td>
</tr>
<tr>
<td>Hypertension</td>
<td>1.1 (0.33–3.20)</td>
<td>0.90</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>1.8 (0.43–7.27)</td>
<td>0.385</td>
</tr>
<tr>
<td>Obesity‡</td>
<td>1.5 (0.49–4.58)</td>
<td>0.473</td>
</tr>
<tr>
<td>Diabetes</td>
<td>1.9 (0.51–6.90)</td>
<td>0.311</td>
</tr>
<tr>
<td>Asthma</td>
<td>1.9 (0.35–9.58)</td>
<td>0.415</td>
</tr>
<tr>
<td>ACEI or ARB use</td>
<td>1.7 (0.41–6.83)</td>
<td>0.430</td>
</tr>
<tr>
<td>White blood cell count</td>
<td>0.9 (0.58–1.17)</td>
<td>0.423</td>
</tr>
<tr>
<td>Lymphocyte count</td>
<td>1.0 (0.99–1.001)</td>
<td>0.066</td>
</tr>
<tr>
<td>Lymphopenia</td>
<td>8.9 (1.27–182.2)</td>
<td>0.058</td>
</tr>
<tr>
<td>Neutrophil count</td>
<td>1.0 (0.99–1.00)</td>
<td>0.958</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>1.0 (0.65–1.54)</td>
<td>0.992</td>
</tr>
<tr>
<td>Platelet count</td>
<td>0.98 (0.96–0.99)</td>
<td>0.032</td>
</tr>
<tr>
<td>Blood urea nitrogen</td>
<td>1.1 (0.99–1.32)</td>
<td>0.312</td>
</tr>
<tr>
<td>Creatinine</td>
<td>1.4 (0.85–2.26)</td>
<td>0.449</td>
</tr>
<tr>
<td>AST</td>
<td>1.0 (0.99–1.08)</td>
<td>0.149</td>
</tr>
<tr>
<td>ALT</td>
<td>1.0 (0.99–1.04)</td>
<td>0.255</td>
</tr>
<tr>
<td>Total bilirubin</td>
<td>0.8 (0.08–6.78)</td>
<td>0.839</td>
</tr>
<tr>
<td>Albumin</td>
<td>0.5 (0.26–0.83)</td>
<td>0.029</td>
</tr>
<tr>
<td>C–reactive protein</td>
<td>1.0 (0.98–1.01)</td>
<td>0.362</td>
</tr>
<tr>
<td>Lactate dehydrogenase</td>
<td>1.0 (0.98–1.01)</td>
<td>0.779</td>
</tr>
<tr>
<td>Ferritin</td>
<td>1.0 (0.99–1.02)</td>
<td>0.579</td>
</tr>
<tr>
<td>Creatinine kinase</td>
<td>1.0 (0.92–1.01)</td>
<td>0.104</td>
</tr>
<tr>
<td>D-dimer</td>
<td>0.007 (0–1.76)</td>
<td>0.191</td>
</tr>
</tbody>
</table>

The values are bolded as they are the statistically significant values (p < 0.05) and hence are made to be easier for readers to find. ACEI = angiotensin-converting enzyme inhibitor; ARB = angiotensin II receptor blocker; ALT = alanine transaminase; AST = aspartate aminotransferase; BMI = body mass index.

*Ethnicity determined by patient or family member report.

†Health care worker status determined by patient report.

‡Obesity determined by BMI ≥ 30 mg/kg².

hospitalized after returning to the ED. Our findings are concordant with recent data demonstrating elevated lactate dehydrogenase as a predictor of more severe COVID-19 disease.\textsuperscript{12} This laboratory value could be useful in assessing disease progression in COVID-19 patients who return to the ED. While platelet count and albumin were inversely associated with odds of hospitalization, the median levels were within the normal ranges, so these findings may not be clinically relevant.

While recent literature suggests a low prevalence of asthma in patients with severe COVID-19, we found...
our cohort was easier to train in using the home pulse oximeter and had better follow-up than the general population. Two patients withdrew from the study due to difficulty understanding how to use the pulse oximeter. Some patients could not be reached after enrollment. These occurrences emphasize the importance of patient selection and patient education when utilizing this intervention.

We standardized the home pulse oximeter used in our study to avoid variability between different brands. If multiple brands of pulse oximeters are used, the findings could be more heterogeneous with variability between home pulse oximeter readings. In a study of three different commercially available pulse oximeters, good correlation was observed for each of the finger pulse oximeters when compared to arterial blood gas samples in 94 patients. However, agreement may vary from device to device.

Patients were called once per day to collect data in real time. It is possible that these patient callbacks highlighted the importance of \( \text{SpO}_2 \) below 92%, which may have increased likelihood of patients returning to the ED. The use of home pulse oximetry monitoring may perform better when paired with some form of telemedicine.

Given the need to censor data to be shared, outcomes of patients may underrepresent ICU status, ARDS, septic shock, or mortality. Hospital length of stay is likely skewed lower as five patients remained hospitalized at time of data censoring. Additionally, our

Table 5
Univariate Logic Regression of Factors Associated With Composite Outcome of ICU Admission, ARDS, and Septic Shock in COVID-19 Patients*

<table>
<thead>
<tr>
<th>Variable</th>
<th>All COVID-19 Patients (N = 77) OR (95% CI) p-value</th>
<th>COVID-19 Patients Who Were Hospitalized (n = 22) OR (95% CI) p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.036 (0.98–1.10) 0.200</td>
<td>1.014 (0.95–1.09) 0.676</td>
</tr>
<tr>
<td>Male sex</td>
<td>1.360 (0.31–7.05) 0.689</td>
<td>0.455 (0.06–3.22) 0.420</td>
</tr>
<tr>
<td>BMI</td>
<td>1.028 (0.86–1.22) 0.753</td>
<td>0.971 (0.77–1.18) 0.775</td>
</tr>
<tr>
<td>( \text{SpO}_2 ) at enrollment†</td>
<td>0.771 (0.50–1.18) 0.222</td>
<td>1.185 (0.68–2.17) 0.553</td>
</tr>
<tr>
<td>Home ( \text{SpO}_2 &lt; 92% )</td>
<td>14.25 (2.90–105.8) \textbf{0.002}</td>
<td>1.667 (0.26–14.42) 0.605</td>
</tr>
<tr>
<td>Hypertension</td>
<td>0.926 (0.13–4.45) 0.929</td>
<td>0.833 (0.09–5.78) 0.857</td>
</tr>
<tr>
<td>Hyperlipidemia</td>
<td>1.074 (0.05–7.39) 0.950</td>
<td>0.611 (0.03–6.09) 0.696</td>
</tr>
<tr>
<td>Obesity‡</td>
<td>1.88 (0.40–8.89) 0.410</td>
<td>1.600 (0.27–10.01) 0.605</td>
</tr>
<tr>
<td>Diabetes</td>
<td>1.900 (0.25–9.71) 0.469</td>
<td>1.222 (0.13–9.56) 0.848</td>
</tr>
<tr>
<td>Asthma</td>
<td>9.450 (1.53–56.79) \textbf{0.012}</td>
<td>3.24 (0.0–35.0) 0.995</td>
</tr>
<tr>
<td>ACEI or ARB use</td>
<td>0.921 (0.05–6.12) 0.942</td>
<td>0.52 (0.02–5.09) 0.605</td>
</tr>
<tr>
<td>White blood cell count</td>
<td>0.748 (0.32–1.18) 0.373</td>
<td>0.439 (0.09–1.34) 0.201</td>
</tr>
<tr>
<td>Lymphocyte count</td>
<td>0.998 (0.99–1.00) 0.201</td>
<td>0.999 (0.99–1.00) 0.732</td>
</tr>
<tr>
<td>Neutrophil count</td>
<td>0.999 (0.99–1.00) 0.628</td>
<td>0.999 (0.99–1.00) 0.220</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>0.621 (0.31–1.08) 0.118</td>
<td>0.524 (0.19–1.03) 0.112</td>
</tr>
<tr>
<td>Platelet count</td>
<td>0.999 (0.99–1.00) 0.327</td>
<td>1.00 (0.99–1.01) 0.678</td>
</tr>
<tr>
<td>Blood urea nitrogen</td>
<td>1.10 (0.99–1.40) 0.288</td>
<td>1.071 (0.97–1.42) 0.455</td>
</tr>
<tr>
<td>AST</td>
<td>1.036 (0.99–1.09) 0.115</td>
<td>1.027 (0.97–1.11) 0.373</td>
</tr>
<tr>
<td>ALT</td>
<td>1.008 (0.98–1.04) 0.590</td>
<td>0.996 (0.95–1.04) 0.860</td>
</tr>
<tr>
<td>Total bilirubin</td>
<td>0.959 (0.04–11.76) 0.976</td>
<td>1.401 (0.01–173.4) 0.884</td>
</tr>
<tr>
<td>Albumin</td>
<td>0.0098 (0.0002–0.38) \textbf{0.042}</td>
<td>0.070 (0.0001–4.36) 0.266</td>
</tr>
<tr>
<td>C-reactive protein</td>
<td>0.988 (0.95–1.01) 0.346</td>
<td>0.987 (0.94–1.01) 0.432</td>
</tr>
<tr>
<td>Lactate dehydrogenase</td>
<td>0.997 (0.98–1.01) 0.697</td>
<td>0.997 (0.98–1.01) 0.697</td>
</tr>
<tr>
<td>Creatinine kinase</td>
<td>0.974 (0.93–1.01) 0.183</td>
<td>0.986 (0.93–1.04) 0.576</td>
</tr>
<tr>
<td>D-dimer</td>
<td>0.045 (0.0–9.46) 0.368</td>
<td>2.539 (0.73–14.64) 0.202</td>
</tr>
</tbody>
</table>

The values are bolded as they are the statistically significant values (p < 0.05) and hence are made to be easier for readers to find.

ACEI = angiotensin-converting enzyme inhibitor; ALT = alanine transaminase; ARB = angiotensin II receptor blocker; ARDS = acute respiratory distress syndrome; AST = aspartate aminotransferase; BMI = body mass index; ICU = intensive care unit.

*Laboratory values not available on all patients on initial visit due to enrollment in non-ED locations or due to no laboratory studies ordered by ED provider. Laboratory values not available on all patients on day of admission to the hospital if they were hospitalized at another institution. Hence, data are available for 11 of 22 patients who ended up hospitalized.

‡\( \text{SpO}_2 \) = home pulse oximeter oxygen saturation.

§Obesity determined by BMI ≥ 30 mg/kg².
study is a small sample size, and larger-scale studies need to be conducted to further investigate the utilization of home pulse oximetry monitoring to identify robust predictors of hospitalization. Such future studies should consider using known risk factors for poor outcomes in COVID-19 including age, sex, preexisting hypertension, diabetes, chronic lung disease, cardiovascular disease, low albumin, elevated C-reactive protein, and lymphopenia. Finally, we opted to exclude patients who tested negative for COVID-19; however, it should be noted that there is a significant false-negative rate with the current iteration of the RT-PCR test. There may be some utility to providing pulse oximeters to patients with high index of suspicion for COVID-19 who test negative; however, we did not investigate this.

**CONCLUSIONS**

This study found that home pulse oximetry monitoring identifies need for hospitalization in initially non-severe COVID-19 patients when resting home oxygen saturation drops below 92%. Half of patients who ended up hospitalized had oxygen saturation of less than 92% without worsening symptoms. Home pulse oximetry monitoring reduces ED utilization, which in turn reduces exposure risk to frontline health care workers and conserves personal protective equipment.

The authors acknowledge the Swedish Hospital Foundation for their donation of pulse oximeters, which made our study possible. We are grateful to Dr. Alan Shapiro and Dr. Alfonso Tafur for providing feedback on our findings and reviewing our manuscript. We appreciate the efforts of Dr. Agnieszka Bar and Dr. Mitchell D’Aloia for helping with data collection. We acknowledge Dr. Keri Robertson, Dr. Tommy Quoc Dang, and Dr. Kate Maxouris for helping us to enroll patients from the emergency department, outpatient testing center, and employee testing center.

**References**


16. World Health Organization. Rational Use of Personal Protective Equipment for Coronavirus Disease (COVID-19) and Considerations During Severe Shortages. 2020. Available at: https://www.who.int/publications/i/item/rational-use-of-personal-protective-equipment-for-coronavirus-disease-

Supporting Information
The following supporting information is available in the online version of this paper available at http://onlinelibrary.wiley.com/doi/10.1111/acem.14053/full
Data Supplement S1. Supplemental material.

E. Brooke Lerner, PhD¹, Craig D. Newgard, MD, MPH², and N. Clay Mann, PhD, MS, MBA³

ABSTRACT
Background: Our objective was to quantify trends in emergency medical services (EMS) incidents as the effects of the COVID-19 pandemic spread across the United States and to determine if there was an increase in EMS-attended deaths.

Methods: We conducted a 3-year comparative retrospective cohort analysis of data from the National EMS Information System. Data were included if care was provided between the 40th and 21st weeks of the next year and compared over 3 years. We included incidents identified through 9-1-1 where patient contact was made. The total number of EMS incidents per week was used as the denominator to calculate the rate of patient deaths and possible injury. We assessed for temporal and seasonal trends.

Results: Starting in the 10th week of 2020 there was a decrease in the number of EMS activations in the United States compared to the prior weeks and the same time period in previous years. The number of activations between week 10 and week 16 decreased by 140,292 or 26.1%. The portion of EMS activations reporting a patient disposition of death nearly doubled between the 11th and 15th weeks of 2020 (1.49%–2.77% of all activations). The number of EMS activations documenting a possible injury decreased from 18.43% to 15.27% between weeks 10 and 13.

Conclusion: We found that early in the COVID-19 outbreak there was a significant decrease in the number of EMS responses across the United States. Simultaneously the rate of EMS-attended death doubled, while the rate of injuries decreased.

The effects of the coronavirus disease 2019 (COVID-19) pandemic on different aspects of the U.S. healthcare system are evolving and emerging in the literature. These effects are presumed to be directly

From the, ¹Department of Emergency Medicine, University at Buffalo, Buffalo, NY, USA; the, ²Department of Emergency Medicine, Center for Policy and Research in Emergency Medicine, Oregon Health & Science University, Portland, OR, USA; and the, ³Department of Pediatrics, University of Utah School of Medicine, Salt Lake City, UT, USA.
Received May 13, 2020; revision received June 9, 2020; accepted June 10, 2020.
This project was supported by a contract with the National Highway Traffic Safety Administration, Office of Emergency Medical Services (DTNH22-15-C-00030). This project was supported by the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services (HHS) as part of an award totaling $2,799,974 with 0% financed with nongovernmental sources. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement by, HRSA, HHS, or the U.S. Government. The authors have no potential conflict of interest to disclose.

Author contributions: EBL—study concept and design, interpretation of the data, drafting of the manuscript, and acquisition of funding; CDN—study concept and design, interpretation of the data, and critical revision of the manuscript for important intellectual content; and NCM—study concept and design, acquisition of the data, analysis and interpretation of the data, critical revision of the manuscript for important intellectual content, and acquisition of funding.

Supervising Editor: Mark B. Mycyk, MD.
Address for correspondence and reprints: E. Brooke Lerner, PhD; e-mail: lerner@buffalo.edu.
or indirectly related to COVID-19, including massive efforts to slow the spread of the disease. Specifically, efforts were made to decrease how rapidly the virus infected the population to avoid placing too high a burden on the health care system, which could have resulted in an inability to meet the demand for care. This message was explained in the popular media as an effort to “flatten the curve.”

As communities began to implement social distancing interventions including “stay-at-home” orders and messaging the concept that we had to work together to reduce the spread of the virus to decrease the burden on the health care system, certain social patterns began to emerge. Patient visits to emergency departments (EDs) began dramatically trending downward.¹ The National Syndromic Surveillance Program found that from the 11th week of 2020 (March 9–15) to the 14th week (March 30–April 5) ED visits dropped from just over 2.5 million to 1.2 million.² Even in the region that includes New York, New Jersey, and Puerto Rico, ED visits went from 223,489 to 144,249 during that same time period.² Simultaneously, in New York, there were three deaths attributed to COVID-19 in week 11 and 3,194 in week 14, while in New Jersey the number of deaths was 2 in week 11 and 756 in week 14 (data on Puerto Rico were not available).³ Anecdotal, emergency physicians and other emergency care providers reported fewer patient visits, but higher patient acuity in the United States.⁴ These reports were supported in other countries where they too saw a drop in ED visits.⁵–⁷

Very limited information is available on the effect of the pandemic on the emergency medical services (EMS) system. In particular, it is unknown if there has been a similar decrease in patient encounters for prehospital care providers. Anecdotal reports suggest that EMS was responding to increasing numbers of out-of-hospital cardiac arrest cases in areas hard hit by the COVID-19 outbreak.⁸ This suggestion is concerning and corroborates reports of declining ED visits, since it may be a sign of the detrimental effects of citizens with emergent conditions not seeking timely emergency care and/or possibly a direct effect of COVID-19 infection. If confirmed, such a finding could be the result of the virus going undetected in such situations, as some communities have reported that medical examiners have limited access to testing for COVID-19 in deceased patients.⁹ If communities are not testing deceased patients, we may not know the full scope of the effect that the virus is having on our communities. The objective of this paper was to quantify the trends in national EMS incidents as the effects of the pandemic spread across the United States, using absolute numbers of EMS activations, activation rates, and types of EMS incidents. We also sought to determine if there was an increase in the number of EMS attended on-scene deaths, as has been reported in the popular media.¹⁰

**METHODS**

We conducted a 3-year comparative retrospective cohort analysis of data submitted to the National Emergency Medical Services Information System (NEMSIS) database. NEMSIS populates a National EMS registry including standardized patient care records (PCRs) submitted by upwards of 10,000 EMS agencies across 47 states and territories in near real time. Transmission of data to the national EMS registry is automated in most systems so that once a PCR is completed, the record populates the associated state and national registries automatically. This project was designated as being exempt from institutional review board review at the State University of New York at Buffalo in Buffalo, NY.

Patient care record data were included for study if patient care was provided by EMS providers between the 40th week of 2017 (October 2–8) and the 21st week of 2020 (May 18–24, 2020). The 40th week in a year was chosen as the study initiation period to preempt the beginning of the traditional flu season for any given year. For this study, time period 1 included PCRs from the 40th week of 2017 through the 21st week of 2018. Time period 2 included PCRs from the 40th week of 2018 through the 21st week of 2019. Time period 3 includes PCRs submitted from the 40th week of 2019 through the 21st week of 2020. These three study time periods allowed us to evaluate trends over time and to compare those trends to prior years to control for normal seasonal variation.

We analyzed all national EMS data that were available in the National EMS Database repository on June 3, 2020. Data were abstracted from the NEMSIS system at the NEMSIS Technical Assistance Center. EMS responses were included if the request for aid originated through the area’s emergency system (i.e., 9-1-1) and patient contact was documented. Patient transfers from one facility to another and nonemergent requests for private transport were excluded from the analysis.
The total number of EMS incidents per week were determined for each of the study time periods. We then used this number as the denominator to determine the rate (i.e., percentage) of specific EMS activation types per week. Documentation of patient deaths was based on a NEMSIS patient disposition variable (i.e., eDisposition.12 [Incident/Patient Disposition]) by combining patients who were identified as dead for whom resuscitation was either attempted or not and transportation was provided to a hospital or not. We included EMS responses for which a possible injury was documented as a concurrent comparison group, which is a common reason for EMS activations and, hypothetically, should substantiate our approach, by demonstrating an opposite effect, due to stay-at-home orders. Potential injury incidents were identified using the NEMSIS element eSituation.02 (Possible Injury).

**Data Analysis**

We analyzed data using descriptive statistics. In the National EMS Database repository, PCRs are provided by all EMS units responding to a request for service. Thus, if multiple units are dispatched to the same event, more than one PCR will be submitted to the national repository. We focused on week-to-week comparisons for a defined period over several years to assess temporal and seasonal trends. Confidence bands around weekly rates are not provided, since the large number of EMS activations associated with these analyses made them nearly indistinguishable from the reported value.

**RESULTS**

The overall study time frame included 37,550,949 9-1-1–initiated EMS activations resulting in patient contact. Time period 1 includes 8,621,423 EMS activations. Time period 2 includes 13,387,829 EMS activations. Time period 3 includes 15,541,697 EMS activations. Figure 1 illustrates that starting in the 10th week of 2020 (March 2–8) there was a precipitous decrease in the number of 9-1-1–initiated EMS activations in the United States compared to the prior weeks and the same time period in previous years. The weekly call volume decreased by 140,292 activations or (26.1%), comparing week 10 of 2020 to week 16.

The portion of EMS activations reporting a patient disposition of death at the scene remained fairly consistent until the 11th week of 2020 (March 9–15), at which point the proportion of scene deaths nearly doubled, increasing from 1.49% to 2.77% among all EMS activations with patient contact by week 15 (Figure 2). Examining the raw numbers of EMS-attended scene deaths reported in 2020, the number increases from 6,294 in week 11 to 8,942 in week 15. Conversely, the proportion of patients attended by EMS for which a potential injury is reported demonstrates an opposing trend for EMS activations, falling from 18.43% in week 10 of 2020 to 15.27% in week 13, with the actual number of EMS activations trending downward for this 3-week period from 98,487 to 66,593 (Figure 3).

**DISCUSSION**

In this study, we found that EMS activations initiated through the emergency response system and resulting in a patient contact declined rapidly since COVID-19 cases were first identified in the United States and social distancing measures were enacted. Further, there has been an increase in the percentage of EMS-attended scene deaths compared to prior weeks during similar time periods in previous years. This is in contrast to EMS activations reporting potential injuries, which decreased during the time frame representing the COVID-19 infection. These findings have both public health and economic implications for the U.S. emergency response system.

From a public health perspective, these findings suggest that individuals are not accessing the emergency medical system with the same frequency as experienced prior to the spread of COVID-19. While some of our findings could be explained by the lifestyle changes related to stay-at-home orders, such as driving less and participating in less risky recreational activities (i.e., fewer injuries), the decrease in EMS activations is likely not entirely explained by societal changes implemented in response to COVID-19. It is possible that changes in social perceptions (i.e., fear of infection) may explain our observed increase in the frequency of scene deaths attended by EMS. Recent publications have documented a decrease in the number of patients presenting to hospitals for acute coronary syndromes during the initial months of the pandemic in the United States, Spain, and Australia. Further, over a third of patients who delayed presenting for care of their myocardial infarction cited fear of COVID-19 or not wanting to burden the hospital as a reason for their delay. This phenomenon appears to be
affecting EDs and EMS systems in the United States, which could be a positive consequence if people whose medical needs do not require those services are seeking other avenues for care. However, our finding of a doubling in the rate of EMS attended deaths suggests that people who are experiencing medical emergencies are not accessing timely care. This conclusion is supported by a publication from Portugal that found an excess number of deaths that were not entirely explained by the reported fatalities due to COVID-19. As well as in Italy where they too saw an increase in cardiac arrests.

Changes in EMS call volume and case mix can have significant negative effects on EMS providers, especially if patient needs become more significant and urgent. More severely ill patients require EMS professionals to provide more technically complex care and increased exposure to high-stress situations related to patients not seeking early care for treatable conditions. These cases place additional stress and anxiety on EMS professionals, potentially resulting in long-term negative consequences to the health, well-being, and longevity of these important frontline responders.

It is also important to consider the economic implications of changing EMS volumes. EMS agencies must schedule units so that there is always additional capacity to respond to the next call. That is, if all EMS units are responding to individual patients at the same time there will be no capacity in the system for the next emergency. However, if too many units are idle then the community’s cost of maintaining the EMS system becomes too high to sustain the service since many agencies only earn revenue when they treat and transport patients. Most EMS agencies are required to have a certain number of units in service to meet contractual obligations or due to geographic factors related to response time requirements across

Figure 1. Comparison of the weekly number of 9-1-1-initiated EMS activations with patient contact from 2017 to 2020. The number of states submitting to the national EMS repository increased over the study period (2017, 32 states; 2018, 40 states; 2019, 44 states). The District of Columbia submitted PCRs in each time period. States enrolling in the National EMS repository commonly begin submitting PCRs at the beginning of the calendar year. No state stopped submitting PCRs once enrolled. PCR = patient care record.
the service area. A drop in call volume will result in agency costs that are likely flat, while agency revenues significantly decline, leading to EMS budgetary shortfalls that will be difficult to recover from. The ability of EMS to quickly and efficiently respond to future emergencies could be jeopardized.

While our reported findings may adjust over time, it will be important to consider how public health messaging regarding the potential burden on health care systems and fears about contracting a novel virus may affect community member’s decision to access care. In future outbreaks and other public health emergencies, it will be important to balance the need for people to seek and receive needed care with the requirement for communities to implement practices (such as social distancing) that are meant to address and contain an emerging threat.

LIMITATIONS

This study is limited by the expansion of states participating in NEMSIS during the time frame that was studied. During the study period, the number of states contributing data to the NEMSIS expanded from 32 to 44. States enrolling in the National EMS Database repository commonly begin submitting PCRs at the beginning of the calendar year (Figure 1). No states began submissions to the National EMS Database repository during the COVID-19 pandemic.

With that expansion, the expectation is that the number of EMS activations per week would increase from year to year. Instead, we found fewer activations from week 10 through week 17 of 2020, during the period of COVID-19 community spread compared to the weeks preceding it and the same weeks in the prior time periods. Another potential limitation is that the analysis is based on EMS unit activations rather than individual patients. In some cases, multiple records may have been completed for the same patient. However, this limitation applies to all years of PCRs submitted to NEMSIS and can therefore be considered consistent across the three time periods. It is unknown what affect the pandemic may have had on the number of units responding to a scene. We could
expect to see the practice of multiple responding units potentially decrease for responses during the pandemic to limit the number of responders and reduce exposure, as was recommended by the American Heart Association during week 15 of 2020. This would result in an opposite effect than that observed, but this assumption cannot be evaluated with the available data so there also may have been an increase in the number of units responding to each patient.

In some areas of the U.S. EMS data are submitted to NEMSIS with a defined lag period to promote data validity. This analysis was done in the 23rd week of 2020. It is possible that some data were not yet submitted, which may affect the activation rates reported for scene deaths and patients with potential injuries. However, the use of a rate (or proportion) should minimize the impact of this potential bias. There is no obvious reason why EMS responders would systematically withhold (or accelerate) the submission of specific types of EMS activations.

**CONCLUSION**

We found that early in the COVID-19 outbreak there was a significant decrease in the number of emergency medical services responses across the United States. Simultaneously the rate of emergency medical services–attended scene death doubled, while the rate of emergency medical service activations related to patient injury decreased.

The authors recognize the analytical support provided by Mengtao Dia and Chris Hoffman and thank all participating EMS clinicians, EMS agencies, and state EMS offices who provided data to the NEMSIS National Database and delivered patient care during this very difficult time in U.S. history.

**References**


Academic Emergency Medicine Physicians’ Anxiety Levels, Stressors, and Potential Stress Mitigation Measures During the Acceleration Phase of the COVID-19 Pandemic

Robert M. Rodriguez, MD1, Anthony J. Medak, MD2, Brigitte M. Baumann, MD3, Stephen Lim, MD4, Brian Chinnock, MD5, Remi Frazier, MS6, and Richelle J. Cooper, MD7

ABSTRACT

Objective: The objective was to assess anxiety and burnout levels, home life changes, and measures to relieve stress of U.S. academic emergency medicine (EM) physicians during the COVID-19 pandemic acceleration phase.

Methods: We sent a cross-sectional e-mail survey to all EM physicians at seven academic emergency departments. The survey incorporated items from validated stress scales and assessed perceptions and key elements in the following domains: numbers of suspected COVID-19 patients, availability of diagnostic testing, levels of home and workplace anxiety, severity of work burnout, identification of stressors, changes in home behaviors, and measures to decrease provider anxiety.

Results: A total of 426 (56.7%) EM physicians responded. On a scale of 1 to 7 (1 = not at all, 4 = somewhat, and 7 = extremely), the median (interquartile range) reported effect of the pandemic on both work and home stress levels was 5 (4–6). Reported levels of emotional exhaustion/burnout increased from a prepandemic median (IQR) of 3 (2–4) to since the pandemic started a median of 4 (3–6), with a difference in medians of 1.8 (95% confidence interval = 1.7 to 1.9). Most physicians (90.8%) reported changing their behavior toward family and friends, especially by decreasing signs of affection (76.8%). The most commonly cited measures cited to alleviate stress/anxiety were increasing personal protective equipment (PPE) availability, offering rapid COVID-19 testing at physician discretion, providing clearer communication about COVID-19 protocol changes, and assuring that physicians can take leave for care of family and self.

Conclusions: During the acceleration phase, the COVID-19 pandemic has induced substantial workplace and home anxiety in academic EM physicians, and their exposure during work has had a major impact on their home lives. Measures cited to decrease stress include enhanced availability of PPE, rapid turnaround testing at provider discretion, and clear communication about COVID-19 protocol changes.

From the 1Department of Emergency Medicine, University of California at San Francisco School of Medicine, San Francisco, CA, USA; the 2Department of Emergency Medicine, University of California at San Diego School of Medicine, San Diego, CA, USA; the 3Department of Emergency Medicine, Cooper Medical School of Rowan University, Camden, NJ, USA; the 4Section of Emergency Medicine, Louisiana State University Health Sciences Center, New Orleans, LA, USA; the 5Department of Emergency Medicine, UCSF-Fresno Medical Education Program, Fresno, CA, USA; 6Academic Research Systems, University of California, San Francisco, CA, USA; and the 7Department of Emergency Medicine, University of California at Los Angeles School of Medicine, Los Angeles, CA, USA.

Received May 18, 2020; revision received June 15, 2020; accepted June 18, 2020.

The authors have no relevant financial information or potential conflicts of interest to disclose.

Author contributions: RMR and RF contributed to study concept, design, and analysis. All authors contributed to data acquisition, interpretation of data, drafting, and revision of the manuscript.

Supervising Editor: Steven B. Bird, MD.

Address for correspondence and reprints: Robert M. Rodriguez, MD; e-mail: robert.rodriguez@ucsf.edu.


ISSN 1553-2712

© 2020 by the Society for Academic Emergency Medicine
doi: 10.1111/acem.14065
BACKGROUND AND IMPORTANCE

Although the effects of the COVID-19 pandemic on the public’s anxiety levels have been well documented by traditional media, the degree to which the pandemic affects physician stress and personal life has not yet been quantified in the United States.\(^1,2\) Investigators reported a heavy psychological toll on health care workers in Wuhan and other regions of China.\(^3\) Anticipating a surge in mental health care needs in U.S. health care workers, others have called for similar systematic assessments of frontline providers.\(^4,5\)

Goals of This Investigation

In mid-March 2020, we initiated a longitudinal survey study to assess multiple factors affecting the psychological health of emergency medicine (EM) physicians in the United States during the COVID-19 pandemic. In our study design, we seek to evaluate different topics that are relevant to three phases of the pandemic: the acceleration phase, the plateau/deceleration phase, and the resolution phase. Herein we report results of the first (acceleration) phase of this study to aid EM physicians and health care systems in their development of programs for stress mitigation in real time. Specifically, we sought to assess home and workplace anxiety, burnout, work-related stressors, changes to home life, and perceptions as to what measures might ease provider anxiety.

Study Design, Setting, and Selection of Participants

This was a cross-sectional survey administered via e-mail from February 23, 2020, to April 10, 2020, to all EM physicians (attending, fellow, and resident) at seven EM residencies and affiliated institutions: University of California San Francisco–UCSF (San Francisco, CA); UCSF–Fresno Medical Education Program (Fresno, CA); Cooper Medical School of Rowan University–CMSRU (Camden, NJ); University of California at Los Angeles (UCLA-Olive View program with affiliated West Los Angeles VA and Santa Monica UCLA Medical Center, Los Angeles, CA); Kern Medical Emergency Medicine Residency (Bakersfield, CA); Louisiana State University Health Science Center (New Orleans, LA); and University of California at San Diego–UCSD (San Diego, CA). Participating sites were primarily recruited through their involvement in the National Emergency X-radiography Utilization Study (NEXUS) network. To broaden the sampling to sites that were experiencing heavy surges of COVID-19 patients, we contacted investigators at two residencies in New York City (NYC) and one in New Orleans; investigators in NYC believed that their staff were too overloaded to meaningfully participate. We excluded non–clinically active physicians. This study was deemed exempt by the respective institutional review boards.

METHODS OF MEASUREMENT

Collaborating with the University of California Stress Network, we developed a survey instrument to assess perceptions and key elements about the following domains: provider estimates of numbers of patients treated with suspected COVID-19 infection, availability of COVID-19 diagnostic testing, home and workplace anxiety, work burnout, identification of work-related stressors, changes in behavior at home arising from their work during the pandemic, and perceptions as to what measures might decrease provider anxiety. Anticipating the difficulty with response rates to lengthy questionnaires during the acceleration phase of the pandemic, we sought to make our instrument pragmatic and succinct; we adapted selected questions from validated stress and burnout assessment scales that would address our particular domains of study.\(^6,7\) For example, to assess emotional exhaustion and burnout, participants were asked to rate on a 1 to 7 scale (1 = not at all, 4 = somewhat, and 7 = very much) “to what extent were you experiencing severe, ongoing job stress where you felt emotionally exhausted, burned out, cynical about your work and fatigued, even when you wake up?” To assess what measures might relieve anxiety related to their work during the pandemic, respondents were presented a list of 11 measures and asked to assign their top five measures (1 = highest priority and 5 = fifth highest priority) that they thought would alleviate some of their anxiety/stress. After pilot testing our preliminary instrument on five physicians to ensure understanding and a completion duration of <10 minutes, our final survey consisted of 32 Likert-type scale, yes/no, multiple-choice, and free-response questions. We sent repeat e-mail requests to nonresponders twice to increase response rate (Data Supplement S1, Table S1, available as supporting information in the online version of this paper, which is available at http://onlinelibrary.wiley.com/doi/10.1111/acem.14065/full).
Primary Data Analysis

Keeping responses anonymous, we managed survey data using REDCap hosted by the University of California at San Francisco. We used STATA v 15.1 for analyses, summarizing patient characteristics and key responses as raw counts, frequency percent, medians, and interquartile ranges (IQRs). We additionally stratified data and used the Wilcoxon rank-sum test for medians and difference (Δ) in proportions with 95% confidence intervals (CIs) for proportions to compare key question responses for the following subgroups: female versus male, faculty versus resident/fellow, children at home versus no children at home, and surge cities (New Orleans and Camden) versus nonsurge cities (California cities). For the question regarding measures to relieve stress, we created a rank summary of aggregate points. Each respondents’ highest priority measure was given 5 points, second given 4 points, and so forth, with the fifth given 1 point; noncited measures were given 0 points.

RESULTS

Characteristics of Participants

We sent the survey to 751 EM physicians and received 426 responses (56.7% response rate; Data Supplement S1, Appendix S1). The response rate among female EM physicians was higher than that from male EM physicians (60.4% vs. 51.9%, difference Δ = 8.5%, 95% CI = 1.4% to 15.5%). Response rates from faculty, fellows, and residents were 57.6, 42.4, and 51.4%, respectively (Table 1, respondent characteristics).

Main Results

Of the 419 (98.4%) respondents who reported patient contact from February 15, 2020, to their survey time, 410 (97.9%) reported seeing patients who they suspected had COVID-19 infections; the median number of patients they suspected had COVID-19 was 20 (IQR = 10–30). Respondents estimated that 40% (IQR = 10%–80%) of these suspected cases had received the swab test for COVID-19; 289 (67.8%) stated that they had a patient test positive and 89 (20.9%) were unsure. On the 1 to 7 scale, the median reported effect of the COVID-19 pandemic on work stress levels was 5 (IQR = 4–6) and on home stress levels was 5 (IQR = 4–6). With regard to emotional exhaustion and burnout, EM physicians reported a before-the-pandemic median of 3 (IQR = 2–4) and a since-the-pandemic-started median of 4 (IQR = 3–6; Δ = 0.8, 95% CI = 1.7 to 1.9). We found no significant differences in key question responses comparing faculty versus resident/fellow, children at home versus no children at home, and surge city versus nonsurge city. Female gender respondents reported a higher effect of the COVID-19 pandemic on work anxiety levels (6 vs. 5; median Δ = 1, IQR = 0–2) and on home anxiety levels (6 vs. 5; median Δ = 1, IQR = 0 to 2) than men (Table 2).

We asked EM physicians’ concerns regarding their work as health care providers during the pandemic. The primary concerns were worries about the adequacy of personal protective equipment (PPE), worries about the ability to accurately diagnose COVID-19 cases quickly, worries about the well-being of coworkers who have been diagnosed with COVID-19, and worries that patients with unclear diagnoses are exposing others in the community (Table 3).

Most EM physicians (81.7%) had discussed the risks of their excess exposure as health care workers during the pandemic. The primary concerns were worries about the adequacy of personal protective equipment (PPE), worries about the ability to accurately diagnose COVID-19 cases quickly, worries about the well-being of coworkers who have been diagnosed with COVID-19, and worries that patients with unclear diagnoses are exposing others in the community (Table 3).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Demographics (n = 426)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>35 (31–43)</td>
</tr>
<tr>
<td>Female</td>
<td>192 (45.1)</td>
</tr>
<tr>
<td>Physician training level</td>
<td></td>
</tr>
<tr>
<td>Faculty</td>
<td>236 (55.4)</td>
</tr>
<tr>
<td>Fellow</td>
<td>19 (4.5)</td>
</tr>
<tr>
<td>Resident</td>
<td>168 (39.4)</td>
</tr>
<tr>
<td>Race and ethnicity</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>14 (3.3)</td>
</tr>
<tr>
<td>Asian</td>
<td>69 (16.2)</td>
</tr>
<tr>
<td>Asian-Indian</td>
<td>3 (0.7)</td>
</tr>
<tr>
<td>Latinx</td>
<td>36 (8.5)</td>
</tr>
<tr>
<td>Middle Eastern</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>Native American</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>Pacific Islander</td>
<td>1 (0.2)</td>
</tr>
<tr>
<td>White</td>
<td>306 (71.8)</td>
</tr>
<tr>
<td>Home living situation</td>
<td></td>
</tr>
<tr>
<td>Alone</td>
<td>63 (14.8)</td>
</tr>
<tr>
<td>With roommate(s)</td>
<td>47 (11)</td>
</tr>
<tr>
<td>With partner(s)</td>
<td>308 (72.3)</td>
</tr>
<tr>
<td>With child &lt; 18 years</td>
<td>166 (39)</td>
</tr>
<tr>
<td>With adult &gt; 70 years</td>
<td>9 (2.1)</td>
</tr>
</tbody>
</table>

Data are reported as median (IQR) or n (%).
when asked how much they believed that friends and family were treating them differently as a result of their work-related potential exposure to COVID-19, with a median level of concern of 4 (IQR = 2–5). The most common reported changes by friends and family were expressions of concern about the EM physician participants’ health (65.3%), expressions of concern about their exposure to COVID-19 because of contact with the EM physician (42.3%), and a reluctance of family members to be in close contact with the EM physician (40.4%).

In Table 4, we present a ranked summary of responses of measures that would alleviate provider stress. The highest ranked measures to alleviate anxiety/stress related to the COVID-19 pandemic were enhanced availability of PPE, rapid COVID-19 testing with physician discretion, clear communication about changes in COVID-19 protocols, and assurance that physicians can take leave for care of family and self.

**DISCUSSION**

In this cross-sectional survey conducted during the acceleration phase of the COVID-19 pandemic, EM physicians in seven cities reported that the pandemic has induced moderate to severe levels of anxiety at work and at home. Their primary work concerns relate to COVID-19 exposure compromising their personal health, availability of adequate PPE, limited rapid diagnostic testing, and risks of community spread of discharged COVID-19 patients. Occupational exposure has changed the vast majority of
Table 3
Physicians’ Concerns Relating to Their Work During the COVID-19 Pandemic

<table>
<thead>
<tr>
<th>Concern</th>
<th>Median</th>
<th>IQRs</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPE is inadequate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>We are not able to accurately diagnose COVID-19 cases quickly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I may be secondarily exposing family members or others because of my work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients with unclear diagnoses are exposing others in the community</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I am being exposed at work and compromising my health</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well-being of coworkers who have been diagnosed with COVID</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I might have to undergo quarantine and will not be able to work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others at home or elsewhere are afraid to come in contact with me because I’m a health care provider</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I may have to quarantine at home and this will affect my family</td>
<td></td>
<td></td>
</tr>
<tr>
<td>We will not have enough staffing as coworkers are quarantined</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our ED, clinic, or hospital is not prepared enough for the pandemic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social isolation and not being able to do things outside of the home</td>
<td></td>
<td></td>
</tr>
<tr>
<td>We are having to send patients home without a clear diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I will not be able to get food and other necessities for me and my household</td>
<td></td>
<td></td>
</tr>
<tr>
<td>My home life will not be the same after resolution of this pandemic</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Median and IQRs to questions “I worry about or that ...” on 1 to 7 scale, in which 1 = “not at all,” 4 = “somewhat,” and 7 = “extremely.” PPE = personal protective equipment.

Table 4
Rank Summary of Measures That Emergency Physicians Believe Would Relieve Their Stress Related to the COVID-19 Pandemic

<table>
<thead>
<tr>
<th>Measure</th>
<th>Aggregate Points</th>
<th>No. (%) of Respondents Citing Measure (N = 426)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enhanced availability of PPE</td>
<td>1637</td>
<td>410 (96.2)</td>
</tr>
<tr>
<td>Rapid turnaround (&lt; 6 hours) testing</td>
<td>1362</td>
<td>392 (92.0)</td>
</tr>
<tr>
<td>Testing for COVID-19 for patients at my discretion (instead of as limited by current protocols)</td>
<td>1054</td>
<td>351 (82.4)</td>
</tr>
<tr>
<td>Clearer communication about changes in protocols</td>
<td>976</td>
<td>313 (73.5)</td>
</tr>
<tr>
<td>Assurances that I can take leave to care for myself and family members</td>
<td>933</td>
<td>306 (71.8)</td>
</tr>
<tr>
<td>Greater clarity regarding my risk for exposure</td>
<td>858</td>
<td>284 (66.7)</td>
</tr>
<tr>
<td>Assurances that my (and my dependents’) medical care will be covered by my employer</td>
<td>799</td>
<td>270 (63.4)</td>
</tr>
<tr>
<td>Ability to request testing of myself for COVID-19 even if I do not have symptoms</td>
<td>787</td>
<td>295 (69.2)</td>
</tr>
<tr>
<td>Assurances about disability benefits</td>
<td>741</td>
<td>243 (57.0)</td>
</tr>
<tr>
<td>Easily available mental health consultations for myself and other health care providers</td>
<td>660</td>
<td>242 (56.8)</td>
</tr>
<tr>
<td>Departmental ZOOM or other video sessions to discuss COVID-19 response and changes</td>
<td>638</td>
<td>236 (55.4)</td>
</tr>
</tbody>
</table>

Respondents were asked: “From the list below, pick the top 5 measures (1 = highest priority) that you think would alleviate some of your anxiety/stress related to the COVID-19 pandemic.” Aggregate Points are the sum of points in which 1 (highest priority) = 5 points, 2 = 4 points, 3 = 3 points, 4 = 2 points, 5 = 1 point. PPE = personal protective equipment.
physicians’ behavior at home, where they are worried about exposing family members and roommates, the possibility of needing to self-quarantine, and the effects of excess social isolation. Respondents’ highest ranked anxiety relief measures included improved access to PPE, rapid turnaround COVID-19 testing at provider discretion, clearer communications about COVID-19 protocol changes, assurances about leave, and ability to request self-testing.

Although several investigators have examined the effects of the COVID-19 pandemic on health care worker mental health in other countries, we were unable to find any similar studies of U.S. physicians. The moderate to severe levels of stress we found have not been consistently replicated in these other international studies. In a study of 906 health care providers in Singapore and India, with 30% physician enrollment, anxiety was documented in 15.7%, depression in 10.6%, and stress in 5.2% of all study participants. Lu et al. documented higher levels of moderate fear in high-risk (emergency, critical care, and infectious disease) health care workers at Fujian Provincial Hospital, when compared to low-risk medical and administrative staff. Our findings are most congruent with those of Lai et al., who found symptoms of depression (50.4%), anxiety (44.6%), insomnia (34.0%), and distress (71.5%) in frontline health care workers at 34 hospitals in China.

Similar to our findings, investigators in China, Italy, and Turkey have reported higher levels of anxiety and depression in female health care providers during the COVID-19 pandemic. While investigators in Turkey found that having a child was associated with lower anxiety and depression levels, we did not find a similar protective effect of parenthood or differences in any of the other factors that we examined.

It is important to note that respondents’ greatest concern and best anxiety relief measure both related to having adequate PPE. Investigators in China reported that lack of PPE was associated with higher levels of anxiety and depression. Although the availability of PPE has increased substantially over the course of the pandemic, the National Nurses United survey of 8,200 U.S. nurses conducted during the time of our study found that only 55% of nurses had access to N95 respirators on their units and only 24% believed that their employer had sufficient PPE stock for a rapid surge in COVID-19 patients.

Of note, this a longitudinal study with different goals in each of the three phases. In this first phase during the acceleration interval of the COVID-19 pandemic, we have quantified high levels of work and home life anxiety experienced by EM physicians in the United States, we have identified sources of this stress, and we have presented a number of anxiety mitigation measures. Although some of our findings may be intuitive, this work provides a critical early template for the design and implementation of interventions that will address the mental health needs of emergency physicians in the COVID-19 pandemic era. Most, if not all, of respondents’ measures to relieve stress are readily actionable items for emergency departments (EDs) and their parent institutions, and the central PPE concern is a fundamental workplace safety issue. As discussed by Wong et al., institutions should act expeditiously to address these root cause workplace stressors and consider programs to improve emotional resilience for EM physicians.

In terms of future directions of this work, our study design and survey instruments are fluid. As the pandemic has progressed, additional important stressors, such as childcare and homeschooling demands, the economic impact of declining ED volumes, and changes in health care delivery (lack of personal connections with patients because of limited time in rooms) have arisen. We plan to address these stressors, along with concerns about the development of long-term posttraumatic stress, in our subsequent follow-up surveys.

LIMITATIONS

Our primary limitation is the moderate response rate of 57%, which we attribute to general e-mail and clinical work overload during the frenetic early stage of the pandemic and inability to provide gift cards or other incentives in this unfunded study. Although waiting for funding and conducting the survey in a less chaotic time (after the pandemic acceleration phase) may have produced a higher response rate, this method would undoubtedly have introduced recall bias in terms of respondents’ self-assessment of anxiety levels and particular stressors. We believe that our survey provides accurate estimates of how the responding physicians were feeling in real time during the acceleration phase. Another limitation is that those who were experiencing more anxiety may have been more likely to respond to the survey request, thus leading to an overestimation of stress; however, it is also possible that those with more anxiety declined to participate.
In terms of spectrum effects, our survey was limited to providers at academic institutions and therefore may not reflect the experience of nonacademic EM physicians.

Additionally, most of our participant sites were in cities in California that had not yet seen large surges of patients as seen in other areas of the country. It is very likely that EM physicians in NYC and other "hot spots" for COVID-19 have been suffering higher levels of anxiety and effects on home life. Nevertheless, median levels of anxiety in the California sites were similar to those of the New Orleans and Camden sites, which were experiencing surges. This suggests that the impact of COVID-19 is pervasive and that measures to mitigate stress should be enacted universally.

CONCLUSIONS

The acceleration phase of the COVID-19 pandemic has induced moderate to severe workplace and home anxiety in academic emergency medicine physicians. The pandemic has had considerable impact on the home life of most physicians, especially in terms of decreased signs of affection and worries about exposing family members and friends to infection. Institutional measures, including enhanced availability of personal protective equipment, rapid turnaround testing at provider discretion, and clear communication about COVID-19 protocol changes, should be enacted to mitigate physician stress.

REFERENCES


Supporting Information

The following supporting information is available in the online version of this paper available at http://onlinelibrary.wiley.com/doi/10.1111/acem.14065/full

Data Supplement S1. Supplemental material.
Relationship Between Pain, Opioid Treatment, and Delirium in Older Emergency Department Patients

Raoul Daoust, MD, MSc1,2, Jean Paquet, PhD2, Valérie Boucher, MSc3,4, Mathieu Pelletier, MD5,6, Émilie Gouin, MD7, and Marcel Émond, MD, MSc3,4,5

ABSTRACT

Objectives: Emergency department (ED) stay and its associated conditions (immobility, inadequate hydration and nutrition, lack of stimulation) increase the risk of delirium in older patients. Poorly controlled pain and paradoxically opioid pain treatment have also been identified as triggers for delirium. The aim of this study was to assess the relationship between pain, opioid treatment, and delirium in older ED patients.

Methods: A multicenter prospective cohort study was conducted in four hospitals across the province of Québec (Canada). Patients aged ≥ 65 years old, waiting for hospital admission between March and July 2015, who were nondelirious upon ED arrival, who were independent or semi-independent in their daily living activities, and who had an ED stay of at least 8 hours were included. Delirium assessments were conducted twice a day during the patient’s entire ED stay and their first 24 hours on the hospital ward using the Confusion Assessment Method. Pain intensity was evaluated using a visual analog scale (VAS = 0–100) during the initial interview, and all opioid treatments were documented.

Results: A total of 338 patients were included; 51% were female, and mean (±SD) age was 77 (±8) years. Forty-one patients (12%) experienced delirium during their hospital stay occurring within a mean (±SD) delay of 47 (±19) hours after ED admission. Among patients with pain intensity ≥ 65 from VAS (0–100), 26% experienced delirium compared to 11% for patients with pain < 65 (p < 0.01), and no significant association was found between opioid consumption and delirium (p = 0.31). Logistic regression controlling for confounding factors showed that patients with pain intensity ≥ 65 are 3.3 (95% confidence interval = 1.4 to 7.9) times more likely to develop delirium than patients who had pain intensity of <65.

Conclusions: Severe pain, not opioids, is associated with the development of delirium during ED stay. Adequate pain control during the hospital stay may contribute to a decrease in delirium episodes.

Delirium is a medical complication that is frequently observed in older patients after surgery, intensive care stay, hospitalization, or a prolonged emergency department (ED) visit. It is characterized by

From the 1Centre d’Étude en Médecine d’Urgence, Hôpital du Sacré-Cœur de Montréal, Montréal, Québec, Canada; the 2Faculté de Médecine, Département Médecine Familiale et Médecine d’Urgence, Université de Montréal, Montréal, Québec, Canada; 3CHU de Québec–Université Laval, Québec, Québec, Canada; the 4Centre d’Excellence du Vieillissement de Québec, Québec, Québec, Canada; the 5Faculté de Médecine, Université Laval, Québec, Québec, Canada; the 6Centre Intégré de Santé et de Services Sociaux de Lanaudière, Joliette, Québec, Canada; and the 7Centre Hospitalier Régional de Trois-Rivières, Trois-Rivières, Québec, Canada.

Received March 3, 2020; revision received May 12, 2020; accepted May 18, 2020.

This study was supported by a grant from the Fonds de recherche des urgentistes de l’Hôpital du Sacré-Cœur de Montréal and the Fonds de recherche du Québec–Santé (FRQS 29307), Canada.

The authors have no potential conflicts to disclose.

Author contributions: RD and ME conceptualized and designed the study and obtained funding; VB managed the study and supervised multicenter recruitment and data collection; JP carried out data management and supervised statistical analysis; RD drafted the manuscript; all remaining authors contributed substantially to its revision; and all authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work. All co-authors have read and agree with the manuscript’s contents.

Supervising Editor: Tammie E. Quest, MD.

Address for correspondence and reprints: Raoul Daoust, MD, MSc; e-mail: raoul.daooust@videotron.ca.

an acute onset of disturbance in consciousness, attention, orientation, memory, thought, perception, and behavior.\(^1\) ED prolonged stay\(^2\) and its associated circumstances (immobility, inadequate hydration and nutrition, lack of stimulation) increase the risk of delirium in older patients. Given its fluctuating nature, delirium is frequently unrecognized by the emergency physicians\(^3,4\) and occurs in approximately 7% to 17% of older ED patients.\(^5–7\) Considering that delirium in older ED patients was independently associated with 6-month mortality,\(^8,9\) prolonged hospitalizations,\(^10–12\) accelerated functional and cognitive decline,\(^13,14\) and ED revisits,\(^15\) it is important to accurately identify its triggering factors. The number and duration of delirium episodes could be substantially reduced by acting on these factors.

Risk factors associated with delirium include advanced age, prior dementia, poor functional status, and multiple comorbidities.\(^5\) Precipitating modifiable risk factors for delirium specific to the ED are pain, urinary retention, constipation, dehydration, environmental variables, and medications.\(^16\) Pain and pain management strategies are important factors related to the development of postoperative delirium in older patients.\(^17–20\) Poorly controlled pain has been identified as a trigger for delirium in hospitalized, nursing home, and older ED patients.\(^21–25\) Paradoxically, exposure to opioids in hospitalized or cancer patients was significantly associated with an increased risk of delirium.\(^26–29\) However, Nandi et al.\(^30\) found that opioids did not increase the risk of delirium in total joint arthroplasty. Another study performed in a surgical context reported that tramadol and meperidine were associated with increased risk of delirium but use of morphine, fentanyl, oxycodone, and codeine were not.\(^31\)

Two studies implemented in the surgical field simultaneously examined the relationship between pain, opioids, and delirium.\(^32,33\) The first study, performed on patients with hip fractures, suggested that untreated pain and opioid underuse were important contributors to the development of delirium. In fact, severe pain significantly increased the relative risk of delirium by nine times.\(^32\) The second study showed that older patients who had high risk of preoperative delirium had higher probabilities of postoperative delirium if they experienced high pain and used high doses of opioids postoperatively.\(^33\) However, the relationship between pain, opioids, and delirium has not been investigated in the ED context which is an enabling environment for delirium.

The objective of this study was to assess the relationship between pain, opioid consumption, and delirium in an older patient population. We hypothesized that when controlling for confounding factors, delirium would be more likely to occur in patients who suffered from severe pain rather than in those consuming opioids during their ED stay.

### METHODS

#### Study Design and Population

This was a planned substudy of a prospective multicenter cohort study conducted in four Quebec (Canada) EDs (two university-affiliated Level I trauma centers and two regional hospitals) between March and July 2015.\(^11\) Patients were included if they 1) were aged 65 and over; 2) had an ED stay of \(\geq\) 8 hours; 3) needed and/or were waiting for admission to any hospital ward; 4) were independent or semi-independent (able to perform five of seven activities of daily living according to the Older Americans Resources and Services [OARS] scale). Even if we know that delirium is more prevalent in patients who are dependent in daily activities, we chose to include patients who are independent or semi-independent because we wanted to study the risk factors during ED stay associated with delirium for the most robust older patients. Patients were excluded if they: 1) had an unstable medical condition requiring admission to the intensive or palliative care units; 2) were unable to consent; 3) were living (or in transition) in a long-term care facility; 4) were unable to speak French or English; 5) presented with delirium within the first 8 hours of ED stay; and 6) had a history of psychiatric disorders (such as schizophrenia, psychotic symptoms, and bipolar disorder).

Potential participants were identified using the ED information system. Research assistants obtained consent and screened the participants for eligibility after their 8-hour exposure to the ED. Sociodemographic and clinical data were collected upon the initial interview by research assistants. Patients’ baseline physical, pain intensity, and cognitive status were assessed in the ED. Patients were screened for delirium during the initial interview and twice daily (with at least 6 hours between each evaluation) during their entire ED stay and up to 24 hours after being admitted to a hospital ward. This 24-hour limit was established by the study’s steering committee, who determined that a delirium that developed more than 24 hours after the...
end of ED stay was unlikely to be caused by factors associated with ED stay. Patients were assessed up to 24 hours postadmission to a hospital ward on the basis that any delirium that developed during those first few hours was most likely the result of the ED stay rather than their time spent on the ward. Potential participants were considered as “missed” when there was no research team member on site for recruitment or the patient was transferred to a hospital bed before the initial interview. Recruitment was ongoing for 12 hours per day, 7 days a week.

The ethics committee of CHU de Québec acted as the centralized research ethics board and approved this study (project MP-20-2015-2130). Informed consent was obtained for each study participant. Patient records/information were anonymized prior to analysis.

**Measures**

Pain intensity was assessed during the initial interview made at least 8 hours after ED admission using the visual analog scale (VAS) rated from 0 to 100. The question asked was: “How would you rate your level of pain?” with “no pain” inscribed on the left of the ruler and “the most intense pain imaginable” on the right end. Patients’ functional status was assessed using the OARS, while the Telephone Interview for Cognitive Status–modified (TICS-m) and the Confusion Assessment Method (CAM) were used to assess cognitive status. All these instruments have been previously used in Canadian elderly populations. Other information on medications, comorbidities (Charlson comorbidity risk index), severity of illness (Acute Physiological and Chronic Health Evaluation II [APACHE II]), and ED environment evaluation were collected in addition to sociodemographic data.

Physicians and ED nurses were blinded to the study’s objectives. ED environmental information, such as presence of proper lighting (according to the research assistant), patient’s hydration, and the presence of physical restraints or medical interventions limiting movement (urinary catheter, oxygen, intravenous [IV] drip) at the initial interview was recorded by the research assistant. ED length of stay and presence of opioid consumption during the ED stay were recorded from administrative databases.

Each site’s research team assistant received standardized training; this included a group training session conducted by the study coordinator and an experienced research nurse and a 5-hour personalized field training. They were also provided with a detailed training manual. Inter-rater reliability was assessed during patient follow-ups at the coordinating site to ensure that the test was administered using structured interview in a standardized manner.

To ensure that the missed patients were similar to our participants, basic clinical and demographic data were collected on those missed patients. The incidence of delirium was also collected for those patients.

**Main Outcomes**

Delirium associated with ED stay was defined as a delirium that occurred either in the ED or within the first 24 hours of the hospital stay. The CAM was administered during the initial interview ensuring that the patient was not already delirious after the first 8 hours of their ED stay. The CAM is the most commonly used tool for the detection of delirium with its sensitivity ranging between 34 and 58% and its specificity between 89 and 94% when performed by a research assistant. However, even if this sensitivity seems low, it has been shown that when the CAM is administered several times during a shift, it is more sensitive than a diagnosis made by a psychiatrist.

The sensitive (SENS) method for interpreting the CAM was used to ascertain delirium; a patient has delirium according to the SENS method if they have either an acute onset or a fluctuation in any of the items evaluated in the CAM, inattention, and either disorganized thinking or altered state of consciousness.

Pain intensity evaluated during the initial interview was dichotomized using a VAS cutoff of ≥65 since it is associated with a severe pain-related interference with functioning. As a sensitivity analysis, we also used a VAS cutoff of ≥35, which represents a moderate pain-related interference with functioning. We searched in the electronic medication database of each ED for opioids given to the patient including codeine, hydromorphone, meperidine, oxycodone, methadone, fentanyl, tramadol, pentazocine, morphine, and their combination with acetaminophen. Hydrocodone is not used for pain management in Canada.

**Data Analyses**

Clinical characteristics of included patients are presented using descriptive statistics. Associations between patients’ characteristics (sociodemographic, clinical
data, and environmental information), opioid consumption, pain intensity, and delirium were evaluated using bivariate logistic regression models. To evaluate the effect of pain intensity and opioid use on delirium controlling for different confounders, a multiple logistic regression analysis with direct entry method was performed. Because of limited cases of delirium, we only selected as confounders four risk factors generally associated with delirium in the literature: age, TICS-m, OARS, and Charlson comorbidity risk index. Frailty was not included in the analysis as a confounding factor because of its association in the literature with pain in older patients. Other basic assumptions for logistic regression (independence of observations, the absence of outliers, and linearity of the continuous covariates to the log odds) were verified. The goodness of fit of the multiple logistic regression was assessed using the Hosmer-Lemeshow test. Unadjusted odds ratios (ORs) and adjusted ORs were calculated with 95% confidence intervals (CIs).

Since our sample is a subanalysis of a study on the incidence of delirium, the 338 ED patients recruited is sufficient to detect with a power of 80% an OR of at least 1.7 using a logistic regression with a baseline probability of 0.12 and an alpha of 0.05 (PASS v11.0). All analyses were performed with an alpha level of 0.05 and using SPSS version 25 (IBM Corp.).

RESULTS

Cohort Characteristics

A total of 338 subjects were recruited in the four participating EDs (Figure 1). Table 1 provides details on sociodemographic, clinical, and environmental variables. For the whole sample, the mean (±SD) age was 77 (±8) years, 51% were female, and median ED length of stay was 32 hours (interquartile range [IQR] = 24–50).

A sample analysis of patients who were missed revealed that they had a similar profile compared to those included in our study; 55% were female, with a mean (±SD) age of 77 (±9) years. The mean (±SD) Charlson comorbidity score was 1.7 (±1.7); 31.7% were considered level 1 or 2 on the Canadian Triage Assessment Scale. The medical notes revealed only one case of incident delirium within 24 hours of triage for this group of patients.

Association Between Pain, Opioids, and Delirium

Forty-one patients (12%; 95% CI = 8.5% to 15.5%) experienced delirium during their hospital stay occurring within a mean (±SD) delay of 47 (±19) hours after ED admission. Thirty-two percent (95% CI = 27% to 37%) of the patients received an opioid during their ED stay and 16% (95% CI = 12% to 20%) had
a pain intensity (VAS) ≥ 65 at the initial interview. On bivariate analysis, being aged ≥ 85, having low triage priority, having lower OARS or TICS-m score, having a glass of water within reach, and having pain intensity ≥ 65 were all associated with higher risk of delirium. Notably, patients who had severe pain at the first interview were 3.3 (95% CI = 1.4 to 7.9) times more likely to suffer from delirium (Table 3). No association was observed between opioids received during ED stay and delirium in multivariate analyses (p = 0.63). Results of the Hosmer–Lemeshow goodness-of-fit test were χ² = 4.72, df = 8, and p = 0.79, indicating good overall performance.

**Sensitivity Analysis**

Among patients with moderate pain intensity (VAS of ≥35), 16% experienced delirium compared to 12% for patients with VAS < 35 (p = 0.31). No association were observed between moderate pain and delirium in multivariate analyses (p = 0.51).

**DISCUSSION**

Results of this study demonstrate, after controlling for several factors known to affect delirium, an association between severe pain intensity and delirium in older ED population. Even though almost one-third of the older patients included in the study received opioids during the ED stay, they did not display higher probabilities of delirium after controlling for pain and other confounders. We also showed that severe and not moderate pain was associated with delirium.

The 12% delirium incidence observed in our older population is within the usual range (7%–17%) reported in other EDs, suggesting that our sample is representative of older ED populations and that delirium occurs frequently in acute care. However, this could be considered a high incidence of delirium since we excluded patients who were living in a long-term care facility (or were in transition) and those who were not independent or semi-independent, which left only the most robust older patients to participate in the study.

Our results confirmed that pain more than opioids is related to delirium, as reported by other studies. However, only severe pain intensity was associated with delirium, which supports the results reported by Morrison et al., who identified patients at risk of delirium as those with a pain intensity score of 4 or 5 on a scale ranging from 1 to 5 (which corresponds to severe or very severe pain). Moreover, studies showed that analgesic and opioids underuse could

---

**Table 1**

Clinical Characteristics of Included Patients

<table>
<thead>
<tr>
<th>Patients’ Clinical Characteristics</th>
<th>Included (N = 338)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age category (years)</strong></td>
<td></td>
</tr>
<tr>
<td>65–74</td>
<td>155 (46)</td>
</tr>
<tr>
<td>75–84</td>
<td>123 (36)</td>
</tr>
<tr>
<td>≥85</td>
<td>60 (18)</td>
</tr>
<tr>
<td><strong>Sex: female</strong></td>
<td>173 (51)</td>
</tr>
<tr>
<td><strong>Triage priority</strong></td>
<td></td>
</tr>
<tr>
<td>High (1 or 2)</td>
<td>107 (32)</td>
</tr>
<tr>
<td>Low (3–5)</td>
<td>231 (68)</td>
</tr>
<tr>
<td><strong>Time of day of ED presentation</strong></td>
<td></td>
</tr>
<tr>
<td>00:00–08:00</td>
<td>64 (19)</td>
</tr>
<tr>
<td>08:00–16:00</td>
<td>193 (57)</td>
</tr>
<tr>
<td>16:00–24:00</td>
<td>81 (24)</td>
</tr>
<tr>
<td><strong>ED length of stay (hours)</strong></td>
<td>32 (24–50)</td>
</tr>
<tr>
<td><strong>OARS</strong></td>
<td>25.9 ± 2.4</td>
</tr>
<tr>
<td><strong>TICS-m adjusted for level of education</strong></td>
<td>29.5 ± 6.1</td>
</tr>
<tr>
<td><strong>Charlson</strong></td>
<td>2.1 ± 2.0</td>
</tr>
<tr>
<td><strong>APACHE II</strong></td>
<td>10.0 ± 3.5</td>
</tr>
<tr>
<td><strong>Adequate lighting according to the research assistant</strong></td>
<td>203 (61)</td>
</tr>
<tr>
<td><strong>Patient hydration</strong></td>
<td></td>
</tr>
<tr>
<td>Fasting</td>
<td>45 (14)</td>
</tr>
<tr>
<td>Glass of water within reach</td>
<td>248 (76)</td>
</tr>
<tr>
<td>Presence of saliva*</td>
<td>184 (56)</td>
</tr>
<tr>
<td>Any IV fluids</td>
<td>276 (84)</td>
</tr>
<tr>
<td>Any physical restraints</td>
<td>174 (53)</td>
</tr>
<tr>
<td><strong>Medical interventions limiting movement</strong></td>
<td></td>
</tr>
<tr>
<td>Urinary catheter</td>
<td>16 (6)</td>
</tr>
<tr>
<td>Oxygen</td>
<td>56 (19)</td>
</tr>
<tr>
<td>Saline lock catheter or IV drip</td>
<td>257 (88)</td>
</tr>
<tr>
<td>Temporal orientation aid (clock, watch, phone, calendar)</td>
<td>202 (60)</td>
</tr>
</tbody>
</table>

Data are reported as No. (%), median (IQR), or mean ± SD. APACHE II = Acute Physiological and Chronic Health Evaluation II; Charlson = Charlson Comorbidity Risk Index; IQR = interquartile range; OARS = Older Americans Resources and Services; TICS-m = modified Telephone Interview for Cognitive Status. *Research assistant verified the presence of saliva under their tongue.
be an important contributor to the development of delirium.25,32 These observations tend to confirm that untreated severe pain can be an important triggering factor of delirium in an acute ED setting.

An unexpected association between having a glass of water within patient’s reach and delirium was observed in bivariate analysis. Since it is difficult to evaluate the real level of patients’ hydration, the presence of bedside accessible water was considered a proxy of that factor known to affect delirium. However, we did not measure the real water intake by these patients, and the absence of a glass of water could simply be because it was consumed. In recent years, the mainstream in pain management whether for chronic or acute conditions has been to avoid opioid use as much as possible. However, opioids remain an important treatment for acute and severe pain.47 In light of our results, the underuse of opioids in the context of acute medicine could be harmful in an older, more vulnerable population. Older patients’ severe pain should probably be treated adequately to prevent delirium, and opioids represent one important treatment alternative to consider.

Table 2
Bivariate Associations Between Patients’ Clinical and Environmental Characteristics and Delirium

<table>
<thead>
<tr>
<th>Patients’ Clinical Characteristics</th>
<th>Delirium</th>
<th>Unadjusted OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No (n = 297)</td>
<td>Yes (n = 41)</td>
</tr>
<tr>
<td>Age category</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65–74</td>
<td>143 (48)</td>
<td>12 (29)</td>
</tr>
<tr>
<td>75–84</td>
<td>110 (37)</td>
<td>13 (32)</td>
</tr>
<tr>
<td>≥85</td>
<td>44 (15)</td>
<td>16 (39)</td>
</tr>
<tr>
<td>Sex: Female</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>147 (50)</td>
<td>26 (63)</td>
</tr>
<tr>
<td>Triage priority</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low (3-4-5)</td>
<td>197 (66)</td>
<td>34 (83)</td>
</tr>
<tr>
<td>High (1-2)</td>
<td>100 (34)</td>
<td>7 (17)</td>
</tr>
<tr>
<td>Time of day of ED presentation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>00:00–08:00</td>
<td>57 (19)</td>
<td>7 (17)</td>
</tr>
<tr>
<td>08:00–16:00</td>
<td>165 (56)</td>
<td>28 (68)</td>
</tr>
<tr>
<td>16:00–24:00</td>
<td>75 (25)</td>
<td>6 (15)</td>
</tr>
<tr>
<td>ED length of stay, h (median ± IQR)</td>
<td>32 (24-48)</td>
<td>40 (25-60)</td>
</tr>
<tr>
<td>OARS</td>
<td>26.1 ± 2.2</td>
<td>24.4 ± 2.6</td>
</tr>
<tr>
<td>TICS-m adjusted for level of education</td>
<td>30.0 ± 5.7</td>
<td>25.9 ± 7.2</td>
</tr>
<tr>
<td>Charlson</td>
<td>2.1 ± 1.9</td>
<td>2.3 ± 2.4</td>
</tr>
<tr>
<td>APACHE II</td>
<td>10.1 ± 3.5</td>
<td>9.3 ± 3.3</td>
</tr>
<tr>
<td>Adequate lighting according to the research assistant</td>
<td>182 (62)</td>
<td>21 (53)</td>
</tr>
<tr>
<td>Patient hydration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fasting</td>
<td>41 (14)</td>
<td>4 (10)</td>
</tr>
<tr>
<td>Glass of water within reach</td>
<td>213 (75)</td>
<td>35 (90)</td>
</tr>
<tr>
<td>Presence of saliva†</td>
<td>160 (56)</td>
<td>24 (60)</td>
</tr>
<tr>
<td>Any intravenous fluids</td>
<td>245 (85)</td>
<td>31 (78)</td>
</tr>
<tr>
<td>Any physical restraints</td>
<td>156 (53)</td>
<td>18 (46)</td>
</tr>
<tr>
<td>Medical interventions limiting movement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urinary catheter</td>
<td>13 (5)</td>
<td>3 (9)</td>
</tr>
<tr>
<td>Oxygen</td>
<td>51 (20)</td>
<td>5 (15)</td>
</tr>
<tr>
<td>Saline lock catheter or intravenous drip</td>
<td>229 (89)</td>
<td>28 (85)</td>
</tr>
<tr>
<td>Temporal orientation aid (clock, watch, phone, calendar)</td>
<td>182 (61)</td>
<td>20 (49)</td>
</tr>
<tr>
<td>Received opioids during ED stay</td>
<td>92 (31)</td>
<td>16 (39)</td>
</tr>
<tr>
<td>Had severe pain during the first interview (VAS ≥ 65)</td>
<td>35 (13)</td>
<td>12 (31)</td>
</tr>
</tbody>
</table>

Data are reported as No. (%), median (IQR), or mean ±SD.
APACHE II = Acute Physiological and Chronic Health Evaluation II; Charlson = Charlson comorbidity risk index; IQR = interquartile range; OARS = Older Americans Resources and Services; TICS-m = modified Telephone Interview for Cognitive Status; Charlson = Charlson comorbidity risk index; VAS = visual analog scale (0–100).
* p < 0.05.
†Research assistant verified the presence of saliva under their tongue.
Our study captured a large multicenter cohort of older patients that were prospectively assessed for delirium using a validated tool during their entire ED stay and first 24 hours of admission. Pain intensity was evaluated using a well-validated and utilized instrument, and all opioid medications were correctly identified as it is mandatory for nurses to enter them into the computerized health management system. In choosing more robust patients and controlling for several environmental factors known to cause delirium in an ED context, precipitating factors were more precisely evaluated, which constitutes a strength in our study.

LIMITATIONS

Our high rate of missed patients is mainly due to the limited resources. However, after comparing sociodemographic characteristics and comorbidities, we have found no significant difference between patients who were included and those who were missed. We excluded patients with moderate to severe dementia, those who lived in long-term nursing homes, those with preexisting psychological conditions, and patients who had a lesser functional level. These exclusions were established because we were mainly interested in investigating the impact of pain and opioids on the most robust older patients. Therefore, our cohort represents only a portion of the older adult population usually seen in the ED and may not be generalized to all seniors. This study can only establish associations because we were unable to account for unmeasured confounders. The CAM was administered by different research assistants, and this could therefore underestinate or overestimatthe acute onset of new symptom frequency. We tried to decrease this potential bias by providing research assistants with standardized training, which was proven effective given our good level of interobserver agreement. The study coordinator also reviewed every single research file to ensure completeness. The misclassification of delirium may have occurred as we excluded patients with delirium using a single initial CAM assessment; this pragmatic approach was used to ensure study feasibility. Pain intensity using VAS was only evaluated once during the initial interview and pain duration was not evaluated. Multiple assessments of pain would have been preferable since pain can fluctuate during ED stay.

CONCLUSIONS

In summary, severe pain, not opioids, is associated with the development of delirium during ED stay. Adequate pain control during ED stay may contribute to a decrease in delirium episodes. We thank all the research assistants who participated in the recruitment of patients. The authors thank Megan Martin and Michel Garner for their contributions to the revision of the manuscript.

References


Table 3

<table>
<thead>
<tr>
<th>Patients’ Clinical Characteristics</th>
<th>Adjusted OR (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age category (years)</td>
<td></td>
</tr>
<tr>
<td>65–74 Reference</td>
<td></td>
</tr>
<tr>
<td>75–84</td>
<td>1.04 (0.42–2.56)</td>
</tr>
<tr>
<td>≥85</td>
<td>3.04 (1.16–7.94)*</td>
</tr>
<tr>
<td>OARS</td>
<td>0.87 (0.75–1.01)</td>
</tr>
<tr>
<td>TICS-m adjusted for level of education</td>
<td>0.93 (0.87–0.99)*</td>
</tr>
<tr>
<td>Charlson</td>
<td>1.04 (0.87–1.25)</td>
</tr>
<tr>
<td>Received opioids during ED stay</td>
<td>1.25 (0.55–2.83)</td>
</tr>
<tr>
<td>Had severe pain during the first interview (VAS ≥ 65)</td>
<td>3.29 (1.38–7.88)*</td>
</tr>
</tbody>
</table>

Charlson = Charlson comorbidity risk index; OARS = Older Americans Resources and Services; TICS-m = modified Telephone Interview for Cognitive Status; VAS = visual analog scale (0–100).

*p < 0.05.


ABSTRACT

Background: A fundamental challenge for emergency department (ED) clinicians is to relieve severe, acute pain while simultaneously avoiding adverse events associated with opioid analgesics. Because there is evidence that intravenous (IV) acetaminophen is an effective adjuvant analgesic in postoperative settings, we examined whether it also has a role in the ED.

Methods: This was a two-arm, double-blind randomized clinical trial. All patients received 1 mg of IV hydromorphone. Patients were then randomized to receive 1 g of IV acetaminophen or placebo. The primary outcome was the between-group difference in change in pain from baseline (before treatment) to 60 minutes after administration of study drugs, measured on an 11-point numeric rating scale (NRS).

Results: Of 828 patients screened, 162 were enrolled and 159 had the primary outcome. Patients allocated to acetaminophen + hydromorphone had a mean decline in pain from baseline to 60 minutes of 6.2 NRS units; those receiving placebo + hydromorphone had a mean decline of 5.4, a difference of 0.8 NRS units (95% confidence interval [CI] = -0.01 to 1.8). Two patients in each group received additional analgesics in the first 60 minutes of the study. At 120 minutes the NRS pain difference was 0.6 (95% CI = -0.4 to 1.6). A total of 26.9% of patients who received acetaminophen wanted more analgesia versus 37.7% of those given placebo (difference = -10.8%, 95% CI = -24.3% to 4.4%). The incidence of adverse effects was similar in both groups.

Conclusions: The addition of 1 g of IV acetaminophen to 1 mg of IV hydromorphone provided neither clinically meaningful nor statistically superior analgesia than hydromorphone alone.

Opioid analgesics are the mainstay of treatment of severe pain in the emergency department (ED). While effective for most patients when titrated to a patient’s pain level, the well-known risk of potentially life threatening, albeit rare respiratory depression and the risk of the more common noxious side effects...
constitute limitations of this class of analgesics. Concern about the increasing prevalence of opioid misuse and overdose ascribed to oral opioids prescribed in the outpatient setting has also resulted in renewed interest in reducing exposure to opioids in the ED.\(^1\)

Combining analgesics with different mechanisms of action is a strategy that can improve analgesia while decreasing the risk of adverse events.\(^2\) A form of acetaminophen that is administered intravenously has been studied as an adjunct to opioid analgesics primarily in postoperative studies. A recent systematic review found that IV acetaminophen contributed to improved analgesia and decreased opioid consumption in patients after surgery.\(^3\) We could only find one study, limited to patients 65 years or older, that evaluated IV acetaminophen as an adjunct to opioid analgesics. In that study, 1 g of IV acetaminophen administered with 0.5 mg of IV hydromorphone did not produce greater analgesia than IV hydromorphone alone.\(^4\)

It is clear that oral opioid misuse is having a devastating effect on individual lives and communities. Reduction in the use of opioids in the ED has been recommended as part of a larger strategy to address this problem. However, it is imperative that this reduction be accompanied by other effective modalities of controlling acute pain so that individual patients do not suffer unnecessarily because of the societal need to contain the opioid epidemic.

The purpose of this study was to compare the analgesia achieved by a combination of 1 g of IV acetaminophen and 1 mg of IV hydromorphone to 1 mg of IV hydromorphone alone in patients younger than 65 years of age presenting to the ED with acute, severe pain.

**METHODS**

**Study Design and Setting**

The study was a randomized clinical trial that examined the analgesic efficacy of IV acetaminophen as an adjunct to IV hydromorphone to treat acute, severe pain in ED patients. It was approved by the Montefiore Medical Center Institutional Review Board and registered at http://www.clinicaltrials.gov (NCT03553498). All patients provided written consent to participate in the study.

The study took place in two EDs of the Montefiore Medical Center. The Moses Division is a community hospital with more than 70,000 adult visits annually, located in the east Bronx, NY. Trained, salaried, full-time, bilingual (English and Spanish) research associates (RAs) staff both EDs 24 hours per day, 7 days per week. Patients were enrolled from December 2018 to June 2019.

**Selection of Participants**

The RAs screened patients aged 21 to 64 years with acute pain (onset within 7 days) that was of sufficient severity to warrant use of intravenous (IV) opioids in the judgment of the ED attending physician. Exclusion criteria were as follows: history of adverse reaction to hydromorphone, morphine, or acetaminophen; hypotension (systolic blood pressure <100 mm Hg); room air oxygen saturation <95%; CO\(_2\) measurement >46 (measured in patients with history of pulmonary disease or with a >20 pack-year smoking history); heart rate <60 beats/min; alcohol or other drug intoxication as judged by the attending physician; use of other opioids within the past 24 hours; use of nonopioid analgesics in past 8 hours; use of a monoamine oxidase inhibitor within 30 days; history of a chronic pain syndrome (such as sickle cell disease, osteoarthritis, or fibromyalgia); use of transdermal pain patches; and pregnancy.

**Intervention**

Patients who met the criteria were randomized after signing a written informed consent in either English or Spanish, depending on patient language preference. The research pharmacist used an online random number generator, www.randomization.com, to allocate patients to treatment group. The research pharmacist made up packets of medication that contained 1 g of IV acetaminophen or 100 mL of normal saline placebo and labeled them with the number generated by the randomization program. Details of the masking of the medications have been previously described.\(^4\) All patients received a single dose of 1 mg of IV hydromorphone. The nurse retrieved the next sequentially ordered packet from the automated medication dispensing system and then administered the study medication, acetaminophen or placebo, to the patient. The acetaminophen was prepared ahead of time and replaced every 3 days if unused.

Measures of pain, vital signs, and presence of side effects were obtained immediately before IV hydromorphone was administered (baseline). These measurements were taken again 5 minutes after the
Methods of Measurement

Patients were asked to rate their pain on a previously validated and reproducible5 11-point standard verbal numeric rating scale (NRS) ranging from 0 (“no pain”) through 10 (“worst pain possible”) at baseline, at 5 minutes after the end of the administration of the study medications, and every 15 minutes until 120 minutes following initial administration of the study medication. Systolic blood pressure, respiratory rate, oxygen saturation, nausea, vomiting, and pruritus were assessed at all time points. Adverse events were defined as respiratory rate <10 breaths/min, room air oxygen saturation <95%, heart rate <50 beats/min, systolic blood pressure <90 mm Hg, or use of naloxone at any time.

Information about all additional medications administered to patients, including name of medication, dose, route of delivery, and time of administration was abstracted from the medical record during the patient’s stay in the ED. Thus, any ambiguities in the medical record could be resolved by speaking directly to the patient and the patient’s providers in real time.

The RAs entered data directly on a standardized data collection instrument using REDCap, a Web-based application designed to support data capture for investigational purposes. Final data were downloaded from REDCap and analyzed using IBM SPSS Statistics version 22.

Outcome Measures

The primary outcome was the between-group difference in NRS pain scores from baseline to 60 minutes postadministration of study medications (i.e., baseline pain score minus 60-minute pain score). Additional outcomes included change in pain score from baseline to 120 minutes, NRS pain scores across the entire study period, a comparison of incidence of side effects and adverse events, and a comparison of the proportion of patients who chose to forgo additional pain medication at 120 minutes postbaseline when asked the previously standardized question, “Do you want more pain medication?”

Primary Data Analysis

All variables are shown as means with standard deviations (SDs), medians with interquartile ranges (IQR), or proportions with 95% confidence intervals (CIs) as appropriate. The independent samples t-test was used to compare the mean difference in NRS pain scores from baseline to 60 minutes between treatment groups. The difference in pain scores is presented as means and 95% CIs. The point estimates of pain at each time point are presented graphically with 95% CIs and are tabulated. We added a post hoc analysis of the difference in receipt of medications used to treat nausea, vomiting, and abdominal pain.

There are several approaches to the analysis of change. The primary analysis specified in the protocol and analysis plan is an analysis of the simple difference between pain scores at baseline and subsequent time points. An alternative approach is to use multivariable linear regression to adjust for any initial differences in the baseline measure of pain. We conducted a post hoc linear regression with the baseline pain score included as a covariate as well as treatment group to take baseline pain into account. For categorical variables we used Wilson’s method of calculating a 95% CI around differences between proportions.7

We estimated that a sample size of 148 (74 per group) would be needed to detect a difference of at least 1.3 NRS units in the improvement in NRS pain score between the two study groups, assuming a common SD of 2.8 (based on prior studies we have performed)8,9 for an alpha of 0.05 and 80% power. A difference of 1.3 NRS units is a previously validated and reproducible measure used to quantify the “minimum clinically significant difference” in change in NRS pain score.5,10,11 To account for potential protocol violations and missing data, we enrolled an extra seven patients in each group. Thus, our final target sample size was 162 subjects or 81 per group. nQuery Advisor version 6.0 was used to calculate the sample size.

RESULTS

Characteristics of Study Subjects

The RAs screened 828 patients. As shown in Figure 1, the two most common exclusions were pain present more than 7 days before the ED visit and attending physician did not want to prescribe opioids. Seventy-five patients refused to be screened. Two patients in the placebo group and one patient in the acetaminophen group did not have primary outcome

- hydromorphone + study medicine (acetaminophen or placebo) were administered and at 15-minute intervals for the next 120 minutes. Additional analgesics were administered if patients requested more pain relief. At 120 minutes all patients were asked whether they wanted more pain medication. Subsequent care of the patients was at the discretion of the treating attending physician.

- Seventy-five patients refused to be screened. Two patients in the placebo group and one patient in the acetaminophen group did not have primary outcome.
data at 60 minutes. Thus, 159 patients were included in the primary analysis. The sample was primarily female and Hispanic and had abdominal or flank pain (Table 1). The two treatment groups had similar distributions of baseline characteristics.

**Main Results**

Compared to patients who only received IV hydromorphone, patients who received both IV hydromorphone and IV acetaminophen had larger decreases in pain from baseline to 60 minutes postbaseline (mean difference = 0.8, 95% CI = −0.1 to 1.8, p = 0.08), although this difference failed to achieve clinical or statistical significance (Table 2). The findings from the alternate statistical analysis that controlled initial pain rating were nearly identical to the analysis of change scores. The mean difference in reduction in pain between groups after controlling baseline pain was 0.8 (95% CI = 0.2 to 1.7, p = 0.11). The group that received IV acetaminophen and IV hydromorphone reported slightly better analgesia on the three secondary pain outcomes than those who received hydromorphone and placebo but these differences were neither clinically nor statistically significant (Table 2). Figure 2 presents the mean pain scores graphically and in tabular format at all time points. The patients who received acetaminophen had lower means at all time points. At 15 and 30 minutes the differences between the means were statistically significant but did not meet the threshold of a minimum clinically significant difference.

A post hoc analysis comparing receipt of other medications between groups was conducted as receipt of nonstudy medications that alleviate pain could mask differences if they are unequally distributed. More patients in the placebo arm received antiemetics or other medication for abdominal pain (36 [45.6%]) than patients given acetaminophen (28 [33.8%]) (difference = 10.6%, 95% CI = −4.6% to 25.1%); however, this difference was not statistically significant (p = 0.17).

Ten patients received additional opioid analgesics, four in the first hour and six in the second. The
difference between treatment groups was not statistically significant (Table 2). As receipt of additional analgesics could affect the outcomes, the analysis of change in pain from baseline to 60 minutes was rerun with the four patients who received additional opioids before 60 minutes excluded. All 10 patients who received additional opioid analgesics during the 2-hour time period were excluded from analyses of change from baseline to 2 hours. The analyses were essentially unchanged after these adjustments, with a between-group difference of 0.9 (95% CI = −0.04 to 1.8) at 60 minutes and 0.8 (95% CI = −0.2 to 1.9) at 120 minutes.

There were few abnormal vital signs. In the group that received both analgesics, one patient had an oxygen saturation below 95%, and three had systolic blood pressures <100 mm Hg. Of those who received IV hydromorphone alone, one had oxygen saturation <95% and one had a heart rate <50 beats/min. All of these vital sign abnormalities were transient. Nausea, vomiting, and pruritus were more common than abnormal vital signs but did not differ between the two groups as shown in Table 3.

**DISCUSSION**

Effective treatment of severe pain is a core responsibility of the ED. There is evidence that a large proportion of ED patients do not achieve adequate pain relief, even with doses of IV opioids that are larger than what is commonly given. A series of studies our research group has conducted indicates that opioids, even given in substantial doses (1–2 mg of IV hydromorphone) do not provide adequate analgesia for many patients. Between 30 and 40% of patients want additional analgesia 60 and 120 minutes after the first dose.12,13 High opioid requirements have also been documented in other settings. A study of postoperative

---

### Table 1
Baseline Characteristics of Sample

<table>
<thead>
<tr>
<th></th>
<th>Hydromorphone + IV Acetaminophen (n = 80)</th>
<th>Hydromorphone + Placebo (n = 79)</th>
<th>Difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>52 (65.0)</td>
<td>48 (60.8)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>28 (35.0)</td>
<td>31 (39.2)</td>
<td></td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>54 (67.5)</td>
<td>53 (67.1)</td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>22 (27.5)</td>
<td>17 (21.5)</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>2 (2.5)</td>
<td>5 (6.3)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>2 (2.5)</td>
<td>4 (5.1)</td>
<td></td>
</tr>
<tr>
<td>Age (years), mean (±SD)</td>
<td>43.3 (±12.0)</td>
<td>43.1 (±13.5)</td>
<td></td>
</tr>
<tr>
<td>Location of pain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abdomen/flank</td>
<td>70 (87.5)</td>
<td>67 (84.8)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>10 (12.5)</td>
<td>12 (15.2)</td>
<td></td>
</tr>
<tr>
<td>Pain intensity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6–7</td>
<td>7 (8.8)</td>
<td>8 (10.1)</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>12 (15.0)</td>
<td>19 (24.1)</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>11 (13.8)</td>
<td>9 (11.4)</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>50 (62.5)</td>
<td>43 (54.4)</td>
<td></td>
</tr>
<tr>
<td>Nauseated or vomited in ED before treatment?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>57 (71.3)</td>
<td>52 (65.8)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>23 (28.7)</td>
<td>27 (34.2)</td>
<td></td>
</tr>
</tbody>
</table>

Data are reported as n (%), unless otherwise specified.

---

### Table 2
Pain Outcomes by Treatment Group

<table>
<thead>
<tr>
<th></th>
<th>Hydromorphone + Acetaminophen (n = 80)</th>
<th>Hydromorphone + Placebo (n = 79)</th>
<th>Difference (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean decline in NRS scores between baseline* and 60 minutes after administration of study drug</td>
<td>6.2</td>
<td>5.4</td>
<td>0.8 (−0.1 to 1.8)</td>
</tr>
<tr>
<td>Mean decline in NRS scores between baseline and 120 minutes postbaseline</td>
<td>6.1</td>
<td>5.5</td>
<td>0.6 (−0.4 to 1.6)</td>
</tr>
<tr>
<td>Want more medication at 120 minutes†</td>
<td></td>
<td></td>
<td>−10.8% (−24.3% to 4.4%)</td>
</tr>
<tr>
<td>% yes</td>
<td>21 (26.9%)</td>
<td>29 (37.7%)</td>
<td></td>
</tr>
<tr>
<td>Rescue opioids received</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–60 minutes postbaseline</td>
<td>2 (2.5%)</td>
<td>2 (2.5%)</td>
<td>0.0% (−6.5% to 6.4%)</td>
</tr>
<tr>
<td>61–120 minutes postbaseline</td>
<td>4 (5.0%)</td>
<td>2 (2.5%)</td>
<td>2.6% (−4.3% to 10.0%)</td>
</tr>
</tbody>
</table>

NRS = numeric rating scale.

*Baseline = time immediately before treatment.
†Two patients in the hydromorphone + IV acetaminophen group and two in the hydromorphone + placebo group were missing data; thus, the sample sizes were 78 and 77, respectively.
pain found that half of the patients required more than 12 mg of IV morphine to control pain.\textsuperscript{14}

Concern over the opioid epidemic related to over-prescribing of oral opioids has spilled over into the ED itself, with some calling for an “opioid-free” ED.\textsuperscript{15} While there is no evidence that a limited exposure to IV opioids leads to abuse, the concept of combining analgesics with different mechanisms of action may have an opioid-sparing effect. A formulation of acetaminophen administered intravenously has been widely used in Europe for more than 20 years as an adjunct to opioid analgesics in postoperative care. The IV form obtained full FDA approval in the United States in 2010. Acetaminophen has a good safety profile and is not associated with the severe adverse effects of opioids, such as respiratory depression or hemodynamic complications. However a serious disadvantage of IV acetaminophen is its cost. Currently a 1,000-mg dose costs approximately $40.

Table 3
Side Effects by Treatment Group

<table>
<thead>
<tr>
<th>Time point (minutes)</th>
<th>Hydromorphone + IV Acetaminophen Mean NRS scores</th>
<th>Hydromorphone + Placebo Mean NRS scores</th>
<th>Difference</th>
<th>95% Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>9.3</td>
<td>9.1</td>
<td>0.2</td>
<td>(-0.1, 0.6)</td>
</tr>
<tr>
<td>5</td>
<td>4.7</td>
<td>5.3</td>
<td>-0.6</td>
<td>(-1.6, 0.4)</td>
</tr>
<tr>
<td>15</td>
<td>3.4</td>
<td>4.6</td>
<td>-1.2</td>
<td>(-2.1, -0.3)</td>
</tr>
<tr>
<td>30</td>
<td>3.2</td>
<td>4.1</td>
<td>-1.0</td>
<td>(-1.9, -0.0)</td>
</tr>
<tr>
<td>45</td>
<td>2.9</td>
<td>3.7</td>
<td>-0.8</td>
<td>(-1.8, 0.2)</td>
</tr>
<tr>
<td>60</td>
<td>3.1</td>
<td>3.7</td>
<td>-0.6</td>
<td>(-1.6, 0.3)</td>
</tr>
<tr>
<td>75</td>
<td>3.2</td>
<td>3.5</td>
<td>-0.3</td>
<td>(-1.3, 0.7)</td>
</tr>
<tr>
<td>90</td>
<td>3.4</td>
<td>3.7</td>
<td>-0.2</td>
<td>(-1.3, 0.8)</td>
</tr>
<tr>
<td>105</td>
<td>3.2</td>
<td>3.8</td>
<td>-0.5</td>
<td>(-1.5, 0.5)</td>
</tr>
<tr>
<td>120</td>
<td>3.2</td>
<td>3.6</td>
<td>-0.4</td>
<td>(-1.4, 0.5)</td>
</tr>
</tbody>
</table>

*Nausea or vomiting in patients who were not nauseated and had not vomited before the study medications were administered.
In this study the addition of 1 g of IV acetaminophen failed to meaningfully increase the analgesia associated with 1 mg of IV hydromorphone. Both comparison groups showed a similar decline over the entire 2-hour time period. The mean decrease in both groups was clinically substantial: 6.2 NRS units in those who received both analgesics and 5.3 in those who only received hydromorphone. The decrease in pain was greater in the group that received both active analgesics at all time periods and was statistically significant at 15 and 30 minutes after the medications were administered. However, the differences were less than a commonly used threshold indicating a minimum clinically significant difference and may be simply due to chance. Furthermore, there were no differences in receipt of rescue medication, request for additional analgesia, or adverse effects in either comparison group.

Despite the lack of clinical and statistical significance of the primary outcome, the patients who received both active medications had better efficacy outcomes than those treated only with hydromorphone. It is possible that the combination analgesic is more efficacious than opioid alone and that a clinical trial designed to detect a difference of the magnitude observed in this study might achieve statistical significance. However the high cost of IV acetaminophen might not be worth the small benefit conferred given competition for limited resources in the ED.

The only study we could find of the adjunctive use of IV acetaminophen in the ED is one conducted by our research group. That study was restricted to patients aged 65 and older. The methodology was similar to that of the current study, with the exception that the elderly patients enrolled in this earlier trial received 0.5 mg of hydromorphone rather than 1 mg given in the current study of patients under 65 years. The results of the two studies, although conducted in different age groups, were essentially identical.

The combination of different classes of analgesics has been used extensively in postoperative treatment of pain. A Cochrane review of IV acetaminophen and IV propacetamol (a prodrug that is metabolized to paracetamol) in a heterogeneous group of surgical procedures found substantial analgesic benefit of IV acetaminophen and consumption of less opioid, usually self-administered via patient controlled analgesia pumps. It is possible that its efficacy may vary by type of surgery. A recent review of IV acetaminophen to treat pain following total knee replacement found no benefit for pain relief or reduced opioid requirement on postoperative day 1. A systematic review of IV acetaminophen to control pain following cardiac surgery found only marginal benefit of this analgesic either for pain control or reduced opioid requirement. The authors concluded that the routine use of IV acetaminophen was not supported by the evidence. Similarly a review of abdominal surgery did not find 12-hour pain scores and 24-hour opioid consumption to differ between patients who received IV acetaminophen or placebo. As approximately 85% of patients in the current study presented with abdominal pain, it is possible that IV acetaminophen is less effective for abdominal pain than for other specific types of pain.

**LIMITATIONS**

The study has several limitations. Patients with acute pain were enrolled regardless of the etiology of the pain. If the efficacy of IV acetaminophen varies by type of pain, site of pain, or condition, then the preponderance of abdominal pain in this study may limit generalization to other types of pain. Further, the demographic composition of the sample, predominantly female and Latina, may also limit extrapolation to other populations.

Patients were eligible for the study if their physicians felt that the severity of their pain warranted use of IV opioid analgesics. Because there is substantial variation in pain treatment practices, this group of patients is likely to be quite heterogeneous and skewed toward patients with more severe pain. This should not threaten the comparison of the two groups but the findings may not hold true for patients with lower levels of pain.

The group difference between change in pain was statistically significant and sizeable at two time points, 1.2 NRS units at 15 minutes and 1.0 at 30 minutes. It is possible that these differences reflect chance factors rather than indicate a true difference particularly in light of the many comparisons that were evaluated.

Four patients received additional analgesics in the first 60 minutes of the study, and six more patients received more analgesics in the 120-minute period. Receipt of additional analgesics by patients with inadequate pain control could bias the measure of change in pain, as pain ascribed to the study medication could be due to receipt of rescue medication. While this was unlikely to bias the results because the distribution of patients who received more pain medication
was similar in the two groups, we performed the primary analyses with these patients excluded. The results were unchanged.

CONCLUSIONS

In conclusion, the results of this study indicate that the addition of 1 g of intravenous acetaminophen to 1 mg of hydromorphone does not augment the analgesia conferred by 1 mg of hydromorphone alone for patients treated in the ED for severe acute pain.

REFERENCES

Depression and Functional Outcomes in Patients Presenting to the Emergency Department With Low Back Pain

Joshua Ring, MD1, Sarah Peskoe, PhD2, Congwen Zhao, MS2, Benjamin W. Friedman, MD, MS3, Steven Z. George, PT, PhD4,5, and Stephanie A. Eucker, MD, PhD1,4

ABSTRACT

Objectives: Low back pain (LBP) is a common reason for patients to present to emergency departments (EDs). Our objective was to describe the associations between depressive symptoms, pain severity, and functional impairment up to 3 months after initial ED presentation for LBP.

Methods: We performed a secondary analysis on an observational cohort of adult patients from a high-volume, urban ED. Initial depressive symptoms (Patient Health Questionnaire-9) and disability (Roland Morris Disability Questionnaire) were collected in person at the time of initial ED visit and by telephone at 1-week and 3-month follow-ups. Pain intensity (Numeric Rating Scale) was collected at 1-week and 3-month follow-ups. Our primary goal was to determine the associations between initial depressive symptoms and pain intensity and disability scores at 3 months. We also investigated the associations of initial and 3-month change in depressive symptoms with change in disability score from initial presentation to 3 months and change in pain score from 1 week to 3 months.

Results: Of the 674 patients initially enrolled, 362 patients had complete depressive symptom, pain, and disability data and were included in the final analysis. Those with higher levels of intake depressive symptoms had worse pain intensity ($B = 0.14, 95\% \text{confidence interval} [\text{CI}] = 0.08$ to $0.21$) and disability ($0.46, 95\% \text{CI} = 0.30$ to $0.62$) severity at 3 months, with less improvement in disability over the 3 months ($B = 0.22, 95\% \text{CI} = 0.05$ to $0.40$). Furthermore, those with worsening depressive symptoms over the 3-month study period experienced less improvement in pain intensity ($B = 0.10, 95\% \text{CI} = 0.05$ to $0.17$) and disability ($B = 0.84, 95\% \text{CI} = 0.66$ to $1.02$) over the same time frame. Except for a slight strengthening of the association between initial depressive symptom severity and 3-month pain score among patients with no prior LBP episodes, history of prior LBP episodes did not moderate these relationships.

Conclusions: Significant positive temporal associations exist between initial severity and 3-month progression of depressive symptoms and 3-month pain intensity and disability outcomes for ED patients with LBP. Future work is needed to investigate whether behavioral interventions initiated from the ED may mitigate the incidence and severity of LBP-related chronic pain and functional impairments.

Background

Low back pain (LBP) is responsible for approximately 2.6 million annual visits to U.S. emergency departments (EDs), representing one of the most common chief complaints among ED patients.\textsuperscript{1} Approximately 40\% of patients presenting to the ED with acute LBP report pain-related functional impairment 3 months later.\textsuperscript{2,3} These patients are at risk of
developing chronic pain, which is often a highly debilitating condition. As a result, investigators have sought to identify risk factors for the development of chronic back pain. Several studies in the general population and primary care settings have shown an association between more severe depressive symptoms and worse LBP outcomes. Depressive symptoms have been measured using various scales such as the Patient Health Questionnaire (PHQ) and Hospital Anxiety and Depression Scale (HADS) and include symptoms of feeling down, fatigue, poor appetite, and loss of interest or enjoyment in activities. Furthermore, patients with higher levels of depressive symptoms and psychological distress have increased risks of short- and long-term missed work and a threefold increase in health care system utilization and account for a larger proportion of health care costs among patients with LBP.

The strength of the association between depressive symptoms and LBP-related disability may be dependent on the presence of additional clinical factors. One such clinical factor is the “chronicity” of LBP or how long the patient has been experiencing LBP. Chronicity is characterized by multiple features, including frequency, duration, and severity of exacerbations as well as the associated functional impairment. A study in workers at risk for LBP found that those with more psychological distress had more frequent LBP episodes as well as greater duration of disability from LBP than those without psychological distress. However, it is unknown whether chronicity impacts the strength of association between depressive symptoms and functional outcomes in LBP.

Goals of This Study
Therefore, the purpose of this analysis was to examine the relationships between depressive symptom severity and 3-month pain and functional outcomes among patients presenting to the ED with a primary complaint of LBP. The association between depressive symptoms and LBP-related functional impairment is well established in primary care but has not been described in the ED setting. Our primary hypotheses were that after an initial ED visit, higher depressive symptomatology at baseline would be associated with: 1) higher 3-month pain and disability scores and 2) smaller improvements in 3-month pain and disability scores. Additionally, we hypothesized that worsening depressive symptoms over 3 months would be associated with less improvement in pain and disability over the same period. Finally, we explored whether these associations were moderated by having had episodes of LBP prior to the ED visit. This factor was selected because of its clinical relevance and potential for increasing the strength of the associations between depressive symptoms, pain, and functional impairment.

METHODS
Study Design
This was a planned secondary analysis of data obtained from an observational cohort of patients presenting to an urban ED with LBP.

Study Setting and Participants
The initial study and data acquisition was performed in an urban, academic, single-center ED with >100,000 visits per year. Paid research associates bilingual in English and Spanish collected data 7 days per week, 18 to 24 hours per day, which encompassed all hours of the day including mornings, evenings, and many overnights, during the study period (July 2009 to March 2010). Study participants were English- or Spanish-speaking adults (≥21 years) who presented with a primary chief complaint of LBP, defined as pain between the inferior tips of the scapulae and the top of the buttocks, and deemed musculoskeletal in etiology by an ED attending. Etiologies of LBP included both blunt trauma and atraumatic LBP. Patients were excluded if their LBP was of nonmusculoskeletal etiology.

Study Protocol and Baseline Measures
Patients who met inclusion criteria were approached by a research associate after their pain had been treated by an ED provider and asked a series of questions in a 20-minute in-person interview. The following baseline data were collected in the ED:

1. Patient demographics including age, sex, race, ethnicity and education, duration of current episode of LBP, and history of prior episodes of LBP.
2. The PHQ to assess the severity of depressive symptoms. The depression subscale of the PHQ (PHQ-9) is a screening tool used in numerous patient populations to accurately quantify the severity of patient-reported depressive symptoms from 0 (none) to 27 (severe), with high internal reliability, test–retest reliability over 2 to 14 days, and criterion validity when compared to clinical assessment by mental health professionals.
3. The Roland–Morris Disability Questionnaire (RMDQ), a bedside functional disability scale developed for LBP research that ranges from 0 (no disability) to 24 (maximum disability).

The RMDQ has demonstrated its test–retest reliability over 1- to 14-day time periods, and it is sensitive to change over 3- to 6-week time periods in both acute and chronic back pain patients with a minimum clinically important difference ranging from 2 to 5 points.

Follow-up and Outcome Measures

Study participants were asked a 10-minute series of follow-up questions by research associates during telephone calls at 1 week and 3 months after the initial ED visit. These included the RMDQ at 1 week and 3 months, PHQ at 3 months, and severity of LBP at 1 week and 3 months using the Numeric Rating Scale (NRS), which ranges from 0 (no pain) to 10 (worst pain possible; Figure 1).

The primary outcomes were RMDQ and NRS pain scores at 3 months. Secondary outcomes were change in RMDQ score from baseline to 3 months and change in pain score from 1 week to 3 months. The hypothesized predictor variables were severity of depressive symptoms using PHQ-9 score at initial ED visit and change in PHQ-9 score from baseline to 3 months.

Data Analysis

Descriptive statistics were used to summarize patient characteristics, primary and secondary outcome measures, and hypothesized predictor variables and effect moderators. PHQ-9, RMDQ, and NRS pain scores were treated as continuous variables in all analyses.

Associations Between Depressive Symptoms, Disability, and Pain. Linear regression models were used to test each of the study hypotheses. Separate regression models were used to determine the associations between: 1) baseline PHQ-9 and RMDQ at 3 months, 2) baseline PHQ-9 and change in RMDQ from baseline to 3 months, 3) baseline PHQ-9 and pain score at 3 months, 4) baseline PHQ-9 and change in pain score from 1 week to 3 months, 5) change in PHQ-9 from baseline to 3 months and change in RMDQ from baseline to 3 months, and 6) change in PHQ-9 from baseline to 3 months and change in pain score from 1 week to 3 months. All models were run both unadjusted and adjusted for potential confounders, including race, ethnicity, age, sex, and education, which have been associated with pain and disability outcomes.

Models testing the 3-month change in PHQ-9 as a predictor were also adjusted for baseline PHQ-9 score. While not strictly continuous, we assume that the NRS, RMDQ, and PHQ-9 are each an approximation of a continuous scale for regression analysis. Variance inflation factors (VIFs) were used to test for multicollinearity, with a threshold of 10 used to determine high correlation for all models. Cook’s Distance and DFBETAS were used to identify potential influential observations, and sensitivity analyses were performed with these points removed for each model. A robust standard error estimator was used to allow for heteroscedasticity of the data in all models.

Influence of Potential Effect Moderator. We also tested the effects of prior LBP episodes as a potential moderator of these relationships. This variable was included as both a parameter and an interaction term with either baseline PHQ-9 or change in PHQ-9 from baseline to 3 months in additional adjusted analyses of each of the regression models. For prior LBP episodes, patients were categorized into either 1) no prior episodes of LBP or 2) any prior episodes of LBP, regardless of frequency or severity of episodes. This grouping was chosen based on the low numbers of patients with prior LBP episodes at least once or more per year.

Sensitivity Analyses. Sensitivity analyses were performed to: 1) assess the impact of using 1-week RMDQ instead of baseline RMDQ and 2) assess the potential for bias due to missing data. For (1), models for change in RMDQ were refit using the difference between 1 week and 3 months. To assess the association and agreement between RMDQ at baseline and at 1 week, we performed the following analyses: 1) compare the distributions using a nonparametric...
Wilcoxon signed-rank test, 2) calculate a nonparametric Spearman’s correlation coefficient to assess the correlation between RMDQ score at baseline and at 1 week, and 3) calculate a weighted kappa statistic and intraclass correlation coefficient.24 If a significant difference was found between baseline and 1-week RMDQ scores, additional linear regression analyses were performed to determine whether the associations between PHQ-9 and RMDQ change from baseline to 3 months could be explained by the change in RMDQ from baseline to 1 week. Specifically, we planned to examine the associations between baseline PHQ-9 and RMDQ change from baseline to 1 week and RMDQ change from 1 week to 3 months and the association between change in RMDQ from baseline to 1 week and change in PHQ-9 from baseline to 3 months. For (2), we compared the demographic and outcome distributions among patients with complete information to those without complete information. If there were statistically significant differences, we performed additional analyses using multiple imputation to create five imputed data sets for each outcome using the following specification. Assuming that a joint distribution exists for all variables included in the regression model, we used the fully conditional specification method to impute the missing values. We included all available explanatory variables (Table 1) and outcomes in the imputation model, separately for each outcome. Continuous variables were imputed using the predictive mean matching regression method while categorical variables were imputed using the discriminant function method. Estimates were then pooled using Rubin’s Rules.25 A two-sided type I error rate of 0.05 was used for all statistical tests. All analyses were conducted using SAS 9.4 (SAS Institute) and R (R Core Team).

RESULTS

After those without complete PHQ-9, RMDQ, and NRS scores at all time points were excluded, 362 adult ED patients with LBP were included in the final analysis (Figure 2; additional figures are shown in Data Supplement S1). Summary statistics of patient characteristics are shown in Table 1, and outcomes are shown in Table 2. Overall, this patient population had high levels of pain and disability initially that improved over time and low levels of depressive symptoms across all time points (Table 2, Figure S1).

**Associations Between Depressive Symptoms, Disability, and Pain**

Regression models indicated that higher baseline severity of depressive symptoms was associated with worse...
levels of both pain and disability at 3 months, as well as with less improvement in function over time from ED presentation to 3 months (Table 3). These associations remained after adjusting for potential confounders. The effect sizes of the associations may be clinically relevant. For instance, the regression coefficients (B) indicate that a 10-point difference in baseline depressive symptoms on the 27-point PHQ-9 scale is associated with a 4.6-point increase in 3-month RMDQ and a 1.4-point increase in 3-month pain NRS, which are thought to represent clinically important differences.

Additionally, worsening depressive symptoms from baseline to 3 months were associated with less improvement in both pain and disability (Table 4). Furthermore, the 3-month change in depressive symptoms was found to have a stronger association than baseline depressive symptom severity with change in pain and disability over 3 months. These associations remained significant after adjusting for age, sex, race, ethnicity, and education and were stronger after adjusting for baseline depressive symptom severity. Calculated VIF for all variables in all adjusted models was <2.5; therefore, there are no concerns with multicollinearity. All adjusted models removing influential points provided point estimates with similar inference, with estimates provided in Data Supplement S1, Tables S1 and S2 (available as supporting information in the online version of this paper, which is available at http://onlinelibrary.wiley.com/doi/10.1111/acem.13957/full).

### Influence of Potential Effect Moderator

When we examined potential moderation of these associations, we found only a small strengthening of the association between baseline depressive symptom severity and 3-month pain score among patients with no prior LBP episodes (B = 0.26, 95% CI = 0.14 to 0.38) than among those with at least one prior episode of LBP (B = 0.11, 95% CI = 0.04 to 0.18; p-interaction = 0.03). No other significant interaction was found between prior episodes of LBP and depressive symptom severity at baseline or change over 3 months on any of the other pain or disability outcomes (Data Supplement S1, Table S3).

### Sensitivity Analyses

Additional analyses revealed differences in RMDQ scores from baseline to 1 week. While there was a
strong association between baseline PHQ-9 and change in RMDQ from baseline to 3 months, we did not find an association between baseline PHQ-9 and change in RMDQ from baseline to 1 week or from 1 week to 3 months (Data Supplement S1, Table S4). It may be that the changes in RMDQ over these shorter time periods were too small to detect a correlation with this sample size. Alternatively, it may reflect the existence of different functional recovery trajectories, in which depressive symptoms are more correlated with longer term than short-term changes. By contrast, there was a weak but significant association between change in RMDQ from baseline to 1 week and change in PHQ-9 from baseline to 3 months (Data Supplement S1, Table S4), suggesting a reciprocal relationship between depressive symptoms and disability.

Finally, demographics were compared between patients with complete information and those without complete information, and the only statistical difference was by age (mean difference of 2 years). Multiple imputation of missing age yielded only slight changes in the CI sizes with overall no change in the associations identified across all regression models.

**DISCUSSION**

This is one of the first analyses we are aware of to describe the relationship between depressive symptom severity and pain-related functional impairments in patients presenting to an ED with musculoskeletal LBP. Consistent with our study hypotheses and prior work in primary care settings, these results indicate that patients with higher depressive symptom severity at their initial ED visit had less improvement in LBP-related disability and worse overall pain and disability outcomes at 3 months. Furthermore, we found that worsening depressive symptom severity over the 3 months post-ED visit had an even stronger association than the baseline depressive symptom level with impaired 3-month pain and disability recovery. Sensitivity analysis of the temporal relationships between change in depressive symptoms and change in disability suggests a reciprocal relationship between these two outcomes. Interestingly, most patients reporting any depressive symptoms reported only mild-to-moderate symptom severity, suggesting that even subclinical depressive symptomatology is an important prognostic factor for chronic pain and disability. In general, chronicity of LBP did not moderate these relationships. The clinically heterogeneity of the subjects, which includes a range of LBP chronicity, prior management, testing, imaging, and pain treatment, serves to increase the generalizability of the study findings.

These findings add to the growing body of literature describing the adverse influence of depressive symptoms on LBP outcomes. Similar correlations between greater depressive symptoms and worse pain and functional outcomes over the course of treatment were found in patients receiving physical therapy for musculoskeletal pain. Additionally, in a cohort of patients seen in primary care for acute LBP, those with higher initial levels of depressive symptoms had worse functional recovery 3 and 6 months later. Furthermore, in the Stepped Care for Affective Disorders and Musculoskeletal Pain (SCAMP) cohort of primary care patients with chronic musculoskeletal pain, a 12-month longitudinal assessment demonstrated that change in depression predicted severity of subsequent pain. Intriguingly, the SCAMP cohort also demonstrated the reciprocal relationship in which change in pain predicted change in severity of subsequent depression.
One prior analysis in our study population examining a two-question depression screening tool did not show an association with disability. This finding may have been limited by subject recall bias, as this screening questionnaire asked patients to retrospectively determine whether they had depressive symptoms before their current episode of back pain. Moreover, it is likely that this two-question tool is insufficiently sensitive compared to the nine-question battery of the PHQ-9, which considers a broader range of depressive symptoms, particularly given the very low scores reported in the 9-item version in our cohort.

In investigating potential effect modifiers, only patients with first-time LBP showed a slightly increased association between depressive symptom severity and severity of subsequent pain at 3 months. While the relationship between depressive symptoms and chronic back pain severity is well-established, only recently has the role of depressive symptoms in the transition from acute to chronic pain been explored. A recent systematic review found that in multiple studies based in primary care, pain clinics, and workplace settings, measures of psychological distress, depressive symptoms, and depressive mood were consistently predictive of the progression from acute to chronic LBP. Furthermore, in the previously mentioned cohort of primary care patients with acute LBP, history of prior back pain did not modify the relationship between depressive symptoms and worse long-term functional outcomes.

LIMITATIONS

Data were collected in a convenience sample at an urban academic center in a largely medically underserved community, so patient characteristics may not reflect those of other patient populations. The majority of patients were female (62.4%) and/or Hispanic (65.5%). A third of patients had missing PHQ-9 scores, and several had missing pain and disability outcomes measures. However, those with complete and incomplete data had similar baseline variables except for a small difference in age, and sensitivity analysis revealed no significant impact on results due to missingness. Although we see strong associations in linear models, additional considerations, including the potential for non-linear effects, could be considered if future models are used for prediction. We were also unable to control for all potential unmeasured confounders and systematic differences in provider prescribing practices. For instance, information on the ED and post-ED interventions and treatment for each subject was not collected. However, given that assessment of depressive symptoms is not routinely used in ED pain management decisions, it is unlikely that the heterogeneity of treatment modalities led to systematic differences between those with fewer or more depressive symptoms. Another potential limitation is that the data were collected nearly 10 years ago. While the current opioid epidemic has begun to change medication prescribing practices, this is only a recent change with variable uptake nationwide. Furthermore, there has been little to no incorporation of psychosocial measures into standard ED pain management practice over the past decade, making these findings relevant to today’s practice. Finally, baseline pain scores were not recorded at the index visit, as patients were approached for study enrollment only after they were treated by an ED provider and their pain was controlled. The lack of association between baseline depressive symptoms and change in pain score may be due to the inability to use a true baseline pain score, because a similar lack of association was found in sensitivity analysis between baseline depressive symptoms and change in RMDQ from 1 week to 3 months.

CONCLUSION

In summary, this is the first analysis we are aware of that describes the temporal associations between severity and progression of depressive symptoms and 3-month pain and disability outcomes in ED patients with low back pain. Our results align with similar findings from other clinical settings including primary care, physical therapy, and osteopathic treatment. Collectively, these studies emphasize the importance of early recognition of depressive symptoms as a potential risk factor for the development of chronic pain and perpetuation of functional limitations. These results further describe how depressive symptoms may be associated with the development of chronic pain after an acute painful episode. Furthermore, these results also indicate the potential for a reciprocal association between persistent pain-related limitations and worsening depressive symptoms. ED clinicians continue to see large numbers of patients with both acute and chronic low back pain that are at increased risk for having persistent functional limitations after ED visits. Thus, future studies investigating the effectiveness of behavioral treatments delivered in the ED that target...
depressive symptoms, pain intensity, and functional outcomes are critically needed. Such treatments may have the potential to prevent the growing burden of chronic pain.

REFERENCES


Supporting Information

The following supporting information is available in the online version of this paper available at http://onlinelibrary.wiley.com/doi/10.1111/acem.13957/full

Data Supplement S1. Supplemental material.
Feasibility Study of a Quasi-experimental Regional Opioid Safety Prescribing Program in Veterans Health Administration Emergency Departments

Nathalie Dieujuste, BA1, Rachel Johnson-Koenke, LCSW1, Melissa Christopher, PharmD2, Elise C. Gunzburger, MS1, Thomas Emmendorfer, PharmD3, Chad Kessler, MD, MPH4, Jason Haukoos, MD, MSc5,6,7, Jason Smith, PharmD8, and Comilla Sasson, MD, PhD1,6,7

ABSTRACT

Background: The Veterans Health Administration (VHA) Opioid Safety Initiative (OSI) was implemented in 2013 and was associated with a 25% relative decrease in the dispensing of opioids. Although emergency department (ED) providers play a role in the initiation and continuation of opioids, the incumbent OSI did not target EDs.

Objective: The goal of this feasibility study was to leverage the existing VHA OSI and test a novel ED-based quality improvement (QI) program to decrease opioid prescribing in multiple ED settings.

Methods: This was a quasi-experimental study of phased-in implementation of a QI ED-based OSI. The general setting for this pilot were four VHA EDs across the Veterans Integrated Services Network (VISN) region 19: Denver, Oklahoma City, Muskogee, and Salt Lake City. We developed and disseminated a dashboard to assess ED-specific prescribing rates and an ED-tailored toolkit to implement the program. Academic detailing pharmacists provided focused audits and feedback with the highest prescribing providers. We measured change in ED-provider prescribing rate of opioids for patients discharged from the ED, by provider and aggregated up to facility level, pre- and postimplementation.

Results: Interrupted time-series analysis of provider-level data from the program implementation sites indicated a significant decrease in the trend for proportion of opioid prescriptions relative to the preintervention trend. The results of the analysis suggest that the intervention was associated with accelerating the rate at which ED provider prescribing rates decreased.

From the 1VA Eastern Colorado Health Care System, Veterans Health Administration; 2National Academic Detailing Services, Veterans Health Administration, San Diego, CA; 3National Pharmacy Benefits Management, Veterans Health Administration, Washington, DC; 4Durham VA Medical Center, National Emergency Medicine Program Office, Veterans Health Administration, Durham, NC; and the 5Denver Health and Hospital Authority, Denver, CO; and the 6Colorado School of Public Health, Aurora, CO; the 7University of Colorado School of Medicine, Aurora, CO; and the 8VA Eastern Colorado Health Care System, Veterans Health Administration, Denver, CO.

Received September 16, 2019; revision received March 3, 2020; accepted March 27, 2020.

The authors have no relevant financial information or potential conflicts to disclose.

Author contributions: CS and ND conceived the study and, together with JS and MC, designed the intervention; CS, CK, MC, JS, and RJK obtained proper institutional and stakeholder support to pilot the program; CS, ND, and JS managed the recruitment and program introduction of program implementation sites and their respective emergency department medical directors; JS and MC were responsible for oversight of academic detailing services to program implementation sites; TE undertook dashboard development; ND monitored data and conducted all data analyses; JH provided statistical advice on the study design and oversight of the data analysis; ECG completed the revised analysis of the data upon manuscript revision and contributed significantly to the revision of the manuscript results; ND, RJK, and CS drafted the manuscript; and all authors have contributed significantly to this work.

Supervising Editor: Zachary F. Meisel, MD, MPH, MSc.
Address for correspondence and reprints: Nathalie Dieujuste; e-mail: Nathalie.dieujuste@va.gov.

Conclusion: Due to the high volume of patients and the vital role the ED plays in patient treatment and hospital admissions, it is evident that the ED is an important site for QI programs as well as the implementation of opioid safety measures. Given the findings of this pilot, we believe that implementation of a national Veterans Affairs ED OSI implementation is feasible practice.

The opioid epidemic is rapidly changing the landscape in which emergency department (ED) providers prescribe opioids for acute and chronic pain conditions. Opioid prescriptions written in the ED/urgent care centers (UCCs) are significantly associated with long-term opioid use. Opioid-naïve patients who receive more than a 3-day supply of opioids or multiple opioid prescriptions are more likely to become chronic opioid users. As such, ED/UCC providers play an integral role in addressing the opioid crisis. As a result, more focus has been placed within the Veterans’ Health Administration (VHA) to administer nonopioid treatment options for pain management as well as decreasing variability in opioid prescribing in the ED.

The VHA is the largest national provider of health care in the United States, with greater than 2.5 million visits annually to VHA ED and UCC. Our prior study looked at opioid prescribing by ED/UCC providers for patients who were discharged from the ED. From October 2014 to June 2017, the national trend in median prescribing rates decreased by 25.5% from 9.1% to 6.4%. The greatest rates of decline occurred between January 1, 2016 and June 30, 2017. The rate of provider opioid prescribing demonstrated wide variability between facilities ranging from 0.5% to 39.1%. The prescribing rate for ED/UCC providers ranged from 0.2% to 100%. Between June 2016 and May 2017, a total 24 VHA ED/UCC providers were the highest opioid prescribers nationally in at least two of the four quarters (22%-70%), with rates two- to threefold higher than their peers. A recommendation for a focused initiative tailored for ED/UCC providers was made to National VA ED leadership to decrease opioid prescribing variability.

In 2013, the VHA launched the Opioid Safety Initiative (OSI), which focused on four broad areas: education, pain management, risk mitigation, and addiction treatment. By 2017, nearly 349,000 fewer opioid prescriptions were written (25% decrease) for veterans who were seeking care in VHA outpatient settings. The incumbent OSI, which included academic detailing, dashboards of provider prescribing patterns, and a toolkit with patient and provider resources was primarily conducted at the facility-level working with primary care providers. While these results were encouraging in primary care, there was still much work to be done in other settings, including EDs.

To date, there has not been a focused VA effort tailored to the unique needs of ED providers and addressing the challenges and key opportunities these providers face with respect to opioid prescribing. The goal of this feasibility study was to leverage the existing VHA OSI and test a novel ED-based quality improvement (QI) program to decrease opioid prescribing in multiple ED settings.

METHODS

Regulatory
The activities described in this article did not meet the definition for human subjects research as defined by our local institutional review board, the Colorado Multi-Institutional Review Board, or our local VA Office of Research and Development.

Study Design and Setting
This is a quasi-experimental study of phased-in implementation of an ED-based OSI program. Four sites were included in the implementation study. Rocky Mountain Regional VA Center (Denver) was the pilot site for Phase 1 of the intervention. Phase 2 sites included three additional ED in the Veteran Integrated Services Network (VISN) 19 region: Muskogee, OK; Oklahoma City, OK; and Salt Lake City, UT. All sites were compared to the national trends in prescribing for all 142 VHA ED/UCCs.

Population
Veteran Integrated Services Network 19, the Rocky Mountain Network, is composed of eight VA health care systems, including seven ED facilities. VISN 19 has one regional academic detailing service (ADS) director, who has two ADS clinical pharmacist specialists working in the VISN. Table 1 shows the characteristics of each site including level of complexity, number of annual ED visits for 2018, number of emergency medicine board-certified physicians, nurse practitioners/physician assistants, and the number of providers included in the analysis. VISN 19 was
chosen as the initial regional site for implementation due to the senior author working in the ED at a site within the region.

**Measurements**

A simple metric for opioid prescribing by ED providers was developed from pharmacy benefits management data for all ED/UCC providers (physicians, nurse practitioners, physician assistants) from October 1, 2015, to September 30, 2018. To remain consistent with the case definition of “opioids” for the VA OSI program, the following medications were considered opioids: butorphanol, codeine, fentanyl, hydrocodone, hydromorphone, levorphanol, meperidine, methadone, morphine, nalbuphine, oxycodone, oxymorphone, pentazocine, and tapentadol. Data were analyzed to assess national opioid prescribing rates and baseline variability in ED provider and facility level opioid prescribing rates. The ED provider prescribing rate was defined as the proportion of patients who received an opioid prescription (number of prescriptions written by each ED provider [numerator] divided by the number of total patients discharged by the ED provider [denominator]). Using the number of patients discharged from the ED as the denominator, we limited our data analysis to those ED providers who had seen, treated, and discharged a minimum of 100 patients. Providers were included in the analysis if they discharged at least 100 patients from the ED to exclude temporary coverage or providers who were rotating through the ED.

**Intervention Development of the ED-based Opioid Safety QI Program.** We worked closely with our national VHA partners from pharmacy benefits management, the office of emergency medicine, and ADS to develop our QI program. Within our program, academic detailing played an integral role as these clinical pharmacy specialists (PharmDs) provide prescription monitoring and outreach education for healthcare providers in the VHA. The Opioid Safety ED QI Program was a “bundled” approach with four steps: 1) sharing of opioid prescribing dashboard with ED medical director and academic detailer, 2) education of ED providers and implementation of toolkit resources; 3) academic detailer conducting audit and feedback.

### Table 1
Site Characteristics at Time of Implementation

<table>
<thead>
<tr>
<th>Site Name</th>
<th>Level*</th>
<th>Annual ED Visits 2018</th>
<th>Physicians</th>
<th>Nurse Practitioners</th>
<th>Physician Assistants</th>
<th>Total Providers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denver, CO</td>
<td>1a-high complexity</td>
<td>29,149</td>
<td>16</td>
<td>2</td>
<td>0</td>
<td>18</td>
</tr>
<tr>
<td>Salt Lake City, UT</td>
<td>1a-high complexity</td>
<td>20,754</td>
<td>16</td>
<td>0</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>Muskogee, OK</td>
<td>1c-high complexity</td>
<td>14,604</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Oklahoma City, OK</td>
<td>1b-high complexity</td>
<td>24,526</td>
<td>14</td>
<td>1</td>
<td>0</td>
<td>15</td>
</tr>
</tbody>
</table>

*Site complexity level. Level 1 (a, b, c) defined as high patient volume and patient risk, high levels of teaching/research (1a-level 5 ICUs; 1b-level 4 and 5 ICUs; 1c-level 4 ICUs). Level 2 defined as medium patient volume and risk, medium levels of teaching/research, and level 3 and 4 ICUs. Level 3 defined as low levels of patient complexity, low patient volume, little or no teaching/research, and level 1 and 2 ICUs.

**Figure 1.** Four steps of the ED OSI QI program. First, the ED medical director is engaged into the program by their local academic detailer or by the ED OSI program manager. Second, ED medical directors and academic detailers assume the responsibility of educating the sites’ ED providers. Third, academic detailers provide detailed audit-and-feedback discussions with the highest prescribing ED providers at the facility. Fourth, ED medical directors disseminate opioid prescribing rates with ED providers quarterly. OSI = opioid safety initiative; QI = quality improvement.
session(s) with highest prescribing providers, and 4) quarterly reporting of opioid prescribing dashboard data to ED providers (Figure 1).

Because there were no specific ED-tailored resources in the VHA OSI toolkit, we created the following resources for inclusion in the ED OSI toolkit: 1) PowerPoint presentation with overview of opioid epidemic and introduction to ED OSI toolkit; 2) one-page handout on alternative treatment options pathway for acute and chronic pain management in the ED (Data Supplement S1, available as supporting information in the online version of this paper, which is available at http://onlinelibrary.wiley.com/doi/10.1111/acem.13980/full); 3) patient education material on “stepwise approach to acute pain management” and “opioid use and misuse”; 4) National VHA EM/Pain Management Clinical Guideline on Pain Management in the ED; 5) sample ED clinical policy template; and 6) resource list for providers looking for additional information about opioids, substance abuse/use, and pain management. Participants in Phase 1 of the program were asked for feedback on toolkit resources using an iterative process for developing and amending ED OSI toolkit resources. Phase 2 participants were also asked for feedback on toolkit materials, as well as any other additional resources that may be needed to assist with implementation. The program received a VHA QI designation. As such, institutional review board review was not required.

Implementation Approach

Step 1: Sharing of ED OSI Dashboard With Medical Director and Academic Detailer.

Since 2015, the VHA has made an investment to require all VHA facilities to have trained ADS clinical pharmacists who can provide individualized, face-to-face ADSs to encourage clinicians to use evidence-based decision making for mental health and pain management treatments. More than 100 VA clinical pharmacists have been trained in audit and feedback techniques. Prior research has shown that the ADS program can be a powerful and effective tool in helping to change the way in which providers prescribe medications. During Step 1, ED OSI dashboard was shared with the ADS clinical pharmacist and the ED medical director in person. This initial discussion served as an opportunity for discussion of the data, external validity on the metrics being used to assess ED provider behavior, how to best present this information to the ED providers at the facility.

Step 2: Education of ED Providers and Implementation of Toolkit Resources.

Using the introductory PowerPoint presentation, ED providers were given an understanding of the opioid epidemic as well as how their own prescribing patterns compared to their colleagues at the ED facility, VISN, and national levels from October 1, 2015, to September 30, 2018. Additional education and training were provided using the ED OSI toolkit on alternative pain management strategies, patient resources and review of the ED clinical policies on pain management. Opportunities for discussion on potential reasons for variability in prescribing were also explored. Feedback was solicited by the study team on additional resources that may be needed by providers as well as any amendments to the current provider resource for alternative pain management options.

Step 3: Academic Detailer Conducts Audit and Feedback Session(s) With Highest Prescribing Providers.

Giving ED providers their own opioid prescribing data, compared to local, regional, and national benchmarks, coupled with focused feedback, has shown positive results in decreasing opioid prescriptions in community settings. The ADS clinical pharmacist identified highest prescribing providers at each site based on the numbers of opioid prescriptions written compared to the national average for VA ED providers and the facility-level ED provider mean for prescribing. The ADS pharmacist used audit and feedback techniques to help drive clinicians to the desired behaviors (e.g., using clinical practice guidelines for acute and chronic pain management, limiting days of opioids prescribed, educating patients on potential risks of starting or continuing opioids). The ADS clinical pharmacist worked with individual ED providers to discuss their specific opioid prescribing data and educate and provide resources on safer opioid-prescribing practices. In the bidirectional communication, a better understanding of organizational, patient-level, and external environment factors that may be influencing prescribing patterns and work collaboratively with the provider to strategize ways to overcome these barriers.

Step 4: Quarterly Reporting of Opioid Prescribing Dashboard Data to ED Providers.

Emergency department providers were shown how their own prescribing patterns compared to their colleagues at the ED facility, VISN, and national levels changed
quarterly. The decision to share data with providers in a blinded or unblinded fashion was left to the discretion of the ED medical director.

Outcome
Outcomes were feasibility of a QI program in the VHA ED and change in ED provider prescribing rate of opioids for patients discharged from the ED, by provider and aggregated up to facility level, across seven quarters.

Data Management and Analysis
Emergency department provider prescribing patterns were descriptively analyzed for a run-in period of 3 months prior to the intervention and then reassessed at 3-month intervals postimplementation. National temporal trends in opioid prescribing were also included. With Stata (version 13.0) and Excel (version 1808) a descriptive statistical analysis was conducted of the ED opioid prescribing rate of the four facilities pre- and postintervention periods. R (version 3.5.3) was used to conduct interrupted time-series methods with Poisson regression to analyze the change in opioid prescribing trend from the preintervention period through the postintervention.

RESULTS
The 21-month pilot of the ED tailored OSI program included four sites in VISN 19. Denver was the first program implementation site followed by three additional sites in the region: Salt Lake City, Oklahoma City, and Muskogee. We conducted a descriptive analysis of the median ED opioid prescribing rate of the four facilities pre- and postintervention. The initial average median ED opioid prescribing rate for the four implementation sites was 6.7%. By the end of the 21-month pilot phase, we saw a relative 47.2% decrease in the ED opioid prescribing rate of the four facilities after implementation of the ED tailored OSI program. Table 2 shows the median prescribing rates, range, and interquartile ranges for each site.

We used an interrupted time-series analysis with Poisson regression to examine the change in trend of percentage of opioid prescriptions in the program implementation sites. The analysis used provider-level data from the preintervention period and the intervention period across each quarter time period. The analysis showed that in the preintervention time period (FY16Q1 [October 2015–December 2015] to FY17Q2

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Denver</td>
<td>6.25 (0.83–13.01)</td>
<td>4.67 (0.61–22.90)</td>
<td>3.82 (0.18–28.62)</td>
<td>3.95 (0.17–30.07)</td>
<td>5.25 (0.15–30.24)</td>
<td>5.02 (0.20–33.35)</td>
<td>3.82 (0.18–28.62)</td>
<td>3.79 (0.21–33.50)</td>
</tr>
<tr>
<td>SLC</td>
<td>5.81 (1.01–23.02)</td>
<td>4.61 (0.23–18.60)</td>
<td>3.95 (0.23–18.60)</td>
<td>3.18 (0.11–11.98)</td>
<td>3.18 (0.11–11.98)</td>
<td>3.18 (0.11–11.98)</td>
<td>3.18 (0.11–11.98)</td>
<td>3.18 (0.11–11.98)</td>
</tr>
<tr>
<td>Muskogee</td>
<td>3.49 (1.15–13.50)</td>
<td>3.49 (1.15–13.50)</td>
<td>3.49 (1.15–13.50)</td>
<td>3.49 (1.15–13.50)</td>
<td>3.49 (1.15–13.50)</td>
<td>3.49 (1.15–13.50)</td>
<td>3.49 (1.15–13.50)</td>
<td>3.49 (1.15–13.50)</td>
</tr>
<tr>
<td>OKC</td>
<td>6.27 (0.61–22.90)</td>
<td>3.82 (0.18–28.62)</td>
<td>3.82 (0.18–28.62)</td>
<td>3.82 (0.18–28.62)</td>
<td>3.82 (0.18–28.62)</td>
<td>3.82 (0.18–28.62)</td>
<td>3.82 (0.18–28.62)</td>
<td>3.82 (0.18–28.62)</td>
</tr>
<tr>
<td>National</td>
<td>5.55 (0.26–33.35)</td>
<td>3.82 (0.18–28.62)</td>
<td>3.82 (0.18–28.62)</td>
<td>3.82 (0.18–28.62)</td>
<td>3.82 (0.18–28.62)</td>
<td>3.82 (0.18–28.62)</td>
<td>3.82 (0.18–28.62)</td>
<td>3.82 (0.18–28.62)</td>
</tr>
</tbody>
</table>

Data are median prescribing rate (range [interquartile range]) for each site over time.
(January 2017–March 2017) opioid prescribing rates were trending downward with a relative decrease of 0.98 per quarter (95% confidence interval [CI] = 0.97 to 0.99, p < 0.01). In that time period, the weighted mean opioid prescribing rate decreased from 11.3% in FY16Q1 to 8.5% in FY17Q2. During the 21-month, or seven-quarter, intervention period the weighted mean opioid prescribing rate decreased from 8.0% in FY17Q3 (April 2017 to June 2017) to 4.1% in FY19Q1 (October 2018 to December 2018). As such, the estimate for the change in the opioid prescribing rate trend from the pre to post period indicated a significant decrease in the proportion of opioid prescriptions relative to the preintervention trend. In the postintervention period, the rate of opioid prescribing decreased on average 0.87 times per quarter (95% CI = 0.84 to 0.89, p < 0.01). The results of the analysis suggest that the intervention was associated with increasing the rate at which ED provider prescribing rates decreased. Figure 2 shows the model of fit for the trend in opioid prescribing rates (actual vs. predicted) and the weighted mean opioid prescribing rates.

**DISCUSSION**

This is the first ED-based QI implementation study in the VA medical system, which serves 2.5 million veterans annually. Nationally, the implementation of QI programs in medical facilities is on the rise. Due to the high volume of patients and the vital role the ED plays in patient treatment and hospital admissions, it is evident that the ED is an important site for QI programs as well as the implementation of opioid safety measures. Similar QI initiatives have been implemented in other EDs in health care organizations outside of the VHA. These QI efforts have focused largely on the features of their interventions and specifically reducing opioid prescribing rates. However, as this was the first initiative of its kind in the VHA ED, we deemed it important to first assess the feasibility of such work in this specific organization and clinical setting as a primary outcome. Given the findings of this study, we believe implementation of a national VA ED OSI implementation is feasible practice. Future iterations of this program will need to better assess the implementation of this program and the impact of specific aspects of the intervention on opioid prescribing rates.

We hypothesize certain key factors contributed to the success of the program implementation sites at reducing individual provider opioid prescribing rates and reducing variability in opioid prescribing. First, information about provider-specific behaviors was central in initiating change. Providers were made aware of their own individual metric on opioid prescribing, compared to others in their peer group at the facility, VISN, and national level. For many of the ED providers, this was an eye-opening experience, because they were not aware of how they compared to others. In addition, academic detailing staff, working with the ED leadership, helped to build a collaborative team-based approach to identifying areas of improvement and potential change.

An additional factor in the success of the program was the development and dissemination of toolkits with easy-to-use resources including an alternative treatment options pathway and training for providers. By creating the one-page reference handout for ED providers on other options for pain management based on current best practices, providers had easy access to other options for caring for their patient’s acute and chronic pain needs.

A part of the ED OSI program implementation was to create resources for providers to facilitate discussions with and educate patients and patient advocates on the risks of opioids for pain management. This helped ED providers easily and consistently counsel patients on why certain therapies for pain management were or were not being utilized.
Finally, one other key component which aided in the success of this pilot was the inclusion of local program implementors in the form of clinical pharmacist. In addition to the ED medical director, who often had many other completing priorities, clinical pharmacists played a key role in the program implementation by providing academic detailing with audit and feedback to providers. Academic detailing inclusive of audit and feedback has been successful in the VA in several areas, thus it stands to reason that the addition of academic detailing in ED OSI efforts would aid in program success. However, further research will need to be conducted on how academic detailing can be utilized to launch a national ED OSI.

LIMITATIONS

This ED OSI pilot is not without its limitations. One such limitation is that sites included in the program implementation group voluntarily self-selected for the program based on the decision from of their facility’s ED medical director. These early adopter sites may be more equipped for change and willing to utilize new programs or were already involved with opioid safety. More research will need to be conducted to see how the opioid prescribing rate varies with national implementation.

Another potential limitation of this study was that the primary investigator is an ED provider in the Denver/Rocky Mountain Regional VA Medical Center. Physician champions within the clinical context in which a program is being implemented can significantly change the course of the implementation. However, in the program implementation sites, there was still a significant change observed in opioid prescribing by site. It may be the addition of both the academic detailer with the ED medical director may be enough to elicit change in opioid prescribing. More research will need to be conducted to see how this varies by site.

Similarly, another limitation of this pilot is that we did not measure how each site implemented the program, which of the program resources were used by providers, the frequency of dashboard utilization, what number of providers received the training, or the frequency or recipients of academic detailing. In addition, we did not measure the attitudes and beliefs of providers about opioid prescribing or opioid safety prior to or after program implementation. It is possible that opinions regarding opioid prescribing prior to program implementation influenced the ED providers’ opioid prescribing upon program implementation. In the next phase of the ED OSI program, we will attempt to better measure these factors to see which are predictive of successful implementation.

While the results of this study do suggest that this program was associated with accelerating the rate at which opioid prescribing decreased, another possible explanation for the observed findings is the Hawthorne, or observer, effect. This could be seen as a confounder or a contributor to the effect given that monitoring providers is part of the intervention. It is possible that awareness of the program, and particularly the monitoring of the ED-specific opioid prescribing dashboard by one’s ED medical director, may have given rise to this effect, thus influencing provider opioid prescribing behaviors and subsequent opioid prescribing rates. Therefore, a limitation of this study is that we did not account for this effect in our design or analysis. Future research on this program implementation should consider the measuring and accounting for the Hawthorne effect.

Our analysis was also limited as we used interrupted time-series methods with Poisson regression. We considered using the remaining three VISN 19 sites that did not implement the program as contemporaneous control sites (Grand Junction, Fort Harrison, and Cheyenne) and conducting a difference-in-difference analysis. However, after examining the data descriptively, we deemed that remaining three VISN 19 sites were inappropriate controls for difference-in-difference analysis methods due to the lack of parallel trends assumption in the pre period. As such, we were ultimately limited to conducting interrupted time-series analysis methods with the program implementation sites, which does not always account for secular trends in data. Given that this is a QI initiative in the VHA, future research could be considered using a pragmatic clinical trial approach to better utilize control sites and to adjust for secular trends.

Finally, we were limited to only capturing the prescriptions that were written from the ED by VA providers. We did not include milligram morphine equivalents to measure the total dose of opioid prescribed, nor did we not capture any handwritten prescriptions for opioids outside of the electronic prescribing system or any prescriptions that were filled outside of the VA system. It is possible that veterans chose to seek outside care or pay for their medications on their own at external pharmacies. We also did not
include prescription drug monitoring program data at this time to determine total amounts of opioids being prescribed to a veteran. Future research will need to be conducted to further evaluate the amount, dosing, and types of opioids being prescribed to veterans and by what providers.

CONCLUSIONS

The findings of this pilot study warrant further exploration and future research on a national implementation of the ED opioid safety initiative in Veterans' Health Administration facilities. In addition, further investigation regarding provider perceptions on barriers and facilitators to opioid safety and the ED opioid safety initiative program is needed. Bearing in mind the next steps to further increase opioid safety in the Veterans' Health Administration ED, another aspect of consideration is increasing naloxone (Narcan) distribution and the initiation of medication treatment for opioid use disorder for opioids in the Veterans' Health Administration ED. Further investigation into feasibility and acceptability as well as provider perceptions on barriers and facilitators into these additional areas of consideration are warranted. Finally, a stepwise implementation approach to iterate on supporting resources and toolkits for ED-tailored intervention on naloxone and medication-assisted treatment is necessary.

Expanding the ED opioid safety initiative program to continue to tailor opioid safety initiatives to VHA EDs nationally could have a significant impact on the lives of veterans suffering from chronic pain and opioid use and help increase more judicious prescribing of opioids for patients with pain. In addition, expansion of this program may help shape how ED providers treat and manage this significant health issue for the millions of veterans receiving care from VHA facilities.

References


Supporting Information

The following supporting information is available in the online version of this paper available at http://onlinelibrary.wiley.com/doi/10.1111/acem.13980/full

Data Supplement S1. Supplemental material.
The Characteristics and Effectiveness of Interventions for Frequent Emergency Department Utilizing Patients With Chronic Noncancer Pain: A Systematic Review

Charles K. Wong, MD, CCFP-EM, Connor M. O’Rielly, Braden D. Teitge, MD, CCFP, Robert L. Sutherland, Scott Farquharson, MD, Monty Ghosh, MBT MSc, MD, Helen L. Robertson, and Eddy Lang, MD, CCFP-EM

ABSTRACT

Background: Patients with chronic noncancer pain (CNCP) present unique challenges to emergency department (ED) care providers and administrators. Their conditions lead to frequent ED visits for pain relief and symptom management, and are often poorly addressed with costly, low-yield care. A systematic review has not been performed to inform the management of frequent ED utilizing patients with CNCP. Therefore, we synthesized the available evidence on interventional strategies to improve care-associated outcomes for this patient group.

Methods: We searched Medline, EMBASE, CINAHL, CENTRAL, SCOPUS, and Web of Science from database inception to June 2018 for eligible interventional studies aimed at reducing frequent ED utilization among adult patients with CNCP. Articles were assessed in duplicate in accordance with methodologic recommendations from the Cochrane Handbook for Systematic Reviews of Interventions. Outcomes of interest were the frequency of subsequent ED visits, type and amount of opioids administered in the ED and prescribed at discharge, and costs. Methodologic quality was assessed using the Cochrane Risk of Bias in Non-Randomized Studies of Interventions and Risk of Bias tools for nonrandomized and randomized studies, respectively.

Results: Thirteen studies including 1,679 patients met the inclusion criteria. Identified interventions implemented pain policies (n = 4), individualized care plans (n = 5), ED care coordination (n = 2), chronic pain management pathways (n = 1), and behavioral health interventions (n = 1). All of the studies reported a decrease in ED visit frequency following their respective interventions. These reductions were especially pronounced in studies whose interventions were focused around individualized care plans and primary care involvement. Interventions implementing opioid restriction and pain management policies were largely successful in reducing the amounts of opioid medications administered and prescribed in the ED.

Conclusions: Multifaceted interventions, especially those employing individualized care plans, can successfully reduce subsequent ED visits, ED opioid administration and prescription, and care-associated costs for frequent ED utilizing patients with CNCP.

From the 1Department of Emergency Medicine, Cumming School of Medicine; the 2Department of Community Health Sciences, Cumming School of Medicine; the 3Department of General Internal Medicine, University of Alberta, Edmonton; and the 4Cumming School of Medicine, University of Calgary, Calgary, AB, Canada.

Received November 25, 2019; revision received January 30, 2020; accepted February 5, 2020.

Presented at the Canadian Association of Emergency Physicians (CAEP) Annual Conference, Halifax, NS, Canada, May 2019; and the University of Calgary, Department of Emergency Medicine Research Day, Calgary, AB, Canada, April 2019.

Support provided through the 2018 Alberta Health Services Emergency Strategic Clinical Network (ESCN) Systematic Review Grant Competition. CW received a grant to perform this systematic review from the Alberta Health Services Emergency Strategic Network, which is a nonprofit organization that promotes knowledge translation in Alberta, Canada. Funding from this grant was utilized to support work from research associates CO and RLS. CO, BT, RLS, SF, MG, HLR, and EL have no relevant financial information or potential conflicts to disclose.

Author contributions: study protocol design—CW, SF, MG, HLR, and EL; article review and appraisal—CW, CO, BT, RLS; analysis and interpretation of the data—CO, CW; drafting of the manuscript—CW, CO; critical revision of the manuscript for important intellectual content—MG, RLS, EL; and acquisition of funding—CW.

Supervising Editor: Shellie Asher, MD.

Address for correspondence and reprints: Charles K. Wong, MD, CCFP-EM; e-mail: charleskhwong@gmail.com.

ACADEMIC EMERGENCY MEDICINE 2020;27:742–752.

doi: 10.1111/acem.13934
# BACKGROUND

Pain is the most common presenting complaint in the emergency department (ED), with patients seeking pain relief and management of other chronic pain-related symptoms accounting for an estimated 12% to 16% of all ED visits.1–3 Because of their high rates of ED utilization, patients with chronic noncancer pain (CNCP) are often categorized within the important subgroup of frequent ED users and recurrent ED service users.4,5 While the definition of a recurrent ED service user is not firmly established in the literature, they are generally considered to be patients who access the ED in disproportionate frequency.6 These patients are overrepresented in groups with low socioeconomic status and are more likely to have physical and mental health comorbidities.7 Within the chronic pain population, the prevalence of opioid use disorder (OUD) varies from 8% to 35%.8 Key characteristics for individuals with both chronic pain and OUD who utilize acute care resources like the ED include female sex, age greater than 65, and low annual household incomes and those with concurrent cannabinoid use disorder.9

Regardless of the drivers of ED visit frequency, the concomitant demands of a population of recurrent ED users with CNCP complaints and an increasing awareness of the potential harms of opioid treatment places ED physicians in a uniquely difficult position.10 While established guidelines do attempt to provide some guidance to emergency physicians in this area, decisions on how to medically manage this patient population are largely left to the physician’s discretion.11 The resulting variability in the ED management of CNCP may lead to uncoordinated, less effective, and higher-cost care in the ED compared to what these patients might receive in a primary care setting.7

While there is a wealth of evidence on the care of recurrent service users in the ED,12–14 there has not yet been a review of ED-based interventional strategies to optimize the management of the subpopulation presenting with acute exacerbations of their CNCP. Therefore, the purpose of this systematic review is to synthesize all available evidence on ED-based interventions aimed at improving the management of recurrent ED utilizing patients with CNCP. This review will aim to provide ED physicians, allied health care professionals, and administrators with the information required to minimize inefficient resource allocation and optimize the ED management of this complex patient subgroup.

# METHODS

The methods of this study were developed in reference to the Cochrane Handbook for Systematic Reviews of Interventions and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.15,16

## Study Design

We set out to perform a systematic literature review to answer the following questions: what are the characteristics of ED-based interventions targeted at patients who present frequently to the ED for chronic pain and what is the effectiveness of these interventions on ED visit frequency as well as other patient-, physician-, and system-important outcomes? Interventions were considered “ED-based” if implemented primarily within the ED, even if components of the intervention reached beyond the department (i.e., referral to outpatient pain clinic or taper-to-abstinence program). An a priori study protocol defined the review methods, search methodology, and analytical goals of the study. The protocol was registered to PROSPERO (CRD42018087030) prior to the commencement of the review.

## Study Protocol

### Search Strategy.

An electronic search strategy was developed by the investigators (CW, CO) and content experts (CS, SF, MG, EL) and further refined by a medical research librarian (HLR). The search strategy included combinations of subject headings, keywords, and synonyms delineating the target population and all types of ED interventions that we sought to identify (Data Supplement S1, Appendix S1, available as supporting information in the online version of this paper, which is available at http://onlinelibrary.wiley.com/doi/10.1111/acem.13934/full). Headings and keywords were adapted for use in each individual database.

We conducted comprehensive searches of Epub ahead of print, in-process and other nonindexed citations, as well as seven electronic databases (Ovid MEDLINE Daily, Ovid MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials, CINAHL, SCOPUS, and Web of Science) from database inception to June 16, 2018. Further, reference lists of the included studies were examined for any further relevant studies not captured in the literature searches.
Study Selection. Two reviewers (CO, BT) independently screened the titles and abstracts retrieved in the search to identify studies with potential relevance to the research question. No publication date, publication language, or country of publication constraints was applied to the search. Further inclusion criteria for eligible studies included: 1) studies must include adult (≥18 years) patients identified by the study investigators as frequent ED utilizers, 2) the frequent ED utilization of included patients must be driven primarily by chronic pain complaints, 3) studies must have implemented (i.e., experimental studies) or evaluated (i.e., observational studies) ED-based interventions aimed at optimizing the management of frequent utilizers with chronic pain, and 4) studies must have assessed at least one of the outcomes described below. Studies examining only patients with sickle cell pain crises were determined a priori to be excluded, as it was felt by the content experts (MG, EL, CS, SF) to be unrepresentative of other chronic pain syndromes.

It should be noted that no established definition of CNCP was used, but in general the representative patient is one who presents with pain of a recurrent nature, for which no acute organic cause is identified over multiple physician assessments. Additionally, due to large variations in the classification of “frequent ED visits” throughout the available literature no fixed definition of the term was used.

All of the potentially relevant studies identified by the reviewers in the title and abstract screening underwent secondary independent full-text review to ensure applicability to the research question. Those meeting all inclusion criteria were included in the final review. A linearly weighted Cohen’s kappa (κ) statistic and corresponding 95% confidence interval (CI) were calculated following the title and abstract screening phase to quantify the level of agreement between the independent reviewers. At both the title and the abstract review as well as full-text review stages, discrepancies were reconciled by group discussion mediated by an unbiased third reviewer (CW).

Outcomes of Interest. The primary outcome for this study was the change in the frequency of ED visits following the implementation of an intervention. Secondary outcomes included postintervention changes in the amount and type of opioids administered in the ED, amount and type of opioids prescribed at ED discharge, total care-associated costs (e.g., physician, pharmacy, hospital bills), and adverse events (e.g., overdoses related to illicit opioids, all-cause mortality).

Study Quality (Risk-of-bias) Assessment. Two reviewers (CO, RLS) independently assessed and rated the methodologic quality of each study. Randomized controlled trials (RCTs) were assessed using the Cochrane Risk of Bias 2.0 (RoB 2.0) tool. Nonrandomized studies were assessed using the Cochrane “Risk of Bias in Non-Randomized Studies of Interventions” (ROBINS-I) tool. Discrepancies between the independent reviewers were again resolved by consensus through adjudication by an unbiased third party (CW). The level of inter-rater agreement for the study quality assessment was quantified using a linearly weighted κ statistic with a corresponding 95% CI.

Data Extraction. Data from the included studies were extracted by a trained reviewer (CO) using a standardized data extraction form that was developed in reference to the Cochrane Collaboration data collection form for intervention reviews: RCTs and non-RCTs (see Appendix S2). The following data were abstracted: authors, country of origin, study design, the operational definition used for frequent ED use, characteristics of the study sample (i.e., sample size, demographic information), type and description of the implemented intervention, and effect of the intervention on ED visit frequency and other relevant outcomes.

Data Synthesis and Analysis. Descriptions of the included studies, patients, and interventions were presented in tables. Additionally, all of the synthesized outcomes data were assessed using a semiquantitative approach. This involves a tabular presentation of the direction of change in the outcomes of interest following the implementation of the intervention as well as a written description of the magnitude of these changes. Statistical significance of intervention-related differences was defined by a two-tailed p-value < 0.05.

The protocol for this project outlined our intention to perform a meta-analysis in addition to the narrative and semiquantitative descriptions presented here. However, early in the review process it was determined that variations in the definitions used in the studies to define the chronic pain and frequent ED utilization populations introduced large amounts of clinical heterogeneity, thus rendering any meta-analysis inappropriate.
RESULTS

Search Results
In total, 7,240 citations were identified from the electronic search. Following the removal of duplicates (n = 2,222), we retrieved 5,018 unique records. Title and abstract review yielded 162 articles for full-text review. The level of agreement between the independent reviewers following title and abstract review was graded as fair (κ = 0.52 [95% CI = 0.34 to 0.69]). After the independent secondary full-text review, 13 studies were used ranging from N = 22 to 745 ED visits in a month. Postintervention follow-up time ranged from 12 months to 39 months. We identified a high risk of bias due to confounding given the lack of a formal control group or a lack of procedures to control at the analysis stage. A more detailed summary of the risk of bias assessments is shown graphically in Figure 2. In nine of the 13 included studies we identified a high risk of bias due to confounding given the lack of a formal control group or a lack of procedures to control at the analysis stage.

Study and Patient Characteristics
As detailed in Table 1, a total of four RCTs19–22 and nine noncontrolled before–after studies (NCBAS)23–31 were included, involving the assessment of 609 patients and 1,070 patients, respectively (total N = 1,679). The sample sizes of the RCTs ranged from 40 to 406 patients while the sample sizes for the NCBAS were slightly smaller ranging from 14 to 314. All but one of the studies were completed in the United States, with one of the NCBAS being conducted in Canada. The included studies were published between 2007 and 2018. The majority of the studies focused on patients with CNCP complaints alone, while some expanded their definitions to include those with CNCP and comorbid problematic substance use disorders. The reported mean ages for the studies ranged from 36.2 to 47.9 years, with female patients making up 35.7% to 78.0% of the included samples. A total of seven different operational definitions of “frequent ED utilization” were used ranging from ≥4 ED visits in 12 months to ≥3 ED visits in a month. Postintervention follow-up time ranged from 12 months to 39 months.

Intervention Characteristics
All of the included studies examined interventions that were ED-based. We identified four common intervention components: electronic medical record (EMR) alerts, primary care contact and referral, individualized care plans or care pathways, and departmental opioid restriction policies. EMR alerts typically took the form of a prompt when opening an enrolled patient’s chart and were implemented to advise ED practitioners that their patient had been recruited for that particular study and often included details about any care plans or protocols involved in their care. Primary care contact involved referral either to a family physician or to the participating ED informing the existing community provider about their patient’s enrollment in the study. Care plans involved the development of individualized treatment plans or treatment pathways in collaboration with multidisciplinary care teams aimed at optimizing the management of individual patients. Opioid restriction policies ranged from prescription pathways aimed at reducing the administration of opioids in the ED to contracts that eliminated all opioid administration and prescriptions in the ED for recruited patients.

Eight of the included studies implemented multifaceted interventions that included two or more of the identified components,19,20,22–26,30 and the remaining five studies implemented interventions that were focused around a singular component.21,27–29,31

Study Quality Assessment
Four randomized studies were assessed using the Cochrane RoB 2.0 tool and nine nonrandomized studies were assessed using the ROBINS-I tool. The two independent reviewers (CO, RLS) assessed study quality with moderate overall agreement (κ = 0.51 [95% CI = 0.35 to 0.67]). The overall study quality was moderate for the RCTs and low for the NCBAS. However, most concerns with study quality were concentrated in the domains of bias due to confounding, deviations from the intended interventions, and selection. A more detailed summary of the risk of bias assessments is shown graphically in Figure 2. In nine of the 13 included studies we identified a high risk of bias due to confounding given the lack of a formal control group or a lack of procedures to control at the analysis stage.23–31 We identified a high risk of bias due to participant selection in three of the 13 studies.23–25 These concerns were most pronounced in the retrospective studies that participant selection for analyses occurred after the intervention had been implemented. Additionally, we found a high risk of bias in two of the 13 included studies that were secondary to higher rates of participant attrition.29,31 There was minimal concern with study quality in the domains related to missing data, measurement, or selection of the reported result.
Outcomes
The number of annual or monthly visits that qualified as “frequent ED use” varied widely across the included studies and, as such, no pooled estimates or meta-analysis for any of the relevant outcomes could be calculated. The direction of the changes that occurred for each outcome secondary to the interventions in each study is shown in detail in Table 2, and the magnitude of these changes are described below in more detail.

**Frequency of ED Visits.** The mean number of preintervention ED visits per individual reported in the included studies ranged from 6.88 visits per year to 47.8 visits per year. The number of ED visits per individual following the implementation of an intervention ranged from 2.0 visits per year to 16.6 visits per year. This corresponded to a reduction in visits by between 48.4 and 89.5%. Each of the studies that measured the frequency of ED visits (n = 12) reported a decrease in the number of visits following the implementation of an intervention.\(^{19-26,28-31}\) Nine of the studies reported statistically significant reductions in the frequency of ED visits (from p < 0.0001 to p < 0.05),\(^{19,22,23,26-31}\) one reported a nonsignificant reduction (p = 0.68),\(^{21}\) and two studies did not report a p value.\(^{19,24}\)
<table>
<thead>
<tr>
<th>Study</th>
<th>Study Setting</th>
<th>Frequent Use Definition</th>
<th>Intervention Type, Description</th>
<th>Age (Mean)</th>
<th>Females (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Noncontrolled Before-After Studies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alburaih, 2018(^{23})</td>
<td>Country: United States ED type: Rural (n = 314)</td>
<td>≥10 visits/year</td>
<td>Pain contract</td>
<td>48 (median)</td>
<td>185 (58.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Patients signed hard copy of pain contract that informed them that they would no longer receive controlled substances for their chronic pain in the ED;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- EMR-based notification to reinforce contract application.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kahler, 2017(^{24})</td>
<td>Country: United States ED type: Urban (n = 243)</td>
<td>≥6 visits/year</td>
<td>Opioid restriction policy</td>
<td>41.0</td>
<td>151 (62.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Patients referred to taper-to-abstinence pain management clinic run by a pain management and addiction specialist;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Patients informed they would no longer routinely receive opioid medications in the ED;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- EMR-based notification to reinforce the policy application.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Masterson, 2012(^{25})</td>
<td>Country: United States ED type: Urban (n = 134)</td>
<td>≥3 visits/m, ≥2 visits/m in 2 consecutive months, ≥6 visits/y</td>
<td>Pain care management program</td>
<td>39.8</td>
<td>84 (62.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Pain care program developed collaboratively with physicians, nurses, social workers, occupational health practitioners, and addiction and pain specialists;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Goals of program were to coordinate ED care with plans of primary care physicians and prescribe only nonnarcotic analgesics whenever possible.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olsen, 2016(^{26})</td>
<td>Country: United States ED type: Urban (n = 46)</td>
<td>≥3 visits/6 months, ≥6 visits/year</td>
<td>Pain protocol</td>
<td>39.9</td>
<td>32 (69.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Contact primary care or pain specialist to review primary care history and develop treatment plan for future ED visits;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Plan typically precluded the use of narcotics or benzodiazepines in the ED;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- EMR-based notification to reinforce the protocol application.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pace, 2018(^{27})</td>
<td>Country: United States ED type: Urban (n = 266)</td>
<td>NR</td>
<td>Opioid prescribing pathway</td>
<td>43.4</td>
<td>167 (62.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Encourage the use of nonopioid analgesics;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Encourage oral routes over IM or IV;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Discourage the replacement of lost, stolen, or forgotten prescriptions;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Encourage short course opioid standard of &lt; 15 tabs hydrocodone.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passarello, 2015(^{28})</td>
<td>Country: United States ED type: Urban (n = 14)</td>
<td>≥20 visits/year</td>
<td>Pain protocol</td>
<td>47.9</td>
<td>5 (35.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Recommended oral as opposed to IV or IM opiates for the treatment of chronic pain.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rash, 2018(^{29})</td>
<td>Country: Canada ED type: Urban (n = 14)</td>
<td>≥12 visits/year</td>
<td>Individualized care plan</td>
<td>36.2</td>
<td>7 (50.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Patients undergo comprehensive assessment by an interprofessional pain assessment team;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Results discussed with patients, primary care physicians, and ED physicians to collaboratively develop a comprehensive treatment plan;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Treatment plan uploaded to EMR for reference and application of changes as needed.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Svenson, 2007(^{30})</td>
<td>Country: United States ED type: Urban (n = 14)</td>
<td>≥10 visits/year</td>
<td>Pain protocol</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Patients informed that they would no longer receive parenteral narcotics in the ED;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Letter sent to personal physician informing them of this decision;</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- Encouraged patients to see personal physician to discuss alternative therapies.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
Opioids Administered in the ED. The studies that reported this outcome \((n = 3)\) assessed the ED administration of hydromorphone, oxycodone, hydrocodone, tramadol, fentanyl, and morphine.\(^{21,27,28}\) Two of the studies reported a decrease in the total administration of opioid medications in the ED, with one study reporting a statistically significant reduction \((p < 0.001)\).\(^{28}\) One study reported no significant change in the amount of opioid medication administered in the ED \((p = 0.6659)\).\(^{27}\) Preintervention opioid administration amounts visit ranged from 10.3 morphine equivalents \((\text{MEQ})\) per patient visit to 17.2 MEQ per patient visit. The opioid administration amounts following intervention implementation ranged from 2.7 MEQ per patient visit to 10.68 MEQ per patient visit.

Opioid Prescriptions at Discharge. The types of opioid medications considered in the analyses of the studies that measured this outcome \((n = 5)\) were hydromorphone, oxycodone, hydrocodone, tramadol, and codeine.\(^{20,22,24,26,27}\) The method of measurement varied widely across studies with some reporting pill count, total number of prescriptions, or proportion of patients who received an opioid prescription. Regardless of the unit of measurement, all studies reported a statistically significant reduction in the number of opioid prescriptions following the implementation of their respective intervention \((p < 0.05 \text{ to } p < 0.001)\). One study did not assess the statistical significance of their result.\(^{24}\) A single included study stratified on the type of opioid being prescribed and only saw a statistically
significant decrease in the number of hydrocodone prescriptions (p < 0.05), but not in other opioid types.

**Total Costs.** All of the included studies that reported on the total cost outcome (n = 4) were completed in the United States. These studies all reported a decrease in the total care-associated costs following the implementation of their respective interventions. They reported preintervention costs for individual patients that ranged from a mean of $800 per visit, to median of $42,305 per year. Not all of the studies reported postintervention costs, but the reductions in cost ranged from cumulative individual cost differentials of −$3,200 per patient per year, to total estimated cost reductions of $204,000 across all patients in the year following the implementation of their intervention. The probability value of these cost reductions ranged from p = 0.88 to p = 0.007. One included study did not assess for statistical significance.

**DISCUSSION**

The purpose of this study was to identify and evaluate ED-based interventions that aimed to optimize the management of patients with CNCP who accessed EDs with high frequency. We found 13 low- to moderate-quality studies indicating that the implementation of a structured intervention within the ED can reduce the frequency of ED visits among high-utilizing patients with CNCP. Additionally, these interventions can reduce care-associated costs as well as the administration and prescription of common opioid drugs like hydromorphone, oxycodone, hydrocodone, and tramadol. The results of the four higher-quality RCTs and the nine nonrandomized studies included in this review reported similar results across all of the measured outcomes.

The dominant finding that interventions targeting frequent utilizers can lead to a reduction in the frequency of ED visits is consistent with previously completed systematic reviews and meta-analyses of the general frequent ED utilizing population. While these reviews did not isolate individuals with CNCP, it is known that patients with primary complaints of chronic abdominal, chest, and back pain can make up > 15% of the general frequent utilizer population. As such, it is likely that these completed reviews and meta-analyses included many individuals with CNCP.
This study also found that interventions targeting CNCP patients with frequent ED utilization can reduce care-associated costs (i.e., costs secondary to ED length of stay, hospitalization, and resource or imaging use) as well as opioid administration and prescription. These findings are in line with the available evidence from studies of the general frequent ED utilizer population.\textsuperscript{34-38}

The intervention modalities seen in the included studies were grouped into four distinct categories: EMR alerts (regarding patient enrollment status, existence of a care plan), primary care contact and referral, individualized care plans, and opioid policies. Contrary to previous research suggesting that singularly focused interventions yielded better patient outcomes,\textsuperscript{39,40} large-scale (i.e., >60%) and statistically significant reductions in ED visit frequency were more common in the eight multifaceted interventions in this study compared to those using a single modality. The findings from this qualitative synthesis suggest that a multipronged approach may be more effective than a single intervention.

LIMITATIONS

We were unable to complete a meta-analysis of our primary outcome of interest due to clinical and methodologic heterogeneity related to variations in the number of monthly or annual visits that qualified their participants as “frequent ED utilizer” and differences in the length of available follow-up and period of assessment.

Only one study mentioned monitoring their participants’ charts for adverse events, but this study still did not report this as an outcome.\textsuperscript{19} The remaining studies did not report such events (i.e., opioid-related overdoses, mental health admissions) and often failed to report any health care utilization beyond their own EDs. As such, our findings on the effects of the interventions on ED visit frequency cannot be generalized to the surrounding primary, walk-in, or outpatient care settings. Further, no conclusions can be drawn about how the interventions might have altered or improved the actual health of participants beyond the treating ED.

The fact that all but one of the included interventions reported strong positive outcomes should be approached with caution, due to a likely contribution of a publication bias. Unfortunately, the differing measures of effect reported by many of the included studies were not conducive to a funnel plot or formal analysis of publication bias. The administrative nature of system-based interventions may very well exacerbate the known difficulties of publishing negative intervention studies. It is possible, if not probable, that the average effect size of the studied interventions is quantitatively lower in reality than described in the included studies. Notably, only studies from the United States and Canada were identified, and as a result these findings may be less generalizable to other countries.

This study was also limited by the overall low to moderate methodologic quality of the included studies. Most concerns with bias were due to potential confounding and resulting lack of adjustment in the non-randomized interventional studies. Any bias secondary to these concerns was likely to cause an overestimation of intervention effectiveness, such that we cannot know how the ED visit frequencies of patients would have varied without any intervention (due to lack of formal control groups) nor can we be confident that changes were due solely to the intervention in question (in cases of deviation from the intervention).

FUTURE DIRECTIONS

Consistent with previous research in the field,\textsuperscript{41,42} results from this study underline the high variability in the definition of frequent ED utilization as well as variable study methodologies and quality of reporting. Consensus on definitions and more consistent outcomes reporting is required to move this field of study forward to generate more precise conclusions. Future research should aim to produce high-quality evidence with a focus on the impacts of ED-based interventions beyond the walls of their own departments such as in surrounding walk-in clinics. Future studies should also look at long-term outcomes to see if the observed improvements are sustained over time. Further, studies should examine some specific health outcomes of individuals enrolled in these ED-based strategies including adverse events. Researchers, journals, and reviewers should encourage the publication of “negative” studies to mitigate the risk of publication bias in future reviews.

CONCLUSION

The evidence synthesized in this rigorous systematic review suggests that the implementation of ED-based
interventions can reduce the frequency of ED visits, care-associated costs, amounts of opioid administration, and amounts of opioid prescription in a high ED utiler population with chronic noncancer pain. Future research should follow patients beyond the ED to ensure that reductions in ED utilization do not correspond to increased visits to surrounding health care locales or higher rates of adverse events.

References


Supporting Information

The following supporting information is available in the online version of this paper available at http://onlinelibrary.wiley.com/doi/10.1111/acem.13934/full

Data Supplement S1. Supplemental material.
Interrupted Time Series of User-centered Clinical Decision Support Implementation for Emergency Department-initiated Buprenorphine for Opioid Use Disorder

Wesley C. Holland¹, Bidisha Nath, MD, MPH², Fangyong Li, MPH, MS³, Kaitlin Maciejewski, MS³, Hyung Paek, MD⁴, James Dziura, PhD³, Haseena Rajeevan, PhD³, Charles C. Lu, MS³, Liliya Katsovich, MS, MBA³, Gail D’Onofrio, MD, MS², and Edward R. Melnick, MD, MHS²

ABSTRACT

Objectives: Adoption of emergency department (ED) initiation of buprenorphine (BUP) for opioid use disorder (OUD) into routine emergency care has been slow, partly due to clinicians’ unfamiliarity with this practice and perceptions that it is complicated and time-consuming. To address these barriers and guide emergency clinicians through the process of BUP initiation, we implemented a user-centered computerized clinical decision support system (CDS). This study was conducted to assess the feasibility of implementation and to evaluate the preliminary efficacy of the intervention to increase the rate of ED-initiated BUP.

Methods: An interrupted time series study was conducted in an urban, academic ED from April 2018 to February 2019 (preimplementation phase), March 2019 to August 2019 (implementation phase), and September 2019 to December 2019 (maintenance phase) to study the effect of the intervention on adult ED patients identified by a validated electronic health record (EHR)-based computable phenotype consisting of structured data consistent with potential cases of OUD who would benefit from BUP treatment. The intervention offers flexible CDS for identification of OUD, assessment of opioid withdrawal, and motivation of readiness to start treatment and automates EHR activities related to ED initiation of BUP (including documentation, orders, prescribing, and referral). The primary outcome was the rate of ED-initiated BUP. Secondary outcomes were launch of the intervention, prescription for naloxone at ED discharge, and referral for ongoing addiction treatment.

Results: Of the 141,041 unique patients presenting to the ED over the preimplementation and implementation phases (i.e., the phases used in primary analysis), 906 (574 preimplementation and 332 implementation) met OUD phenotype and inclusion criteria. The rate of BUP initiation increased from 3.5% (20/574) in the preimplementation phase to 6.6% (22/332) in the implementation phase (p = 0.03). After the temporal trend of the
An estimated 2.1 million people nationally suffer from opioid use disorder (OUD), contributing to nearly 50,000 overdose deaths each year.\(^1,2\) With 605,000 opioid-related emergency department (ED) visits in 2011 and a 30% increase in visits for opioid overdose-related visits between 2016 and 2017, the ED is a major and increasingly utilized setting for OUD treatment.\(^3,4\) People with OUD not only seek emergency care in high-acuity situations like overdose and withdrawal, but also for comorbid or general health issues.\(^5\) Given the stigma associated with OUD, the ED may serve as the primary access to health care for this vulnerable patient population.\(^5,6\) Thus, the ED provides a unique opportunity to initiate appropriate treatment for OUD.\(^7\)–\(^9\)

Opioid agonist medications, such as buprenorphine/naloxone (BUP) and methadone, are the current standard of treatment for OUD and have been shown to reduce withdrawal symptoms, craving, relapse, overdose, and mortality (all cause and opioid related).\(^10\)–\(^12\) A 2015 randomized clinical trial involving 329 ED patients with OUD demonstrated that BUP can be safely initiated in the ED and demonstrated that patients receiving BUP in the ED were twice as likely to remain engaged in formal addiction treatment at 1 month (78% vs. 37%, \(p < 0.001\)).\(^13\) Despite this evidence, BUP initiation in the ED has been slow to be adopted into routine emergency care to replace the current standard of care that historically has included symptomatic treatment for opioid withdrawal symptoms and referral for addiction treatment without addressing the underlying disorder.\(^5,19\) Furthermore, the rate of naloxone prescription upon ED discharge following nonfatal overdose remains low even as an evidence-based practice known to decrease mortality and risk of future overdose.\(^11\),\(^14\) Just as the ED is a unique setting to increase rates of BUP initiation, it is also an opportunity to implement other harm reduction strategies such as naloxone prescribing.\(^15\)

Numerous patient-side barriers currently limit the adoption of BUP initiation, including confusion and cultural stigma surrounding medication therapy for OUD and patient perceptions that such treatment is harmful, inferior to detoxification, and even incompatible with being truly “drug-free.”\(^16\)–\(^18\) The lack of adoption of ED initiation of BUP into routine emergency care has been attributed to emergency clinicians’ lack of training in addiction treatment and perception that BUP initiation is unfamiliar, complicated, and time-consuming.\(^5\),\(^19\) One potential solution previously shown to provide effective guidance for drug therapy is clinical decision support (CDS), computerized systems that provide patient-specific guidance.\(^20\)–\(^22\)

To address these barriers to implementation and to simplify the practice of ED-initiated BUP, we developed a user-centered CDS called EMBED (EMergency department-initiated BuprenorphinE for opioid use Disorder).\(^23\) To assess the feasibility of implementation and to evaluate the preliminary efficacy of the intervention to increase adoption of ED initiation of BUP, it was implemented in a single ED as the intervention in an interrupted time series study. The lessons learned from this study, particularly the qualitative feedback regarding intervention improvement, can be applied to a subsequent pragmatic group randomized trial involving 20 EDs across five health care systems. This multisystem pragmatic trial will determine the effectiveness of the EMBED intervention on the adoption of ED-initiated BUP for OUD.\(^24\)

**METHODS**

**Study Design and Setting**

A single-site time series study evaluating the preliminary efficacy of the EMBED intervention was conducted during April 2018 to December 2019 in an urban, academic Level I trauma center ED with 103,000 annual patient visits. The time series was divided into three phases for analysis: 1)
preimplementation phase (April 2018–February 2019), 2) implementation phase (March 2019–August 2019), and 3) maintenance phase (September 2019–December 2019). The phased rollout of the CDS intervention began with a soft go-live in mid-January 2019. Users were then made aware of the CDS’s availability when full functionality was achieved in early March 2019. Provider feedback on the CDS was collected in the implementation phase to assist with planning the subsequent trial. The study protocol was reviewed and approved by our institutional review board (Protocol 2000022749).

**Subjects**

Eligible patients were adult ED patients meeting the criteria of a computable phenotype derived from electronic health record (EHR) data that was developed to capture ED patients likely to have OUD and not actively on medication for OUD (MOUD, i.e., methadone, BUP, or naltrexone). The phenotype is comprised of two algorithms: one based on clinician and billing codes (Algorithm 1) and the other based on structured EHR data of the chief complaint (Algorithm 2). Additionally, the phenotype excludes patients who are admitted to the hospital or pregnant. In this way, the phenotype was designed to maximize specificity in identification of patients eligible for BUP initiation. Development, internal validation, and external validation of the phenotype occurred across two large health care systems containing 13 EDs. The phenotype has an externally validated positive predictive value of 0.95 and a negative predictive value of 0.92. A waiver of informed consent was obtained given that data were only collected retrospectively and did not involve patient interaction or identifiable information. Regarding clinician subject inclusion criteria, attending ED physicians practicing at the intervention site who cared for the phenotype-positive patients were eligible for inclusion. Given the additional burden of the consent process and to ensure the validity of the intervention’s efficacy on changing routine care, clinician demographic information was not collected. Therefore, all patient and physician identifiers were removed from EHR data by an honest broker and not shared with the investigative team. As a result of this deidentification process, it was not possible to match physician study data to our faculty roster to determine which physicians had an X-waiver to prescribe BUP. Therefore, a separate emergency medicine faculty roster was used to determine the proportion of physicians with an X-waiver (Figure 1).

**Intervention**

The study intervention included an integrated Web application for decision support and automation of EHR workflow that streamlines the practice of initiating BUP in the ED (Figure 2). Full details of the intervention’s design and IT integration have been previously reported. Briefly, the intervention is launched at the clinician’s discretion for patients who they suspect may have OUD by clicking the “EMBED” button on the navigation bar of a patient’s chart (the phenotype did not flag or alert clinicians of OUD cases). This opens a Web application within the EHR that offers three optional decision support tools to inform clinicians’ selection of the appropriate care pathway through the diagnosis of OUD, assessment of withdrawal severity, and motivation of patient readiness to start treatment for OUD. The clinician can choose to use all, some, or none of these tools. Once the care pathway is selected, the Web application automates a series of EHR activities specific to that pathway including appropriate orders, prescriptions, documentation, discharge instructions, and referral to a community provider of MOUD. Early audits in the implementation period identified low intervention use despite targeted e-mail communication and group lectures.

To enhance use, midway through the implementation period, 5-minute one-on-one tutorials were performed by author WCH to provide just-in-time training as an additional academic detailing component of the intervention. Compared to CDS alone, academic detailing combined with CDS has been shown to increase use when introducing a new CDS. Used in both industry and medicine and usually occurring in a real-time work setting, just-in-time training is an approach that involves presenting relevant information for immediate application. Clinicians were compensated with a $10 Starbucks gift card for their participation in the tutorial and a brief follow-up interview (Data Supplement S1, Appendix S1, available as supporting information in the online version of this paper, which is available at http://online library.wiley.com/doi/10.1111/acem.14002/full). A convenience sample of physicians completing the tutorial was utilized to gather qualitative feedback via semistructured interview questions. A formal qualitative analysis was not performed, but feedback on the
CDS obtained through these interviews was synthesized and categorized according to common, recurring themes as displayed in Data Supplement S1, Appendix S2.

Outcomes
The primary outcome of this study was BUP initiation rate in the ED, defined as whether or not an eligible patient was administered BUP in the ED and/or prescribed BUP on discharge from the ED. Secondary outcomes to evaluate preliminary efficacy of the intervention’s implementation included the attending physician adoption rate of the practice of ED initiation of BUP at least once in the study phase as well as the following patient-level rates in the cascade of care for treatment among eligible patients:27,32 1) launch of the intervention, 2) referral to follow-up for ongoing MOUD treatment as documented in the EHR, and 3) prescription for naloxone at ED discharge.

The RE-AIM (reach, effectiveness, adoption, implementation, and maintenance) framework was used to evaluate the success of the intervention.33–35 The reach of the intervention—the proportion of the target population that participated in the intervention—was the proportion of unique attendings who ever launched the intervention. Effectiveness was assessed based on the rate of BUP initiation in the ED during the implementation period (the primary outcome). Adoption was based on the proportion of unique attendings who initiated BUP, and Implementation—the extent to which the intervention is implemented as intended—was the proportion of phenotype-positive patients for whom the intervention was launched. Maintenance was the rate of BUP initiation in the ED after training ended (September–December 2019).

Data Collection
Quantitative outcome data from the three study periods were extracted from the study site’s EHR database using structured query language (SQL). The SQL query (Data Supplement S2, Appendix S3) included all data elements specified in the master data dictionary created for the subsequent trial (Data Supplement S2, Appendix S4).

Additional qualitative and quantitative data regarding clinician perceptions of barriers to implementation...
and usability of the intervention were collected via one-on-one interviews with clinicians (attendings, residents, APRNs, and PAs) performed by author WCH in the ED. The interview consisted of three parts: 1) system usability scale (SUS), 2) net promoter score (NPS), and 3) three open-ended questions focusing on barriers to implementation and ways to address them (Data Supplement S1, Appendix S1). The SUS is a 10-item usability assessment that is widely used and considered the industry standard for rapid assessment of health IT usability; a score of >70 is considered “acceptable.” NPS is a single-item measure of how likely an individual would be to recommend a product, company, or service to a friend or colleague.

**Data Analysis**

Patient characteristics were summarized as means and standard deviations (SDs) or frequencies and percentages as appropriate for the preimplementation and implementation periods. For patients with multiple visits, only the first visit was used to analyze patient-level outcomes. However, analysis of physician-level outcomes (i.e., BUP initiation, launching the intervention) included multiple visits made by a single patient. T-tests and chi-square or Fisher’s exact tests were used to make unadjusted comparisons of demographics, primary outcome rates of BUP initiation and secondary patient outcomes between periods. McNemar’s test was used to compare the number of unique attendings who initiated BUP who were present during both the preimplementation and the implementation phases, while the generalized estimating equation method was used as a supportive analysis of all attendings (i.e., including those not present in both time periods). Following unadjusted analysis, a multivariate logistic regression model was used to adjust for age, race, sex, the number of waivered physicians, naloxone prescription within the past 24 months, and OUD diagnosis on problems list. To further contrast BUP initiation between study phases while “detrending” the time effect, Poisson regression was utilized for the interrupted time series adjusted analyses. For this analysis, an offset was used for the monthly volume of patients presenting to the ED with OUD. Relative risk and 95% confidence interval (CI) are reported, with values above 1 corresponding to greater relative rates of ED BUP initiation. Physician X-waiver status was used as a covariate in this analysis. All analyses were

---

**Figure 2.** Screenshot of EMBED with care pathways welcome screen within EHR workflow. EHR = electronic health record; BUP = buprenorphine.
performed using SAS 9.4. Statistical significance was set as $p < 0.05$, two-sided.

RESULTS

Subject Characteristics

Of the 141,014 total ED visits during the two study phases used for analysis of main outcomes, 906 (574 preimplementation and 332 implementation) met inclusion criteria for analysis. Of these 906 OUD phenotype–positive visits, 98 patients in total (11%) had more than one ED visit, including a maximum of four visits (of which four patients had). Across these two phases, patients had a mean age of 39.9 years, 31.2% were female, 71.7% were white, and 73.4% had Medicaid insurance (Table 1).

Main Results

Primary Outcome. The rate of BUP initiation (i.e., BUP administered in the ED and/or prescribed on discharge) was 3.5% (20/574) in the preimplementation phase and 6.6% (22/332) in the implementation phase ($p = 0.03$; Table 2, Figure 1). Compared to the beginning of the implementation period, the rate of BUP initiation was higher after just-in-time training was offered as an additional component of the intervention (7.9% vs. 4.9%, $p = 0.28$). After adjusting for age, race, sex, number of waivered physicians, naloxone prescription within the past 24 months, and OUD diagnosis on problems list, the odds of ED-initiated BUP was 1.83 in the implementation phase compared to preimplementation (95% CI = 1.03 to 3.25). BUP initiation relative risk adjusting for the same covariates as well as the time trend with Poisson regression were 2.73 (95% CI = 0.62 to 11.99, $p = 0.18$) for implementation versus preimplementation phase. Of note, the significant difference between BUP initiation rates in the implementation and preimplementation phase persisted with adjustment for time trend, age, race, sex, naloxone

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Preimplementation ($n = 574$)</th>
<th>Implementation ($n = 332$)</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (year)</td>
<td>40.2 (±12.6)</td>
<td>39.4 (±12.1)</td>
<td>0.34</td>
</tr>
<tr>
<td>Sex, $n$ (% female)</td>
<td>172 (30.0)</td>
<td>111 (33.4)</td>
<td>0.28</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black or African American</td>
<td>90 (15.7)</td>
<td>43 (13.0)</td>
<td></td>
</tr>
<tr>
<td>White or Caucasian</td>
<td>425 (74.0)</td>
<td>225 (67.8)</td>
<td></td>
</tr>
<tr>
<td>Asian, American Indian, or Alaska</td>
<td>2 (0.3)</td>
<td>6 (1.8)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>57 (9.9)</td>
<td>58 (17.5)</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td>0.32</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>91 (15.9)</td>
<td>65 (19.6)</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>481 (83.8)</td>
<td>264 (79.5)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>2 (0.4)</td>
<td>3 (0.9)</td>
<td></td>
</tr>
<tr>
<td>Insurance information</td>
<td></td>
<td></td>
<td>0.56</td>
</tr>
<tr>
<td>BCBS or commercial</td>
<td>32 (5.6)</td>
<td>22 (6.6)</td>
<td></td>
</tr>
<tr>
<td>Managed care</td>
<td>35 (6.1)</td>
<td>15 (4.5)</td>
<td></td>
</tr>
<tr>
<td>Medicaid</td>
<td>426 (74.2)</td>
<td>239 (72.0)</td>
<td></td>
</tr>
<tr>
<td>Medicare</td>
<td>55 (9.6)</td>
<td>35 (10.5)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>26 (4.5)</td>
<td>21 (6.3)</td>
<td></td>
</tr>
<tr>
<td>Phenotype</td>
<td></td>
<td></td>
<td>0.57</td>
</tr>
<tr>
<td>Algorithm 1</td>
<td>368 (64.1)</td>
<td>219 (66.0)</td>
<td></td>
</tr>
<tr>
<td>Algorithm 2</td>
<td>206 (35.9)</td>
<td>113 (34.0)</td>
<td></td>
</tr>
<tr>
<td>Naloxone prescribed during encounter as inpatient medication</td>
<td>24 (4.2)</td>
<td>17 (5.1)</td>
<td>0.51</td>
</tr>
<tr>
<td>Prescribed naloxone within past 24 months</td>
<td>30 (5.2)</td>
<td>27 (8.1)</td>
<td>0.08</td>
</tr>
<tr>
<td>OUD diagnosis on problems list at time of encounter</td>
<td>105 (18.3)</td>
<td>83 (25.0)</td>
<td>0.02</td>
</tr>
<tr>
<td>Urine drug screen</td>
<td>88 (15.3)</td>
<td>56 (16.9)</td>
<td>0.54</td>
</tr>
<tr>
<td>Positive for opioids</td>
<td>52 (9.1)</td>
<td>26 (46.4)</td>
<td>0.14</td>
</tr>
<tr>
<td>Positive for oxycodone</td>
<td>9 (10.2)</td>
<td>7 (12.5)</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Data are reported as mean (±SD) or $n$ (%).
BCBS = Blue Cross Blue Shield; OUD = opioid use disorder.
prescription within the past 24 months, and OUD diagnosis on problems list and was only lost after adjusting for X-waiver status over time.

**Secondary Outcomes.** More subjects received a prescription for naloxone at discharge from the ED in the implementation period (6.5% vs. 11.5%; p < 0.01, Table 2). The rate of referral for ongoing MOUD treatment was 16.9% preimplementation and 18.1% in the implementation phase (p = 0.65).

The number of unique attendings who were present in both study phases (inclusive of all physician participants, not just faculty on the roster used to determine X-waiver status) who initiated BUP did not change significantly from 14/58 (24.1%) in the preimplementation phase to 16/58 (27.6%) in the implementation phase (p = 0.65). After just-in-time training was added for the second half of the implementation phase, the number of unique attendings who initiated BUP increased from 7/53 (13.0%) to 13/57 (22.8%, p = 0.10). The addition of just-in-time training was also associated with an increase in the proportion of unique attendings who launched the intervention in the implementation period (7/53, 13.0% vs. 15/57, 25.9%; p = 0.07).

Evaluation of the intervention’s implementation using the RE-AIM framework (Table 3) shows that it reached 44% of the target population (unique ED attending physicians who launched the intervention at least once), resulting in 32.3% of attendings adopting the practice of ED initiation of BUP. The rate of BUP initiation in the ED after training ended in the maintenance phase was 5.5% and associated with a decline in BUP initiation rates over time as displayed in Figure 1 (adjusted time trend of BUP initiation relative risk of 1.19 (95% CI = 0.03 to 56.3; p = 0.93)) for implementation versus maintenance phase. Of those interviewed, 23 responded to the SUS and NPS questionnaire items, the mean SUS score of 82.0 (95% CI = 76.7 to 87.2) is considered an acceptable score, and the NPS of +61 is a score consistent with more respondents indicating they were promoters than detractors of the intervention.

**Qualitative Interview Feedback.** As shown in Data Supplement S1, Appendix S2, categorization of qualitative feedback according to common themes revealed that the CDS decision support tools could be improved by tracking and displaying calculated scores

---

**Table 2**

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Preimplementation n = 574</th>
<th>Implementation n = 332</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary outcome</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BUP administered in ED or prescribed on discharge</td>
<td>20 (3.5)</td>
<td>22 (6.6)</td>
<td>0.03</td>
</tr>
<tr>
<td>Secondary outcomes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prescription for naloxone at ED discharge</td>
<td>37 (6.5)</td>
<td>38 (11.5)</td>
<td>0.009</td>
</tr>
<tr>
<td>Receipt of discharge instruction on opioid use, overdose education, naloxone education, and BUP education</td>
<td>218 (38.0)</td>
<td>115 (34.6)</td>
<td>0.32</td>
</tr>
<tr>
<td>Referral for ongoing MOUD</td>
<td>97 (16.9)</td>
<td>60 (18.1)</td>
<td>0.65</td>
</tr>
<tr>
<td>Number of unique attendings present in both phases who initiated BUP</td>
<td>14/58 (24.1)</td>
<td>16/58 (27.6)</td>
<td>0.65*</td>
</tr>
<tr>
<td>Rate of physician intervention launched per 100 phenotype-positive patients</td>
<td>7.3 (4.8–9.8)</td>
<td>—</td>
<td></td>
</tr>
</tbody>
</table>

Data are reported as n (%) or mean (95% CI). BUP = buprenorphine; MOUD = medication for opioid use disorder; OUD = opioid use disorder.

*McNemar’s test.

**Table 3**

<table>
<thead>
<tr>
<th>Type of Outcome</th>
<th>Specific Outcome</th>
<th>Value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reach</td>
<td>Unique attendings who launched the intervention at least once in the implementation phase.</td>
<td>19/68 (27.9%)</td>
</tr>
<tr>
<td>Effectiveness</td>
<td>Rate of BUP initiation in the ED (Implementation: Mar–Aug 2019).</td>
<td>22/332 (6.6%)</td>
</tr>
<tr>
<td>Adoption</td>
<td>Unique attendings who initiated BUP in the implementation phase.</td>
<td>16/68 (32.3%)</td>
</tr>
<tr>
<td>Implementation</td>
<td>Phenotype-positive patients for whom the intervention was launched in the implementation phase.</td>
<td>28/332 (8.4%)</td>
</tr>
<tr>
<td>Maintenance</td>
<td>Rate of BUP initiation in the ED after training ended (Sep–Dec 2019).</td>
<td>11/208 (5.5%)</td>
</tr>
</tbody>
</table>
(DSM-5, COWS, etc.), on the main screen for easy reference and automatically highlighting the best care pathway based on these scores. Additional suggestions included making the decision support tools more visible and clarifying that they are optional and not required to launch a care pathway. Regarding EHR workflow, a common area of feedback was to increase clarity of what happens when a care pathway is launched. Other interview feedback focused on the need to decrease confusion of the referral process, particularly details of the referral timeline, coordination with external providers, required next steps, and how to explain this process to patients. Additional miscellaneous suggestions include clarification of which features are available to providers without an X-waiver, increasing awareness of availability, continuation of one-on-one training to promote use, and addition of a feature alerting providers to possible OUD patients likely to benefit from BUP.

**DISCUSSION**

In this interrupted time series evaluating the preliminary efficacy of the EMBED intervention at a single site, implementation of a user-centered CDS with a brief, just-in-time training was associated with close to a doubling in the BUP initiation rate in the ED for patients with OUD and receiving a prescription for naloxone at discharge. After adjusting for the temporal trend of physician waiver training, the increased rate of BUP initiation was no longer statistically significant. The primary outcome was lower in the maintenance phase compared to the implementation phase (Figure 1). Given the ED’s significant role in caring for patients affected by the opioid epidemic, these results suggest that a user-centered, well-integrated CDS like EMBED is an efficacious approach to increase adoption of an effective treatment for OUD.

Despite the ED’s potential to initiate treatment for a large number of patients with OUD, a cohort study conducted in Massachusetts indicates that this potential has yet to be realized. Among the 17,000 patients who had an ED visit for nonfatal opioid overdose between 2012 and 2014, only one in three received a MOUD in the 12-month period following their overdose. Compared to no MOUD treatment, both methadone and BUP were associated with decreased all-cause mortality and opioid-related mortality. In light of these findings, successful initiation of BUP in the ED for the subjects in this study could have had a significant mortality benefit for these victims of nonfatal opioid overdose. Although the need for increased MOUD utilization in the ED is clear, the path to meeting that need is unfortunately not so simple.

A number of barriers currently limit the rate of BUP initiation in the ED. One such barrier directly addressed by the EMBED intervention is how the majority of emergency physicians feel unprepared to provide OUD care; a 2019 survey of physicians from two urban, academic EDs found that fewer than half of respondents felt prepared in several components of OUD emergency care. Only 39% of physicians self-rated themselves as prepared to determine the level of care needed by an OUD patient, while 29% felt prepared to connect OUD patients with outpatient treatment. Of all surveyed components, emergency physicians felt least prepared to initiate BUP, with only 27% self-reporting themselves as prepared. Specific features of the EMBED intervention are available to assist physicians feeling unprepared in each of these components, transforming what could be a time-consuming and unfamiliar task into a simpler and more feasible one for all ED clinicians. The results of this interrupted time series study demonstrate the preliminary efficacy of this intervention on physician adoption rate of ED-initiated BUP. This effect size could indeed be much larger as we did not collect data on other reasons for not starting treatment (e.g., patient readiness) given the pragmatic nature of the intervention.

Although 93% of emergency medicine faculty had an X-waiver by the end of the preimplementation phase, the proportion of unique attendings who had adopted the practice of ED initiation of BUP was low (19.2%). In the implementation phase, the proportion of X-waivered physicians increased slightly to 97%. However, during the implementation phase, we observed a close to a doubling in the proportion of unique attendings who had adopted the practice of ED initiation of BUP (19.2% vs. 32.3%, p = 0.53). Although this increase was not statistically significant, it seems more clinically significant and suggestive that the barrier of X-waivered physicians actually adopting the practice of ED-initiated BUP may require a user-centered CDS with just-in-time training. Similarly, although not statistically significant, the proportion of unique attendings who initiated BUP after we began just-in-time training nearly doubled. Taking all of these findings together, we hypothesize that all three (X-waiver training, user-centered CDS, and just-in-time
training) may be necessary and complementary to increase adoption of ED-initiated BUP.

Creation of a CDS with these capabilities is the result of our choice to employ a user-centered design process which involved identifying users’ needs and incorporating their feedback in each phase of iterative prototype development.23 The value of user feedback and collaboration also motivated changes to the intervention. After it was discovered that intervention usage during the first half of the implementation phase was relatively low, one-on-one tutorials and qualitative data collection were added to the intervention to enhance its use and performance. More in-depth qualitative analyses of implementation barriers are under way in four additional urban, academic EDs to determine additional ways to promote adoption of ED initiation of BUP.41

The decision to incorporate training as part of the intervention turned out to be one of the study’s strengths. Although not statistically significant, increased rates of the primary outcome with one-on-one training component suggests that this feature of the intervention is necessary to increase user recognition of its presence and value. The training component also leverages the science of “diffusion of innovations,” as a way to promote intervention adoption via communication and sharing by local champions and among colleagues within a social system.42 To generate a tipping point for adoption through visibility and diffusion, this training feature will be encouraged as an intervention component in the upcoming trial to facilitate implementation at study sites. Despite the increase in BUP initiation associated with training, the sustainability of its effect is questionable. Data collected in the maintenance phase show a decrease in the rate of BUP initiation over the months following the conclusion of the training period. Examining the sustainability of such training as a long-term solution versus a transient benefit as well as its impact on intervention scalability could be an area of future work.

Planning for the upcoming trial is also supported by the collection of qualitative data that resulted in a richer data set and a better understanding of barriers to adoption and possible solutions. Some of these findings include the need to increase awareness of the intervention, for the CDS to display calculated scores on the home screen, and to increase clarity of the referral process.

LIMITATIONS

There are many barriers to adoption of ED initiation of BUP into routine emergency care. For example, the EMBED intervention does not address negative attitudes toward addiction, the inconvenience of obtaining an X-waiver to prescribe BUP, the limited number of providers with an X-waiver, and access and availability to MOUD in the community. This study was also conducted in a single ED with a high rate of X-waivered physicians compared to nonacademic or rural EDs. Because physicians without an X-waiver are limited in their ability to prescribe BUP for home induction, this could have had an impact on the primary outcome of rates of ED BUP initiation. Another weakness of our study’s design is that it was neither randomized nor controlled, so causality between the intervention and outcomes could not be established. Given the urgency of the opioid crisis, additional temporal trends could have occurred that were not adjusted and remain as confounders. The increase in X-waivered physicians, although identified and adjusted during analysis, is an example of one trend which alone could have impacted study results if overlooked. Finally, in an effort to avoid unintended consequences from a hard-stop alert that triggered the CDS, not all attending physicians utilized the intervention during the implementation phase.43 Although the phenotype identified potentially eligible patients, the less intrusive nature of the intervention also meant that we did not collect data on actual presence or absence of OUD, withdrawal severity, and readiness for treatment. Therefore, the full potential and effect of the intervention on various outcomes may be underestimated or inaccurate.

Future research could explore alternative approaches to triggering the intervention in a nonobstructive manner and whether lack of a hard-stop alert represents another barrier to physician adoption. Additional investigative efforts are necessary to explore the extent to which findings of the study are generalizable across different patient populations, health care systems, and EHR platforms. Similarly, the scalability of EMBED as a solution to increasing adoption of BUP initiation in the ED will require further investigation, particularly across different health care systems using different EHRs. The upcoming pragmatic trial to be launched in 20 EDs across five health care systems using different EHR vendors may provide answers to some of these remaining questions.
CONCLUSION

This interrupted time series study demonstrates that implementation of the EMBED intervention in a single ED is feasible, acceptable, and estimated to be efficacious at increasing rates of buprenorphine initiation for the treatment of opioid use disorder in the setting of a temporal trend of increased physician X-waiver training. The intervention was associated with a doubling of the buprenorphine initiation rate, emphasizing the importance of user-centered health IT to change practice around evidence-based medicine that may be slowly adopted. Our findings suggest that we should proceed with the larger effectiveness trial to explore whether these preliminary findings are significant and generalizable. Knowledge gained from the study will continue to inform the trial, particularly the finding of how one-on-one training is a necessary part of the intervention to increase and sustain adoption of clinical decision support.

References


Supporting Information

The following supporting information is available in the online version of this paper available at http://onlinelibrary.wiley.com/doi/10.1111/acem.14002/full

Data Supplement S1. Supplemental material.
Data Supplement S2. Supplemental material.
Diagnostic Performance of a Rapid Point-of-care Test for SARS-CoV-2 in an Urban Emergency Department Setting

Samuel McDonald, MD, MS¹, D. Mark Courtney, MD, MSc¹, Andrew E. Clark, PhD, D(ABMM)², Alagaraju Muthukumar, PhD, DABCC², Francesca Lee, MD², Jyoti Balani, MD², Lenin Mahimainathan, PhD, MB(ASCP)², Anjali Bararia, MSc, M(ASCP)², Dwight Oliver, MD², Ravi Sarode, MD², and Deborah Diercks, MD, MSc¹

The ability to rapidly and accurately identify a patient’s COVID-19 status has had significant impact on emergency departments (ED) and health systems globally. Since the identification of SARS-CoV-2 illness in the United States, there has been rapid development in patient testing capacity following initial challenges including sparse availability. This was made possible by increasing availability of diagnostic molecular tests in several formats, from laboratory-based traditional, RT-PCR methods to near-patient testing rapid point-of-care (POC) PCR tests. Recent reports have shown the occurrence of false negatives at a higher rate with some of these tests.¹ False-negative results can potentially result in the spread of disease in the community, hospital patients, and critical personnel. Conversely, a false-positive diagnosis can result in potential exposure of a COVID-negative patient while receiving care in a COVID ward and unnecessary use of personal protective equipment (PPE) that is currently in limited supply. Much of the attention by infectious disease services and hospital leadership has been on minimizing false-negative results; however, paramount to effective testing is the overall concept of accuracy, which minimizes both false-negative and false-positive results. Measurement of SARS-CoV-2 test accuracy is complicated by the lack of a consensus reference method (or criterion standard) to compare results from newer assays.

Initially, turnaround times for SARS-CoV-2 testing results took approximately 5 to 7 days.² This was because SARS-CoV-2 testing was first established in reference and academic clinical laboratories with capacity for high-complexity test development. As testing was brought in house following commercial reagent availability, batched results from high-throughput assays became available within 24 hours. While these strategies were a major step forward and still have utility in an outpatient setting, such a timeframe cannot support most ED decision making. There is a need for rapid POC molecular tests that can be readily and safely deployed in an ED setting that generate reliable results in < 2 hours. Such tests provide clinically actionable results in the ED setting, facilitating diagnosis and rapid decision making. The ID NOW COVID-19 assay performed on the Abbott instrument platform is one such rapid
diagnostic test, capable of delivering results in 5 to 13 minutes.\(^3\)

Because of the novelty of this virus as well as the recent introduction of these tests into the health care market, there are little comparison data in any emergency setting to measure performance. Given the criticality of maximizing accuracy and getting some quantification of a false-negative rate, the aim of this study was to evaluate the agreement of the diagnostic performance of the ID NOW COVID-19 with the Abbott m2000 real-time PCR, the institutional presumed reference standard, for SARS-CoV-2 in patients presenting to an urban ED. The ID NOW COVID-19 uses isothermal nucleic acid amplification technology for detection of SARS-CoV-2 on a POC platform.

This is a retrospective analysis of data for prospectively collected specimens from symptomatic ED patients for standard-of-care decision making and was approved by the institutional review board. The health system recommended testing all patients presenting with a COVID-19-like illness with the final decision at the clinician’s discretion. All subjects had dry nasal swab (NS) testing with the Abbott COVID-19 assay on the ID Now platform (ID NOW, Abbott Diagnostics) paired with nasopharyngeal swab collected in viral transport medium tested on m2000 instrument (m2000, Abbott Molecular). All dry NSs were tested and results were obtained within 1 hour of collection. Positive results from the ID NOW were accepted as a final result. This was supported by prior in-house testing, validation data, and recent literature.\(^3\) If ID NOW returned negative, the paired NP sample was then tested on the m2000 instrument for concordance. During the study period, results were used to determine patient disease status in standard care. Diagnostic performance is presented as positive agreement and negative predictive value with associated 95% CI.

During the evaluation period, April 28, 2020 to May 13, 2020, a total of 585 patients were tested having 597 samples collected and evaluated on both the ID NOW and the m2000 platforms. This represented 100% of all COVID tests in the ED during this time and 43% of all ED encounters. The cohort had a mean (±SD) age of 53 (±19) years with an admission rate of 62%. Of those who were admitted, 9% were admitted to the ICU. Only the first valid sample pairs per encounter were included in the analysis. Additionally, six observations were removed due to an invalid result on the ID NOW or no corroborating m2000 result leaving a total of 579 samples. Within this cohort, the prevalence of COVID-19 was 5.7% (95% CI = 4.0% to 7.9%). There were a total of seven false-negative tests (7/33) using the ID NOW with a positive agreement of 78.8% (95% CI = 61.0% to 91.0%). The negative predictive value was 98.7% (95% CI = 97.4% to 99.5%) (Table 1).

Our study described the diagnostic performance of a rapid molecular test that was introduced to improve the evaluation of patients with symptoms concerning for COVID. Findings suggested ID NOW has a positive agreement of 79% when comparing to RT-PCR testing with the m2000 instrument, resulting in a probability of false-negative testing in 1% to 2% of patients when disease prevalence is around 6%. Our results are similar to those reported in other studies comparing the ID NOW to the Abbott m2000.\(^3\) A subsequent smaller comparison study reported the sensitivity to be 87%.\(^4\) However, a preprint study recently reported a sensitivity of ID NOW to be 51.6% when using the Cepheid Xpert Xpress SARS-CoV-2 as the reference standard.\(^5\)

Statistically, no test will likely result in 100% accuracy in all settings and thus some level of discordance can be anticipated, particularly for a disease in which optimal clinical and diagnostic testing parameters have yet to be defined. In a higher prevalence of disease setting when 20% to 25% of tests are positive, this false-negative rate would have significant implications. However, in the context of a 6% pretest prevalence of disease, this likely 1% to 2% rate of false-negative results can be mitigated by thoughtful operational decisions.

While there is rising discussion with concern for false-negative testing, it remains to be shown whether this is a clinically significant failure to detect active disease or merely failure to detect low levels of viral RNA of uncertain clinical significance. Additional work could be performed to determine where patients are in the course of their illness trajectory and how a bedside clinician could utilize this knowledge to better direct testing modalities to mitigate misleading testing information.

There are several potential approaches to address the lower accuracy of ID NOW testing in the ED.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>2 × 2 Table of Diagnostic Performance of ID NOW SARS-CoV-2 Assay Compared to the m2000 Using Dry Nasal Swabs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Abbott ID NOW</td>
</tr>
<tr>
<td></td>
<td>Positive</td>
</tr>
<tr>
<td>Positive</td>
<td>26</td>
</tr>
<tr>
<td>Negative</td>
<td>7</td>
</tr>
</tbody>
</table>
setting. Requiring all negative test results to be verified on an auxiliary platform (i.e., m2000) is one strategy. Patients who are discharged can simply be discharged with strict self-isolation precautions until confirmatory results are available the following day utilizing standard follow-up mechanisms. Unfortunately for admitted patients, most hospitals want strict separation in wards of COVID-positive or COVID-negative patients. This would result in significant ED boarding while awaiting confirmatory testing. For effective implementation, hospitals need to have a data-driven risk stratification step after an initial negative result to determine specific hospital location or simply assign all patients to a COVID ward with prompt deescalation procedures with a concordant negative test performed by the reference standard. This strategy of reflexing all negative result samples to a traditional RT-PCR test would require retesting almost 94% to 95% of patient samples. At our current disease prevalence, this may be of questionable value.

Alternatively, one could also simply rely on the more rapid ID NOW and tolerate a 1% to 2% false-negative rate in populations with low disease prevalence as ours, with lesser impact for discharged patients who receive self-isolation instructions and could be verified with additional testing for inpatients who deteriorate or fail to improve. This requires maintaining awareness of the current region’s disease prevalence.

We choose to rely on the ID NOW with the caveat that an ED physician can confirm result using Abbott m2000 if there is suspicion that ID NOW test is a false negative. This results in repeat testing in only those patients where there remains a high posttest probability and relies on physician gestalt estimation or a priori knowledge of disease probability. This raises the potentially important role that formal or scientifically validated pretest probability tools are developed appropriate for the acute care setting.

In our ED cohort of patients, we found 1% to 2% will have a false-negative result when using ID NOW compared to the Abbott m2000. Developing a process for retesting those patients in whom the physician identifies as having a high pretest probability of disease may address the concerns around false-negative results.

REFERENCES

Quality clinical research remains a high priority during the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) coronavirus disease (COVID-19) pandemic response; however, the logistic barriers of conducting research across institutions with differing departmental policy responses and varying resources is a documented problem. The COVID-19 pandemic has resulted in a surge of more than 1,000 newly formed clinical trials registered on clinicaltrials.gov with “COVID-19” in their study title as of May 8, 2020. While the need for new and innovative research remains high, traditional operations to support the execution of these studies must rapidly evolve to overcome the newly formed barriers of clinical research.

Most notably, reserves of personal protective equipment (PPE) dwindled and have forced hospitals to reevaluate the distribution of these supplies in an effort to conserve resources and minimize waste. Due to the lack of PPE, the logistics of screening and consenting patients in the emergency department (ED) must be revised while continuing to make safety of colleagues, patients, and the general public a main priority. Best practices need to be established to adapt clinical research into the rapidly changing environment of the COVID-19 response.

As the pandemic continues, new recommendations for conducting research have been made by the National Institutes of Health (NIH) and the U.S. Food and Drug Administration (FDA) for the COVID-19 public health emergency. These recommendations suggest that all ongoing studies during the COVID-19 pandemic should incorporate procedures that are compliant with regional management policies for controlling the spread of the disease; however, the logistics of executing that research is left to individual institutions. We are hesitant to provide specific examples of studies relevant to emergency medicine that should be conducted during the COVID-19 pandemic due to the context-dependent issues across institutions and for fear that while an example might be suitable in one setting, it may be deemed inappropriate in another. Some variables such as the seriousness of the disease, investigational product supply, and whether there are reasonable alternative treatments will be relatively consistent across institutions, but other variables such as viability, available resources, and the ability to safely administer an interventional product might vary widely across sites. Instead, we want to detail who should be judging the criteria to continue research so that individual institutions can make an informed decision that best suits their specific scenario. The FDA has recommended that sponsors, in consultation with clinical investigators and institutional review boards/independent ethics committees, judge whether a study should be open to enrollment during the pandemic.

Ongoing clinical trials raise an additional area of concern during pandemic response. Although there is
some guidance from local and national organizations, the assessment of an appropriate response weighing ethical principles, staffing constraints, risk to participants, and ultimately the course of action to take resides on the principal investigator and the institutional review board.\textsuperscript{4,5} It is widely believed that critical clinical care research during a pandemic must have a different approach than during nonemergent circumstances.\textsuperscript{6} Logistic adaptations to mitigate risk for patients, health care workers, and the general public are paramount for fostering an environment for clinical research to continue during a pandemic. A review of the literature through PubMed using the search terms “COVID-19 research guidelines” and “COVID-19 research best practices” did not produce any guidelines or best practices pertaining to navigating the logistic barriers that emergency medicine (EM) research faces during the COVID-19 response.

Reevaluating how research is conducted in the ED and redefining clinical research operations has been necessary. Our objective was to evaluate the essential elements needed to keep clinical research operations open and active to enrollment and then systematically determine how research teams could immediately adapt these elements to the current pandemic environment while still maintaining safety measures for our workers, patients, colleagues, and families. The term “essential elements” was defined by our working group as the necessary measures that needed to be taken to conduct vital research operations. While financial support is a critical and a core component of research logistics, it was decided that funding was a byproduct of research enrollment and that defining safety measures for staff, colleagues, and patients should take priority; however, we acknowledge that in the long-term funding issues must be addressed to maintain research viability.

The Clinical Researchers’ United Exchange (CRUX) is an innovative national interest group affiliated with the Society for Academic Emergency Medicine (SAEM) established in 2018. Participation is open to all SAEM members, but was designed specifically to engage nonclinical research staff in the execution of research within EM. On April 22, 2020, the group convened an emergency meeting to discuss operational logistics and share best practices related to staffing, screening, enrolling, and processing of biospecimens during the COVID-19 pandemic. A CRUX chair moderated the session in which each topic was addressed systematically, while providing members a platform to share their institutions’ newly adapted procedures and concerns. Following each section, members were able to discuss best practices for each topic. Comments were recorded and subsequently circulated among CRUX members to confirm consensus. Responses were reviewed and collated into recommendations that are summarized by Figure 1.

Staffing for clinical research projects during the pandemic was one of the largest barriers to research we identified. Variables such as personal health and living situations came to the forefront of staffing discussions as comfort levels and continuation of job duties for nonclinical research staff was evaluated. The group identified that these types of difficult conversations were best fostered by creating the environment for managing up. This was accomplished by informing teams of what tasks need to be done and allowing them to volunteer for roles that they feel comfortable executing. It was suggested that those who do not fit into the current needs of the department be transferred into institutional labor pools to fill new or vacant roles. Members on the call described filling emerging roles in newly adapted research studies with furloughed non-EM health care providers. Those who were interested received research training to support ongoing clinical studies with their clinical skill sets. The CRUX discussion group identified that even when staffing needs are met, a hand-off strategy should be developed with standard operating procedures to ensure that the experiences of newly adapted roles are passed on through the anticipated summer turnover of staff leaving the department. Additionally, many institutions have undergone hiring freezes, and once the immediate needs of critical research studies have been met, efforts should be shifted toward reestablishing paused studies in a timely manner.

Screening procedures of patients who met criteria for ongoing clinical studies were impacted by institutional restrictions. At some hospitals, nonclinical research staff were not allowed in the ED and limited to electronic medical record review studies. For institutions that did allow research staff into the ED, virtual study screening through electronic medical chart review or analyzing electronic screening surveys via tablets in patient rooms were recommended as remote alternatives to in-person discussions for determining study eligibility. Action at a distance may continue to be the new normal for clinical researchers in the time of COVID-19.\textsuperscript{7}

Study enrollment procedures required the greatest number of modifications during the pandemic
response. Institutions with many COVID-19 research studies should incorporate a COVID trials’ triage system, which could be as simple as a phone number to call at time of screening, staffed by a clinical research professional versed in all the ongoing research studies in the ED. The triage system can act as a central processing unit for identifying which research study would be most beneficial for the patient, the institution and the pursuit of knowledge. To save PPE and reduce the number of people exposed to COVID-19, all institutions should consider implementing telemedicine devices, whether it be a tablet, smartphone, or a landline telephone to gauge a patient’s interest in participating in a research study without making physical entrance into the room. Some institutions reported utilizing research physicians who are working clinically to conduct discussions of consent; while this is convenient and can minimize excess contact with a patient, it is not sustainable as research efforts grow. Teleconsent was found to be most well received at hospitals that already have telemed services integrated into their current standard of care. \(^5\) Paperless enrollment is recommended, but when not possible, the recommendation to use disposable pens, sterilize clipboards, and minimize entrances into the patient room were cited as strategies that could aid in infection control. Where appropriate, instead of leaving to make a photo copy of a signed consent form, consider holding the document up to the window and having a colleague taking a photo using a HIPAA-compliant device to document consent. If the paper consent form must be removed from a patient’s room, do so in a sealed bag that could be decontaminated upon exit.

Procedures for processing biospecimen samples and study data needed to adapt to newfound recommendations of social distancing for COVID-19. At the consequence of cross-functional training, larger institutions with multiple people working per shift should look to have individual task assignments with the same employees working together during shifts (i.e., team A or B). Staff working in the ED should have minimal movement throughout the hospital and should be limited to duties within the ED. Administrative tasks or laboratory processing should ideally be done by another team member. Coordination of necessary work among teams should be done to minimize entry or egress to parts of the hospital as well as the individual patient rooms. Trip consolidation should be applied where enrollment windows permit. Contact with one another when not necessary (i.e., biospecimen sample handoff) should be replaced with a designated, safe location drop-off. Communication of staff location should be maintained throughout the workday to minimize staff cohorting in laboratory or office settings.

With the outbreak of COVID-19, we have seen unprecedented changes, innovation, and reevaluation of our health care system. During these tumultuous times, one thing that remains steadfast is that health care workers in the ED continue to do what they do best—adapting and overcoming any situation that presents itself. Some things will be forever changed by this pandemic, but the innovations and reevaluations of workplace efficiencies and operational logistics can propel us to the next level of innovative care for the future. We advocate that every institution have an open discussion of their essential research operations and use some of the recommendations presented here to find a revised, innovative solution to the challenges presented by COVID-19 that best suit their research endeavors.
References


The first cases of Coronavirus disease 2019 (COVID-19) were reported in Wuhan, China, in December 2019.\(^1\) The literature demonstrates geographic variation with regard to estimates of infection incidence, suggesting that COVID-19 has been under-diagnosed in certain regions.\(^2,3\) The rate of asymptomatic infection has been estimated to be as high as 30.8%, which may help explain variation in incidence, particularly in regions with differing screening practices.\(^3\) Transmission of COVID-19 by asymptomatic carriers has been reported in multiple family units, indicating that this mode of infection is important in understanding disease epidemiology and population risk.\(^4,5\) In one study, in individuals who were asymptomatic at the time of confirmed COVID-19 infection, the median communicable period (defined as time from positive test to negative test) was 9.5 days (range = 1–21 days), and approximately 21% of these patients went on to develop symptoms, suggesting that individuals may be infectious prior to the development of symptoms.\(^5\) As such, identifying asymptomatic individuals is likely important in containing the spread of the virus. With the temporary closure of many ambulatory care settings, the emergency department (ED) has become a focal point in COVID-19–related healthcare delivery in many communities. However, due to various factors including limited testing capacity, data regarding the yield of routine screening of asymptomatic individuals are limited. Furthermore, while the utility of screening asymptomatic individuals has been explored in high-prevalence regions, studies have not been published for low-prevalence regions.\(^6,7\) In this study, we aimed to describe screening for COVID-19 in asymptomatic ED patients in a low-prevalence region.

This was a retrospective cohort study of data from the study site’s electronic health record (EHR). The study was approved by the study site’s institutional review board. The study site is an urban, Level I trauma center and tertiary referral center with an annual ED volume of 82,000 patients. It resides in a county with a population of 1,527,718 people. We included all ED patients who had a COVID-19 test ordered under the site’s asymptomatic screening protocol, either in the ED or within 24 hours of admission from the ED. The study site was also performing asymptomatic COVID-19 testing at one of its clinics, for patients scheduled for surgical procedures. Patients were excluded if they incorrectly presented to the ED for this asymptomatic screening. On April 4, 2020,
the study site began testing for COVID-19 in asymptomatic ED patients who were placed on involuntary psychiatric holds. This testing process began due to concerns that COVID-19 infections would be difficult to control in the close living environments of psychiatric hospitals, and the study site wanted to ensure that patients did not have COVID-19 prior to transfer to these facilities. On April 14, 2020, the asymptomatic screening protocol expanded to include testing for all patients admitted to the hospital from the ED and all patients awaiting placement in other close living environments (i.e., skilled nursing facility, jail, or other congregate living facility). When ordering asymptomatic COVID-19 testing, the provider was required to enter the indication for testing, including “other” if the reason for testing was not listed. All patients undergoing asymptomatic testing were required to wear surgical masks, and all ED providers were also required to wear surgical masks. No further isolation (contact or airborne) was implemented as part of the asymptomatic screening protocol.

Detection of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was performed by reverse transcriptase polymerase chain reaction, using samples obtained from nasopharyngeal swabs. Two assays were used during the study period, both of which were sanctioned by the Food and Drug Administration’s Emergency Use Authorization and validated for use with nasopharyngeal swab samples. One assay was run on a high-throughput platform (Cobas 6800, Roche Diagnostics) and one assay was run on a medium-throughput platform (BDMax, Becton Dickinson). Data were collected from April 4, 2020, to May 13, 2020 (40 days). The following data were explicitly defined and collected from the EHR for each patient meeting inclusion criteria: age, sex, Emergency Severity Index (ESI) triage criteria, homeless status, indication for asymptomatic COVID-19 testing, COVID-19 test order, COVID-19 test result, and admission team. ESI is a triage scheme that stratifies patients into five categories from one (most severe) to five (least severe).

Homeless status was categorized as “yes” or “no,” based on patient self-report. Indications for asymptomatic COVID-19 testing included the following: 1) admission to the hospital, 2) involuntary psychiatric hold, 3) placement in a close living environment, and 4) “other.” When the indication of “other” was selected, providers were required to insert the reason as free text. In the case of canceled COVID-19 tests, data regarding the reason for cancelation were not captured for individual patients. However, tests were usually canceled because of an inadequate sample or because the patient refused testing. The primary outcome measure was the result of COVID-19 testing. The infectious disease service was consulted to assist with management of patients with positive test results. Data are described with simple descriptive statistics. Continuous data are presented as mean ± one standard deviation (SD). Ninety-five percent confidence intervals (95% CIs) are presented, where appropriate. Data analysis was conducted using Stata 15 (StataCorp, 2017).

A total of 1,342 ED patients had an order placed for COVID-19 testing using the asymptomatic testing protocol. The mean ± SD age was 47.4 ± 22.9 years and 823 (61%) were male. Of these, 96 (7.1%, 95% CI [5.8%, 8.7%]) patients had their test canceled (either specimen was inadequate or patient refused testing). Patient characteristics, stratified by testing status (completed vs. canceled), are presented in Table 1. Thirty-six patients had a testing indication of other (clinic request = 3, homeless = 3, predialysis = 2, high-risk = 2, clearance for return to work = 2, possible admission = 1, resolved symptoms of COVID-19 = 1, chills/sick contact = 1). In the 1,246 patients who completed asymptomatic testing for COVID-19, two (0.2%, 95% CI [0.02%, 0.6%]) tested positive. One of these cases underwent repeat testing the next day (while admitted) and tested negative and remained asymptomatic during hospitalization. This case was considered by the infectious diseases consult team to be either a false-positive test or an asymptomatic infection. The other positive patient, when evaluated by the infectious diseases consult team, was determined to have both a cough and close contact with COVID-19–infected individuals and thus was deemed a symptomatic infection. During the 40-day study period, the study site tested 886 symptomatic patients for COVID-19 and 37 (4.2%, 95% CI [3.0%, 5.7%]) tested positive.

As COVID-19 testing capacity increases with time, more institutions will likely seek to adopt an asymptomatic screening protocol for admitted patients or for patients who will be transferred to other inpatient settings, to control nosocomial spread of the virus. In this study, we describe the testing of asymptomatic ED patients for COVID-19 in a low-prevalence region of the United States. Prior studies on asymptomatic screening are from high-prevalence regions. During the study period, 526 cases of COVID-19 were
diagnosed in the study site’s county (40-day incidence = 34 cases per 100,000 population).9 The Centers for Disease Control and Prevention reported a national incidence that varied significantly by region, ranging from 20.6 to 915.3 cases per 100,000 population, during a time period just prior to this study (February 12 to April 7, 2020).10 Thus, the study site represents a low-prevalence region. In our study population, we found a very low true-positive rate among asymptomatic patients undergoing testing. While the rate of asymptomatic infection was low, testing led to the diagnosis of COVID-19 in one inpatient, who then was placed in appropriate isolation. Had this patient not been tested via the asymptomatic screening protocol, this COVID-19 infection would not have been identified, potentially exposing health care workers and other patients. When performing disease surveillance in low-prevalence regions, the presence of false-positive test results is inevitable. In our study, one of the two patients with positive test results was likely a false-positive. This asymptomatic individual was retested the next day and found to be negative.

This study must be interpreted in the setting of its limitations. This was a single-center study, so our results may not be generalizable to other settings. As the prevalence and impact of COVID-19 is very regional, local COVID-19 infection trends must be considered. This study was retrospective, and thus it is limited by the data available in the EHR and subject to the limitations of a retrospective study. As the data were downloaded from the EHR, errors in manual data abstraction were minimized. Finally, a small percentage of patients (7%) had their test canceled. It is possible that some of these patients may have tested positive.

Screening asymptomatic ED patients for COVID-19 in a low-prevalence region is of very low yield. The decision to test asymptomatic ED patients should be made on an institution-by-institution basis, dependent on regional prevalence and test availability.

### References


### Table 1

Characteristics of Asymptomatic Patients Tested for COVID-19

<table>
<thead>
<tr>
<th></th>
<th>Test Completed (n = 1,246)</th>
<th>Test Canceled (n = 96)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td>47.9 ± 22.9</td>
<td>40.3 ± 21.9</td>
</tr>
<tr>
<td><strong>Sex (male)</strong></td>
<td>758, 61% (58%–64%)</td>
<td>65, 68% (57%–77%)</td>
</tr>
<tr>
<td><strong>Homeless</strong></td>
<td>189/1035, 16% (14%–19%)</td>
<td>19/82, 23% (15%–34%)</td>
</tr>
<tr>
<td><strong>Indication for test</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Admission</td>
<td>956, 77% (74%–79%)</td>
<td>65, 68% (57%–77%)</td>
</tr>
<tr>
<td>Psychiatric hold</td>
<td>229, 20% (18%–23%)</td>
<td>21, 22% (14%–31%)</td>
</tr>
<tr>
<td>Placement</td>
<td>25, 2% (1%–3%)</td>
<td>2, 2% (0%–7%)</td>
</tr>
<tr>
<td>Other</td>
<td>29, 2% (2%–3%)</td>
<td>7, 7% (3%–14%)</td>
</tr>
<tr>
<td>Unknown (missing)</td>
<td>10, 1% (0%–1%)</td>
<td>1, 1% (0%–6%)</td>
</tr>
<tr>
<td><strong>ESI triage category</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>36, 3% (2%–4%)</td>
<td>3, 3% (0.7%–9%)</td>
</tr>
<tr>
<td>2</td>
<td>71, 6% (5%–7%)</td>
<td>5, 5% (2%–12%)</td>
</tr>
<tr>
<td>3</td>
<td>626, 51% (48%–54%)</td>
<td>43, 45% (35%–56%)</td>
</tr>
<tr>
<td>4</td>
<td>337, 27% (25%–30%)</td>
<td>40, 42% (32%–53%)</td>
</tr>
<tr>
<td>5</td>
<td>32, 3% (2%–4%)</td>
<td>4, 4% (1%–10%)</td>
</tr>
</tbody>
</table>

Continuous variables described as mean ± SD. Categorical variables described as number, percent (95% CI).

ESI = Emergency Severity Index.

*Data are self-reported. Data were missing for 225 patients.
†Data were missing for 12 patients.


BACKGROUND

Acute otitis media (AOM) is the second most commonly diagnosed illness in children and the most common indication for antibiotic prescription. More than one-third of children experience pain, fever, or both 3 to 7 days following treatment, and nearly 75% of parents identify pain and disturbed sleep as the most important sources of AOM-related burden.

There is significant parental uncertainty regarding treatment of AOM and less than 30% of U.S. parents receive instructions on appropriate analgesia for their children. Discharge instruction complexity and inadequate comprehension is associated with medication errors, suboptimal postdischarge care, and unnecessary recidivism. Medication errors can be reduced using standardized discharge instructions, and parents prefer these to verbal summaries. Video discharge instructions have been shown to be preferred over paper instructions in many pediatric presentations; however, no study has explored the effectiveness of video instructions for AOM.

ARTICLE SUMMARY

This study evaluated parents of children aged 6 months to 17 years with a chief complaint of otalgia in the setting of upper respiratory tract infection and where the treating physician ultimately made a diagnosis of AOM. In the intervention group, parents were shown a 5-minute video detailing discharge instructions, while parents in the control group received the same discharge instructions via paper handout. The primary outcome was the AOM Severity of Symptom (AOM-SOS) score on day 3 postdischarge from the ED.

QUALITY ASSESSMENT

Overall, this was a well-performed randomized study performed in a single Canadian tertiary care pediatric emergency department. Allocation and the randomization process was concealed, although it was unblinded given that one group watched a video and the other received a paper handout.

There are several limitations to this study. They used a convenience sample of patients, recruiting 7 days a week, but only between 10 AM and 10 PM. The authors acknowledge that it is possible that children presenting in the middle of the night are experiencing more pain than those that present during daytime or evening hours, or their symptoms may just be more disruptive to patient and caregiver sleep. Furthermore, parental understanding of discharge instructions may differ significantly when those instructions are given in the middle of the night.

Another limitation is external validity, as the study was performed in a single tertiary care pediatric ED.
with more resources than rural and nonpediatric EDs. Furthermore, they excluded non–English-speaking patients from the study for feasibility purposes and thus application is currently limited in parents without English fluency. However, the video discharge instructions are freely available online and could be shown to patients in any center with internet access.

Loss to follow-up was also a concern in this study as only 68% of patients in the trial completed the primary outcome. The authors did anticipate a high loss to follow-up in their power calculation, but the loss of patients could skew the results through attrition bias.

KEY RESULTS

219 parents were randomized and 149 completed the primary outcome (77 video; 72 paper instructions). The children had a mean age of 2.9 years, 49% were female, and 18.7% were not offered analgesia prior to arrival. There were no crossovers in the trial.

The primary outcome was the AOM-SOS score on day three. The AOM-SOS score is a 7-question survey that assesses the child’s symptoms over the last 24 hours as reported by the caregiver. The questions enquire about things such as crying, ability to sleep, appetite, and activity level. The score ranges from 0 to 14 with higher scores indicative of greater symptoms severity.

The median score at day 3 was 1 in the video group as compared to 3 in the paper group \((p = 0.004)\), and the difference was maintained after adjusting for pre-intervention AOM-SOS and medication use.

There were no significant differences in secondary outcomes, which included parental knowledge gain, functional outcomes, and the number of children receiving antibiotics or analgesics following discharge.

AUTHORS’ COMMENTS

Discharge instructions are an important part of any ED visit, particularly for parents of children who cannot speak or understand instructions. Video discharge instructions have been used and found to be effective for other pediatric illnesses, and this study highlights their potential benefit for AOM. However, video instructions had no effect on parental knowledge, so it remains unclear how symptom scores were lowered through use of video.

TOP SOCIAL MEDIA COMMENTARY

On Twitter we (@socmobem) asked:

What do you think needs to be done to help practitioners transition to using a video?

Shawn Dowling (@Shawnkdowling) responded:

Novel and consistent ways to disseminate these! Playing on hospital TV’s, having these materials linked to patient facing websites, pt education app, etc! Telling families to "google-patient education video" doesn’t instill a lot of confidence that they will get to the right resource.

Teresa Chan (@TChanMD) responded:

I just ask to borrow [patients] phones and help them load it. I probably could make a QR code card! Imagine a series of discharge instruction videos that we curate and make QR code-based cards for ppl to take home?

David Barbic (@DavidBarbic) responded:

Why reinvent the wheel? Go to http://bcemergncynetwork.ca for dozens of handouts and links to websites (many reviewed by patient partners), reviewed by EPs @BCEmergMedNtwrk

TWITTER POLL

The twitter poll results showed that no one is currently using video discharge instructions for AOM.
TAKE-TO-WORK POINTS

For English fluent parents of children with AOM, providing them with video discharge instructions can improve symptom severity at 72 hours after ED discharge.

References

Anticonvulsants for the Treatment of Low Back Pain and Lumbar Radicular Pain

Michael Gottlieb, MD\(^1\), Alex Koyfman, MD\(^2\), and Brit Long, MD\(^3\)

<table>
<thead>
<tr>
<th>NNT color recommendation</th>
<th>Red (harm &gt; benefit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary heading</td>
<td>When compared to placebo, anticonvulsants do not reduce pain or disability in low back pain but increase the risk of adverse events.</td>
</tr>
<tr>
<td>Benefits in NNT</td>
<td>No one was helped</td>
</tr>
<tr>
<td>Benefits in percentages</td>
<td>No one was helped</td>
</tr>
<tr>
<td>Harms in NNT (NNH)</td>
<td>1 in 6 were harmed (adverse event)</td>
</tr>
<tr>
<td>Harms in percentages</td>
<td>16% more were harmed (adverse events)</td>
</tr>
<tr>
<td>Efficacy endpoints</td>
<td>Reduction in pain or disability in the immediate, short, intermediate, and long term</td>
</tr>
<tr>
<td>Harm endpoints</td>
<td>Adverse events</td>
</tr>
<tr>
<td>Who was in the studies</td>
<td>9 trials comprising 859 total adults with low back pain with or without radiation to the leg, sciatica, or neurogenic claudication of any duration</td>
</tr>
</tbody>
</table>

NARRATIVE

Low back pain is a common cause of presentation to the emergency department, comprising over 4% of all presentations.\(^1\) Globally, low back pain was found to rank highest in disability and sixth overall in disease burden, with a prevalence of 9.4%.\(^2\) However, the majority of episodes do not require a surgical intervention and are treated conservatively with medical therapy.\(^3\) While clinical guidelines generally recommend nonpharmacologic interventions and nonopioid analgesics,\(^4\) one large study found that the prescription of anticonvulsant medications has nearly doubled from 2000 to 2010.\(^5\) However, it is important to determine whether these medications are safe and effective for use in this patient population.

The systematic review discussed here included nine randomized parallel and crossover controlled trials (859 subjects in aggregate) investigating the efficacy of anticonvulsants compared with placebo in adults with nonspecific low back pain with or without radiation to the leg, sciatica, or neurogenic claudication of any duration.\(^6\) The systematic review excluded studies investigating patients who were pregnant or postsurgical or who had mixed conditions (e.g., low back pain and neck pain). Outcomes included pain intensity, disability, and adverse events. Pain and disability were assessed at the following time points: immediate (≤2 weeks after randomization), short term (2 weeks to 3 months), intermediate term (3 to 12 months), and long term (≥12 months).

Seven trials evaluated a gabapentinoid (e.g., gabapentin, pregabalin), and two evaluated topiramate. The mean age of participants was 50.8 years in the treatment group and 51.5 years in the placebo group. Four trials recruited participants with chronic low back pain with or without radiating leg pain and five trials recruited participants with lumbar radicular pain. Only one trial included participants with acute symptoms.
There was no difference in pain scores in the short-term, intermediate-term, or long-term endpoints among the included studies (moderate- to high-quality evidence). Gabapentinoids also did not reduce the risk of disability (based on validated disability scales) in study participants (moderate-quality evidence). However, there was an increased rate of adverse events in the treatment group (53.0% vs. 36.7%; pooled relative risk = 1.4, 95% CI = 1.2 to 1.7, absolute risk difference = 16.3%; number needed to harm of 6; high-quality evidence). The most common adverse events reported in participants taking a gabapentinoid were drowsiness or somnolence, dizziness, and nausea. The most common adverse events reported in participants taking topiramate were paresthesia, drowsiness or somnolence, dizziness, and diarrhea.

CAVEATS

While this meta-analysis found that anticonvulsants did not improve pain or disability, there are several limitations that must be considered. First, while most studies included gabapentinoids, the dose and frequency varied significantly within the included trials. Additionally, there were limited data on topiramate, so it remains unclear whether this medication is effective in this population and further studies are needed. The populations also varied with regard to the underlying pathology and inclusion criteria. Of note, some studies required radiographic imaging to exclude underlying pathology, while others relied exclusively on clinical criteria. Moreover, the time periods varied with regard to the chronicity of the symptoms. This is important because chronic back pain is associated with worse outcomes than acute back pain. While most studies evaluated chronic back pain, Maher et al. included predominately acute (i.e., less than 3 months’ duration) back pain patients and also found no difference in pain or disability in their study. Finally, there were differences in the pain assessment endpoints in the study groups. As a result, many of the assessments included relatively smaller groups.

Based on the above data, this analysis suggests that gabapentinoids are associated with an increased risk of adverse events without an improvement in pain or disability. We have therefore assigned a color recommendation of red (harm > benefits) to this intervention.

References

Ketamine Versus Opioids for Acute Pain in the Emergency Department

Michael J. Duhaime, MD and Allan B. Wolfson, MD

<table>
<thead>
<tr>
<th>NNT color recommendation</th>
<th>Yellow (unclear if benefits)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary heading</td>
<td>Ketamine is not inferior to morphine for acute pain control in the ED</td>
</tr>
<tr>
<td>Benefits in NNT</td>
<td>Not applicable (ketamine was not inferior to morphine for acute pain control)</td>
</tr>
<tr>
<td>Benefits in percentages</td>
<td>Not applicable</td>
</tr>
<tr>
<td>Harms in NNT (NNH)</td>
<td>No one was harmed (no serious or life-threatening adverse events) 1 in 9 were harmed (experienced neuropsychological adverse events)</td>
</tr>
<tr>
<td>Harms in percentages</td>
<td>No one was harmed (no serious or life-threatening adverse events) 11% higher risk of neuropsychological adverse events compared to the opioid group</td>
</tr>
<tr>
<td>Efficacy endpoints</td>
<td>Change in Visual Analog Scale (VAS) for pain</td>
</tr>
<tr>
<td>Harm endpoints</td>
<td>Adverse events (hypoxia, neuropsychological effects such as agitation, hallucination, dysphoria, and confusion)</td>
</tr>
<tr>
<td>Who was in the studies</td>
<td>A total of 870 adult patients who presented to the ED with acute pain from traumatic and nontraumatic causes, in two meta-analyses comprising a total of 11 trials</td>
</tr>
</tbody>
</table>

**NARRATIVE**

Acute pain is one of the most common complaints in the emergency department (ED). With recent efforts to find effective nonopioid analgesics, ketamine has surfaced as a potential option for ED analgesia.\(^1^\)\(^3^\) While ketamine is typically used as a sedative agent, several studies have shown that when it is administered in subdissociative doses, both as a stand-alone agent and as an adjunct to opioids, it may also provide analgesia.\(^4^\) Two recently published meta-analyses compared ketamine with opioid analgesics in adult ED patients with acute pain.

The patient-level meta-analysis, by Karlow et al.,\(^5^\) included randomized controlled trials that directly compared a single bolus, slow push, or slow infusion of a subdissociative intravenous (IV) dose of ketamine with a single IV dose of opioid/opiate analgesia. This comprised three studies that included a total of 261 adult ED patients. The primary outcome studied was the change in patient-reported pain scores after administration of ketamine (dose range = 0.3 to 0.5 mg/kg IV) or morphine (dose = 0.1 mg/kg IV). The pooled estimate for the mean difference in reported pain score reduction between the ketamine and morphine groups was 0.42 (95% confidence interval [CI] = −0.70 to 1.54). Because of heterogeneity in the methods and timing of pain assessment and in event assessment, adverse events were reported as raw data. Ketamine was associated with a higher rate of adverse events than morphine in all of the individual studies. However, the only reported acute life-threatening adverse event was decreased oxygen saturation, which was reported in a single patient in the opioid trial arms and in none of the patients in the ketamine groups.

Ghate et al.\(^6^\) performed a systematic review and meta-analysis comparing low-dose ketamine with opioids in adults with acute pain in the ED. The authors included eight studies (six RCTs and two...
observational studies), with a total of 609 ED patients. The major outcome studied was change in patient-reported pain scores 30 minutes after treatment. Both low-dose ketamine (dose range = 0.1 to 0.6 mg/kg IV/SC/IM) and morphine (dose = 0.1 mg/kg IV or 0.5 mg hydromorphone IV) appeared to provide some level of analgesia in individual studies (compiled data were not reported), but no significant difference was demonstrated between the two agents. The study also reported rates of neuropsychological adverse events of 15.4% in the ketamine group and 4.4% in the opioid group (relative risk [RR] = 3.44, 95% CI = 1.81 to 6.55, absolute risk difference [ARD] = 11%, number needed to harm [NNH] = 9) Neuropsychological events were defined as agitation, hallucination, dysphoria, and confusion.

CAVEATS

The meta-analysis by Ghate et al. included more studies than the meta-analysis by Karlow et al. and a larger sample of patients (n = 609). However, the former included two observational studies, raising concern about the validity of the results and the appropriateness of pooling the data. The studies included in Ghate et al. varied significantly in the dose (0.1 to 0.6 mg/kg) and route of administration (IV/SC/IM) of ketamine, as well as in the use of adjunctive analgesia (one study’s protocol included a dose of midazolam with ketamine). The incorporated studies also varied in the choice of the compared opioid, with one study utilizing 0.5 mg IV hydromorphone and the others 0.1 mg/kg IV morphine.

Regarding the meta-analysis by Karlow et al., the clinical heterogeneity among the included trials is a major limitation. One of the studies included only patients with long-bone fractures, while the other two included patients with musculoskeletal pain and abdominal pain, raising the question of whether specific etiologies of pain may respond differently to specific analgesics. Patients receiving ketamine appeared to have more acute adverse events, but it was difficult to draw conclusions about harm endpoints because of the small sample size.

Neither of the two meta-analyses included data on the rate of administration of the medications, which has been shown to correlate with adverse side effects. It must also be noted that pain control commonly requires redosing and titration. Comparing a single dose of opioid analgesia to a single dose of ketamine might not be appropriate for determining the efficacy of either for pain control. In addition, this meta-analysis included only patients with acute pain; the efficacy of ketamine in patients with chronic pain (e.g., chronic back pain) is not addressed here. Finally, the small sample size of the included trials and the meta-analyses limits the validity of the findings.

In summary, ketamine appears to be comparable to opioids for acute pain control. However, because of the small sample size of the meta-analyses and limitations of the included trials, we have assigned a color recommendation of yellow (unclear if benefits) to this intervention. Larger high-quality studies are needed to further support the routine use of ketamine for pain control in the ED.

References

Friends, fellow triagers, colleagues, lend me your eyes—I come to bury Peacock et al., but also praise them. For this one article has added to the literature and furthered our understanding of the diagnostic utility of high-sensitivity troponin (hsTn), yet leaves us desirous of so much more. Allow me, in the paragraphs that follow, to explain.

As this investigative team aptly notes, uptake of hsTn in the United States has lagged compared to the rest of the world. While reasons for this lie primarily with approval delays at the Food and Drug Administration (FDA), there remains a reluctance in many institutions to shift from conventional assays to their more sensitive counterparts. When pressed, decision makers at such facilities point to the absence of U.S. data and feign uncertainty about potential accuracy in “their” patient populations. It is within this void that Peacock et al. present new(ish) data on Beckman Coulter’s high-sensitivity troponin I (hsTnI) assay, showing that it can reliably identify emergency department (ED) patients with suspected acute coronary syndrome (ACS) who have not sustained myocardial injury. Such work is important from a validation perspective, though incremental at best, providing information that is really a mirror image of itself (i.e., absence of a protein that reflects myocardial injury indicates absence of myocardial injury). Said another way, publication of data showing that the Beckman Coulter hsTnI assay can effectively rule out myocardial injury may be necessary to convince skeptics of its safety profile but the exercise is merely revealing an inherent truth based on known characteristics of the underlying pathophysiology and laboratory medicine. The fact that this analysis included an all-comer prospective cohort of ED patients does make the findings unique and adds to their applicability, although it does not alter this perspective.

However, there are nuances to be gleaned from the presented data that warrant further discussion. Foremost, the authors show that low risk, defined as a negative predictive value (NPV) for myocardial infarction (MI) of 99% or greater, can be identified using either the manufacturer delineated limit of quantification (LoQ) of 2.3 ng/L or the 10% coefficient of variation (CV) cut-point of 5.6 ng/L with the Beckman Coulter hsTnI assay. Given that reporting of hsTn in whole integers (rather than with decimals) is the industry standard, it would have been nice to see data with rounding up to 3 and 6 ng/L, respectively, although it probably would not have made much difference. Nonetheless, the high NPV for both cut-points was preserved on singular and serial measurement, provided for the latter that at least one result was below the LoQ or 10% CV and no repeat value was above the upper reference limit (URL) of 17.9 ng/L. Interestingly, as also shown by Peacock et al (see figures 1 and 2 from the manuscript), NPV was consistently 100% regardless of the pattern of hsTnI change for these patients, suggesting independence from the actual delta value. Such a finding stands in distinction from non-U.S. data where absolute deltas with this assay that exceed predefined thresholds of 4 ng/L at 1 hour and 5 ng/L at 2 hours, even among those with low initial troponins, were associated with acute MI. This may be attributable to the cohort enrolled by Peacock et al., where the prevalence of MI was lower (11.2% vs. 15.4% and 14.5%) and the time from symptom onset to first blood draw which exceeded 3 hours for the vast majority of enrolled patients. From a care delivery perspective, demonstration that the 10% CV cut-point provides acceptable safety is key as it enables rule-out in more than twice as many people (58% vs. 24%) compared to the LoQ.

Relevant to evaluating patients with suspected ACS, the work by Peacock et al. also reinforces a critical paradigm shift, namely, recognition that the URL, which has long been the cut-point for ruling out MI, is no longer sufficient. Regardless of the assay or
Although not necessarily mutually exclusive, the ability to modify either may not be directly influenced by hospitalization or further diagnostic testing making the question of “stay or go” perhaps a suboptimal one. Unfortunately, data beyond the initial hospital encounter were not collected, limiting interpretation of results to diagnostic potential. Alas poor Peacock (et al)! Prognosis matters too.

**Phillip D. Levy, MD, MPH**

(plevy@med.wayne.edu; philliplevy1@icloud.com)

*Department of Emergency Medicine and Integrative Biosciences Center, Wayne State University, Detroit, MI, USA*

Supervising Editor: Jeffrey A. Kline, MD

---

**References**


The word “emergency” is defined by time. We work in time-dependent windows of diagnosis and treatment. Consider the enormous impact of biomarkers and imaging now standard in the ED (troponin, D-dimer, and lactic acid and lightning-fast panscanning CT and bedside ultrasound). Faster detection of critical illness has also led to more ED administration of pharmacologic interventions (aspirin, thrombolitics, antibiotics, fluid resuscitation, pressors, TXA, four-factor PCC, etc.). Helicopter transport, intraosseous vascular access, and pelvic binders all improve time-dependent patient care.

The COVID-19 pandemic overwhelmed health systems due to an enormous number of critically ill patients presenting all at once. The disease had insidiously spread far wider than governments were aware. We have learned that the associated pneumonia also advances insidiously, and by the time patients present to the ED they have moderate to severe ARDS.

COVID-19 simply does not fit with our prior clinical experience with patients who have severe lung injury and hypoxia. Patients with COVID-19 pneumonia often have alarmingly low oxygen saturations (~50%–80%) but frequently do not feel short of breath. Patients who become acutely hypoxic, like those who choke or drown, rapidly become unconscious or seize. Respiratory failure patients with rapid-onset hypercarbia become narcotized and lethargic. Most patients we need to intubate in emergency have either precipitous hypoxia, hypercarbia, or shock that leads to compromised mental status. They have subjectively and clinically evident shortness of breath and dyspnea with increased work of breathing.

COVID-19 pneumonia patients often do not subjectively appreciate their lung injury. Through the virus’ effect on surfactant and resultant alveolar collapse, patients have a progressive drop in PaO₂ and an incremental increase in their respiratory rate. This process develops over days. The lungs initially remain “compliant” and patients effectively ventilate, lowering PaCO₂. They develop a large right-to-left shunt. I believe much of the lethality of COVID-19 has to do with the lack of subjective symptoms despite the advanced underlying lung injury that is occurring. It has been postulated byGattinoni et al. that the increase in respiratory drive exacerbates the inflammation and lung injury caused by the virus itself. Gattinoni has described the intact gas volume of the lung and high shunt fraction of COVID-19 pneumonia as an atypical form of ARDS, although there remains much debate about phenotyping COVID-19 cases.

Eventually, the cycle of worsening hypoxia, increasing respiratory rate, and the underlying lung injury precipitates overt respiratory failure. Acute respiratory failure, cardiac dysthymia due to severe hypoxemia, and thrombosis may explain the alarming number of COVID-19 patients found dead at home. The lack of subjective symptoms found in COVID-19 pneumonia also explains the many cases of incidentally discovered pneumonia in ED patients who present with syncope, fatigue, and other medical complaints. I wrote of this phenomenon that I called “silent hypoxia” in a New York Times opinion piece that was published on April 20, 2020.

In the two short months since COVID-19 exploded in our health care system, we have learned much about the disease. COVID-19 kills through its attack on the lungs in almost all patients. Thrombosis, renal failure, and neurologic injury largely correlate with severity of lung injury and also prolonged mechanical ventilation. The onset of pneumonia is between 5 and 10 days postinfection. Although there
are many laboratory abnormalities involving abnormal blood counts and inflammatory markers, the single most reliable marker of critical illness (ICU care, mechanical ventilation, and death) found in a large health care system in New York City involving more than 4,100 COVID-19 cases was the level of hypoxia on presentation.5

This month’s AEM article confirms the utility of home pulse oximetry monitoring as a screening tool for COVID-19 pneumonia.6 This study validates pulse oximetry for determining the need for hospitalization. It also confirms the phenomenon of silent hypoxia, because 50% of patients who returned requiring treatment for COVID-19 pneumonia in this study had no subjective worsening of symptoms. They only returned because of close pulse oximetry monitoring.

In all areas of emergency medicine, we know that earlier detection and intervention minimizes end-organ injury and improves outcomes. I believe that this will be shown with COVID-19 pneumonia too. Last month in AEM, Caputo et al.7 reported that awake proning and positioning maneuvers coupled with non-invasive oxygenation reduced the need for intubation in two of three patients with moderate to advanced COVID-19 pneumonia. Hopefully, such techniques will work even better if we identify pneumonia earlier with only mild hypoxia and before severe lung injury. This month’s study by Shah et al.6 supports the growing body of literature that pulse oximetry monitoring should be a standard of care for discharging known or suspected COVID-19 patients.

We just crossed the 150,000 dead mark. But amidst the pessimism of this pandemic, this study of home oximetry monitoring coupled with last month’s AEM publication on proning points to progress and hope. We have learned much about how COVID-19 kills in a short period of time. It would be great if magic bullets arrive that can stop this virus instantly or vaccines appear that prevent further infections. In the short term, though, we must focus on incremental gains related to COVID-19 pneumonia: earlier detection through pulse oximetry and supportive and adjunctive care that reduces the need for mechanical ventilation.

Richard M. Levitan, MD
(airwaycam@gmail.com)
Littleton Regional Healthcare, Littleton, NH

Supervising Editor: Jeffrey A. Kline, MD

References
Guidance and Patient Instructions for Proning and Repositioning of Awake, Nonintubated COVID-19 Patients

Prior studies on proning awake, nonintubated patients with hypoxemic acute respiratory failure, as well as evolving study of similar COVID-19 patients, coupled with experience and dramatic anecdotal evidence from the COVID-19 pandemic, suggest the importance of proning all such patients with COVID-19 to improve oxygenation and reduce respiratory effort. Literature and experience from health care teams in the midst of the pandemic suggest that any COVID-19 patients with respiratory compromise severe enough to warrant admission should be considered for proning. We additionally suggest that these patients should be considered for proning as well as ongoing patient repositioning (e.g., right lateral decubitus, seated, and left lateral decubitus positions). Figure 1 represents the proning and positioning instructions developed at New York City Health + Hospitals/Elmhurst, a large, inner-city, tertiary public hospital in the epicenter of the COVID-19 pandemic in New York City and later adapted and utilized at facilities across the United States.

Additionally, we suggest the use of these proning and repositioning instructions for mild and discharged COVID-19 patients to be completed independently. Mild COVID-19 patients often still have respiratory involvement that may benefit from these exercises. Studies are needed to evaluate if it may also stave off disease progression. Like all medical interventions, it remains a risk/benefit analysis and such positioning most commonly presents very little risk in appropriately selected patients. The Intensive Care Society Guidance for Prone Positioning of the Conscious COVID-19 Patient includes the following absolute contraindications: acute respiratory distress (requiring higher level intervention, e.g., immediate need for intubation), hemodynamic instability, agitation or altered mental status, unstable spine, thoracic injury, or recent abdominal surgery. Relative contraindications to consider include facial injuries, neurologic conditions (e.g., seizure disorder), morbid obesity, pregnancy due to gravid abdomen, and pressure sores/ulcers or high risk for pressure sores/ulcers due to positioning.

New York City Health + Hospitals/Elmhurst successfully developed and implemented the proning and positioning guide with awake, nonintubated patients as well as provided it to patients suspicious for or with confirmed COVID-19 who were discharged from the emergency department (ED). The guide was developed via expert consensus of an interdisciplinary team of emergency medicine physicians and physical therapists and iteratively revised based on usability testing and understandability feedback from providers and patients, with whom it was piloted through several rounds of revisions. Distribution of the guide was implemented by the emergency medicine team and physical therapist collaborating in the department and provided to ED patients awaiting admitted bed availability and was utilized by the majority of appropriate patients in the ED independently (occasionally with verbal reminders). Majority of awake, nonintubated moderately ill COVID-19 patients in the ED are on continuous pulse oximetry monitoring and patients have oxygen saturation, oxygen requirement (e.g., number of liters oxygen via nasal cannula), and position recorded hourly on a sheet attached at patient bedside (Figure 2). A vital signs documentation sheet was in use in the department and a column was added to denote patient positioning upon development of this proning guidance.

The authors have no relevant financial information or potential conflicts of interest to disclose.
Instructions for patients with cough or trouble breathing:

Please try to **not** spend a lot of time lying flat on your back! Laying on your stomach and in different positions will help your body to get air into all areas of your lung. You may notice improvement in breathing immediately or several minutes after positioning change. Please do not stay in any position that causes discomfort or pain; skip such positions in the rotation. It is most important you do not just lay flat in bed and this guide is designed to help you change positions in bed.

Your healthcare team recommends trying to change your position every 30 minutes to 2 hours and even sitting up is better than laying on your back. **If you are able to, please try this:**

1. 30 minutes – 2 hours: lying on your belly
2. 30 minutes – 2 hours: lying on your right side
3. 30 minutes – 2 hours: sitting up
4. 30 minutes – 2 hours: lying on your left side; then back to position #1.

PHOTOS BELOW TO DEMONSTRATE THIS:

1. 30 minutes – 2 hours: laying on your belly

2. 30 minutes – 2 hours: laying on your right side

3. 30 minutes – 2 hours: sitting up

4. 30 minutes – 2 hours: laying on your left side

Then back to Position 1. Lying on your belly!


Figure 1. Patient handout.
<table>
<thead>
<tr>
<th>DATE</th>
<th>TIME</th>
<th>02 SAT</th>
<th>PULSE</th>
<th>02 Requirement/setting</th>
<th>Patient Positioning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P=prone; L= L sidely; R = sidelying; S= supine</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P     L   R     S</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P     L   R     S</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P     L   R     S</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P     L   R     S</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P     L   R     S</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P     L   R     S</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P     L   R     S</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>P     L   R     S</td>
</tr>
</tbody>
</table>

**Figure 2.** Vital signs and positioning flowsheet.
For patients in the ED, inclusion criteria for recommending self-conducted repositioning and proning began if oxygen saturation was < 92% (including if < 92% on room air prior to improved saturation with oxygen delivery while in ED) and patient deemed to be independent in bed mobility (Figure 3). Patients
requiring noninvasive ventilation (BiPAP or CPAP) in the department were generally encouraged to side lie and sit up; however, proning and repositioning was accomplished with many of these patients in conjunction with direct assistance by the physical therapy team or other health care team members due to concern about mask or equipment dislodging with proning independently.

The procedure and highlighted benefits should be explained to the patient, emphasizing goal to maintain each position for 30 minutes to 2 hours. Specific illustrated instructions were provided via patient handout (Figure 1). If a patient does not tolerate lying prone or a specific position, then alternative positions are encouraged and patients are reminded not to stay in any positions of discomfort. Ten to 15 minutes are allotted for adjustment in oxygen saturation and saturations and positions were recorded as noted above (Figure 2). If no improvement in oxygen saturation with a change in position, escalation of medical care is necessary. Generally, patients that desaturate, look uncomfortable, or request assistance to change position should all be urgently repositioned with repeat saturation monitoring until patient is in a position of comfort with improvement in saturation. It is prudent for staff to check for integrity of oxygen delivery device (e.g., position and length of tubing) after position changes or immediately if oxygen desaturation. Additionally, using pillows or rolled sheets and towels can increase comfort and relieve pressure on bony prominences. All members of patient care team must be made aware patients have been encouraged to self-position and prone.

In conclusion, awake, nonintubated patients appear to greatly benefit from self-proning and alternating positioning, with many being able to independently change positions without disrupting flow of oxygen. Special consideration for position changes in patients requiring noninvasive ventilation (BiPAP or CPAP) is required but patient positioning is similarly achievable in such patients with direct assistance from physical therapists or other health care team members.

Suzanne K. Bentley, MD, MPH1,4 (bentleys@nychhc.org)
Laura Iavicoli, MD2,4
David Cherkas, MD2,4

1 Departments of Emergency Medicine and Medical Education, Icahn School of Medicine at Mount Sinai
2 Department of Emergency Medicine, Icahn School of Medicine at Mount Sinai, New York
3 Department of Rehabilitation Medicine, NYC Health + Hospital/Elmhurst
4 Attending Physician, NYC Health + Hospital/Elmhurst, Elmhurst, Elmhurst, NY, USA

References

In Response to “The Novel Use of Home Pulse Oximetry”: An Australian Offer of Support

To the Editor,

We write to thank Shah et al. for sharing their article “Novel Use of Home Pulse Oximetry Monitoring in COVID-19 Patients Discharged From the Emergency Department Identifies Need for Hospitalization” and provide an international perspective from Australia.

From our antipodean vantage, we have watched with concern the rapid spread of the disease through North America and the devastating toll it has taken on colleagues, health infrastructure, and the broader community. The combination of the adoption of early public health interventions and our relative isolation have left us in good stead to attenuate, or at least defer, the ravages of this disease domestically, just as they did for us during the 1919 influenza pandemic.

Independent of Shah’s work, we have established a cache of local pulse oximeters to service our community particularly in the event of a significant surge of patients, with similar aims to identify silent clinical deterioration and rationalize emergency department attendances.

The lack of an early clinical surge has allowed us to develop a computer application that prompts patients via SMS to enter their vital signs including oxygen saturations via a smartphone, notifies clinicians of clinical deterioration, and advises patients to present to emergency based on certain triggers, all while conserving direct clinician interaction.

Our system was built within the REDCap research electronic data capture platform, and we offer it free for use and adaption to any health institution that feels it may be of utility. We hope this may be of assistance to our international colleagues.

Martin Dutch, BMedSci (Hons), MBBS (Hons)1
(martin.dutch@mh.org.au)
Jonathan Knott, MBBS, PhD, FACEM2
1Emergency Department, Royal Melbourne Hospital, Parkville, Victoria, Australia
2Clinical Sub-Dean (Emergency Medicine), University of Melbourne, Melbourne, Victoria, Australia

Supervising Editor: John Burton

References

The Editor:

We read with great interest the study by Bijur et al.\(^1\) that evaluated the analgesic efficacy of intravenous (IV) acetaminophen as an adjunct to IV hydromorphone for the treatment of severe, acute pain in the emergency department (ED). The study found that there was no significant difference in analgesia when 1 g of IV acetaminophen was added to 1 mg of IV hydromorphone. While we are in complete agreement with the authors about the utilization of combinations of analgesics from different therapeutic classes, we were surprised by the study of these two specific drugs.

Intravenous acetaminophen is a nontitratable and weak analgesic that does not result in opioid sparing or reduce opioid-related adverse effects. Additionally, it does not provide analgesic superiority over the oral and rectal routes beyond the initial 15 minutes after administration. Furthermore, a prohibitive acquisition cost would preclude most EDs from utilizing this medication routinely for pain control. The addition of a single dose of this weak analgesic to an extremely potent opioid is unlikely to affect the analgesic efficacy of the latter and can be misinterpreted to bolster support for this opioid.

The authors acknowledge that reducing opioid use and exploring nonopioid modalities are important in curbing opioid-related consequences. However, the choice of opioid is also important, as individual opioids have different abuse liability and adverse effect profiles. IV hydromorphone's high lipophilicity allows it to cross the blood brain barrier rapidly, resulting in euphoria and reinforcing effects. As such, hydromorphone has a higher potential for addiction and abuse than other opioids such as morphine. Previous research has demonstrated that oral and parenteral hydromorphone produced significantly more euphoria and reinforcement than morphine in individuals with a history of opioid dependence. These findings have also been replicated in individuals without a history of opioid use disorder. Additionally, hydromorphone has had a 438% increase in nonmedical use from 2004 to 2011 and a significant increase in street value compared to less abuse-liable opioids.\(^3\) Hydromorphone is used as a heroin substitute in “heroin” vending machines in Vancouver, Canada.\(^4\)

In addition to increased recreational use potential, parenteral hydromorphone causes significantly higher rates of excessive sedation, hypoxia, respiratory depression, and the need for naloxone administration compared with other opioids. Many of the adverse events have been attributed to excessive hydromorphone dosing.\(^3\)

At this time, we do not know how the acute parenteral administration of hydromorphone contributes to subsequent opioid use or addiction, but there are indicators that hydromorphone is more dangerous than other opioids. Hydromorphone should not be the first-line opioid used in the ED. While we do advocate for continued research examining combinations of opioid and nonopioid therapies for both efficacy and harm from opioid analgesics, we hope that future studies will focus on studying less abuse-liable but equianalgesic opioids as part of opioid stewardship efforts in cost-conscious manner.

Maryann Mazer-Amirshahi, PharmD, MD, MPH, PhD\(^1,2\)
Sergey Motov, MD\(^3\)
Lewis S. Nelson, MD\(^4\)

\(^1\)Georgetown University School of Medicine, Washington, DC

© 2020 by the Society for Academic Emergency Medicine
doi: 10.1111/acem.13990
ISSN 1553-2712
793
References

Response to Mazer-Amirshahi et al, Intravenous Acetaminophen and Hydromorphone: The Bad and the Ugly of Emergency Department Pain Management

Intravenous (IV) administration of opioid analgesics is the mainstay of treatment of severe pain in the ED. The search for adjunctive analgesics is due to the desire to limit exposure to opioids and reduce side effects at the same time as increasing pain relief. While agreeing with the premise of combining analgesics from different therapeutic classes, Mazer-Amirshahi et al. question the choice of IV acetaminophen and IV hydromorphone.

We feel that several of their reservations are not consonant with existing literature. In contrast to the assertion that IV acetaminophen does not result in opioid sparing, a 2016 Cochrane review of 75 randomized trials of postoperative use of IV acetaminophen found that there was a 26% reduction in opioid use over the first 4 hours following surgery. At the outset of our study it was not known whether the postoperative experience with IV acetaminophen would transfer to the treatment of acute, severe pain in the ED. The lack of empirical data about whether an effective strategy worked in the ED setting provided the rationale for the study.

In an earlier study, our group performed a head-to-head comparison of IV acetaminophen to IV hydromorphone. As would be expected, hydromorphone conferred more analgesia (reduction in an 11-point numerical rating scale of 5.3) than IV acetaminophen. However, the reduction associated with acetaminophen was still substantial (3.3 NRS units). As the analgesics come from different therapeutic classes one might expect some degree of additivity resulting in better pain control. We agree with Mazer-Amirshahi et al. that the high cost of IV acetaminophen is a serious barrier to its wholesale adoption. However, if the effect of the two analgesics were found to be additive the cost might be worth the pain relief for patients with the most intense pain such as pain from kidney stones, ovarian torsion, aortic dissection, or long-bone fractures.

As with almost all medication there is practice variation in the treatment of acute pain in the ED. In national data hydromorphone is the most frequently administered IV analgesic following IV morphine. Analysis of data from the National Hospital Ambulatory Care survey provided estimates of the percent of all visits to the ED during which either of these analgesics was prescribed. In 2010 morphine was administered to patients during 6.2% of all ED visits, while hydromorphone was administered at 5.3% of visits. A meta-analysis of eight studies that compared morphine and hydromorphone found hydromorphone to be slightly more effective than morphine, a difference that was statistically significant but of small magnitude. Both had similar frequencies of adverse effects. It is unlikely that the inferences from our study would be any different if the opioid used had been morphine.

We disagree with the contention that hydromorphone is an unsafe medication when used in an ED. Our group has performed a series of studies of IV hydromorphone involving more than 1,000 patients. No patients in any of these studies needed naloxone to reverse the effect of the analgesic and fewer than 1% required supplemental oxygen.

We also disagree with the assertion that hydromorphone is uniquely prone to abuse. The U.S. Drug Enforcement Administration describes the pleasurable effects of hydromorphone including euphoria and its
abuse potential as being “similar to other schedule II opioids.” We are aware of the data referenced by the correspondents in which healthy pain-free volunteers and patients with opioid use disorder but without acute pain rated the pleasurable sensations induced by various opioids. We believe these data have little relevance for ED patients in acute severe pain.

It is clear that use of oral opioids is associated with unwanted sequelae including ED recidivism, worsening of underlying pain disorders, and ultimately opioid use disorder. There is little doubt that diversion of discharge prescriptions for oral opioid analgesics, notably hydrocodone and oxycodone, from the ED is a serious problem. In contrast there is little opportunity for diversion of the formulation of hydromorphone that is used in IV administration by ED patients.

We agree with Mazer-Amirshahi et al. that IV acetaminophen is not a promising adjunct to opioid analgesics in the ED, given the lack of additional analgesia and its high cost. In the clinical setting we believe the choice of opioid is at the discretion of the treating provider.

Polly E. Bijur, PhD, MPH
(polly.bijur@einsteinmed.org)
Benjamin W. Friedman, MD, MS
Andrew K. Chang, MD, MS

1Department of Emergency Medicine, Albert Einstein College of Medicine, Bronx, NY

2Department of Emergency Medicine at Albany Medical Center, Albany, NY

Supervising Editor: Jeffrey A. Kline, MD

References


Prehospital Care at the Epicenter of a Pandemic: The New York City EMS Response

The New York City 911 system recorded 5,700 medical calls on Tuesday, March 24, 2020, the most in recorded history. On March 25, the call number increased to 5,800, then 6,200 calls on March 27 and 6,500 calls on March 30.

By April 2, over 3,000 FDNY members were on leave for illness. A workforce reduced by nearly 20% was therefore charged with responding to a greater than 40% increase in calls. The average prepandemic response for life-threatening medical emergencies was 6.5 minutes. In April, there was a 9.5-minute average response time for medical calls in NYC. In the Bronx, it was closer to 11 minutes.

Through the pandemic surge, EMS protocol changes were implemented to protect first responders and improve response to increasing call volume. However, according to many paramedics, these revisions added confusion and frustration to an already disorienting and frightening time and undermined morale as many felt helpless, restricted from intervening as patients deteriorated in front of them.

On March 6, the NYC EMS advisory committee, composed of EMS physicians, published advisory 2020-03 recommending all nebulized medications be administered via breath-actuated nebulizer, a delivery device that limits aerosolization into the environment but requires patient cooperation and is likely less effective in the critically ill.

On March 17, advisory 2020-04 recommended that EMS interview patients from six feet away and place a surgical mask on all patients with infectious symptoms, altered level of consciousness, or in cardiac arrest. The document further recommended that EMS providers wear face masks, eye protection, and gowns (if available).

On March 20, advisory 2020-05 recommended supraglottic airways in all situations when endotracheal intubation is indicated. Intubation is considered to be the highest aerosol risk of EMS-performed procedures.

On March 30, advisory 2020-07 temporarily reduced staffing standards for disaster response, permitting an ALS ambulance to respond with one paramedic and one EMT. A BLS ambulance was permitted to operate with one EMT and one certified first responder. Prior to this policy, NYC ALS ambulances in the 911 system were required to have two paramedics and BLS ambulances required two EMTs, but staffing shortages during the pandemic made this impossible.

On March 31, advisory 2020-08 implemented temporary cardiac arrest standards for the pandemic disaster response. To reduce the risk to EMS and downstream providers, adult patients in nontraumatic or blunt traumatic arrest were not to be transported to the hospital unless there was ROSC or a medical control order.

On April 1, advisory 2020-09 implemented the EMS Viral Pandemic Triage Protocol. It stipulated the following criteria for not transporting patients, i.e., leaving patients at the scene: Patient age less than 65, heart rate less than 110, systolic blood pressure greater than 100, respiratory rate less than 22, and SpO₂ at least 95%.

On April 17, advisory 2020-10 dictated that if EMS did not witness the cardiac arrest and BLS or ALS did not find a shockable rhythm, resuscitation should not be initiated. If cardiac arrest was witnessed or a shockable rhythm was present, resuscitation should be terminated after 20 minutes.

On April 23, after alarming discontent with advisory 2020-10, advisory 2020-11 reverted the NYC region’s policy to resuscitate patients in compliance with advisory 2020-08.

The authors have no relevant financial information or potential conflicts to disclose.
On April 27, advisory 2020-12 rescinded the implementation of the EMS viral pandemic triage protocol. Maimonides Medical Center (MMC) is the largest hospital in Brooklyn. The 14 MMC ambulances in the 911 system typically respond to between 0 and 1 cardiac arrests per day. Starting on March 11, this increased to a peak of 19 cardiac arrest calls on April 2 (Figure 1).

Below are the reflections of several Maimonides EMS providers who worked through the COVID-19 pandemic.

**Christopher Oliveri**

Encountering the first critical patient who was infected with SARS-CoV-2, I was faced with a difficult choice: Place her on CPAP and increase my risk of viral exposure or watch her suffocate. After placing her on CPAP and seeing no improvement, I knew I needed to transport her quickly to the hospital. I didn’t have IV access to sedate and intubate her and I soon became very frustrated. While she was in our ambulance, I tried to rush her family out so that we could start transport. After asking them several times to leave while watching my patient deteriorate, I raised my voice so they would get out of the ambulance before she died right there. When I arrived at the hospital, I was shell-shocked. I saw similar patients in her same condition and I then knew then that she wouldn’t recover. She was going to die in the hospital alone. Her family’s last moment with her were cut short by me. This weighs heavily on me now knowing that giving her family a few extra minutes wouldn’t have made a difference for my patient’s outcome but would have made a difference for them.

**Jill Eby**

One thing I am hearing a lot of around our EMS garage is, “Coronavirus took away our ability to have a win.” EMS providers are always searching for a win... Getting ROSC, catching a STEMI on EKG, reversing hypoglycemia, getting a stroke patient to the ER on time, getting the tube. With this virus, there was no win... no chance. We understood that the frequent changes in EMS protocols were implemented as a safety measure for first responders but it still broke our collective hearts every time we couldn’t try our very best to bring back someone’s loved one. Three or four times a shift we had to tell someone that their loved one died without feeling confident that we did absolutely everything for them. That will stick with each and every one of us for a very
long time. While at work, the virus took our hands and our minds away from us while instilling fear that we will kill our own family members if we go home and hug them. It hit us from all sides and I fear it will have a lasting effect on many first responders, even if they can’t or won’t admit it. That is the culture, unfortunately. We didn’t falter during this pandemic because that’s not what we do. Seeing the 911 call numbers go down and being able to actually help again—that’s what will help us recover, I’m sure of it.

**Kelly Vyater**

When going into EMS, everyone says “you know what you signed up for.” Honestly, a pandemic was not something I thought would be waiting for me. Walking into work everyday, I felt more and more ethically challenged. There were constant protocol changes asking me to deviate from what I was taught. I attempted to explain to patients that the ER, which is normally seen as a safe place, is not so safe these days.

Going into EMS, you have to be emotionally resilient. I’m usually the “roll with the punches” type of person. After working up a cardiac arrest, I’ve seen families cry as we tell them that we did everything we could and that takes a toll on you the first couple of times. But after a while, it becomes second nature. However, watching a family cry during this pandemic as they hug and kiss their loved one as we wheel them out the door, watching the families realize it will probably be the last time they see their loved one, that is the most psychologically painful thing I have ever experienced as a paramedic.

**Avi Merl**

This was in one word crazy. I have never been this physically and mentally drained ever in my life as I was through the prime time of COVID-19. There was a 40-year-old man who reported a fever and not feeling well. I applied a pulse oximeter which revealed a pulse oxygenation of 58%. The first thing I thought was, “that’s impossible.” The sheer amount of death we responded to was never something that I’d ever thought I see. Listening over the EMS radio, “Central, it’s an 83 [patient pronounced dead]”; I heard that more times than I have heard it through my entire 25-year EMS career.

I had to explain to multiple family members that they can’t come with us as we transported their loved one to the hospital as they won’t be allowed in. The worst sight for me was a little, old Italian man in his early 90s with an oxygen saturation in the 50s and his wife of 63 years scrambling around to make sure he had his hat and everything else he needs to go to the ER. And then as she grabs her coat, we explain to her that she can’t accompany him because it’s unsafe for her. She became visibly upset and said to me, “We are married 62 years. He has never gone to the doctor alone and he doesn’t know what to say.” I stood there knowing that this man is going to die alone in the hospital. And the chances are that if he’s sick with COVID-19, his wife is going to be as well. I hope we never see anything like this in our time again.

**Rebecca Solomon**

In the middle of March 2020, life changed not just for us in health care but for everyone. My kids went from riding the bus to school to home school and online classrooms. Every day I went to work armed with new information on how this disease is making us, our colleagues, and friends sick. We have questions with no answers but we are the frontline and need to keep going. Our communities rely on us as we are the ones they seek answers from in a world that no longer makes sense.

As a paramedic in NYC during this pandemic, I watched how COVID-19 changed us. Our PPE protocols changed, our patient care protocols changed, our hospitals changed to meet the demands of a system that is being stressed to the point of nearly breaking. The worst part for me is our patients are changing. They are dying and I don’t think they know it. As we respond to the residences of our patients, the families turn to us for answers. How do you tell the family that you believe their loved one has this virus? How do you tell the family that they may also have COVID-19? There is a good chance that once I transport a patient to the hospital, the family may never have their chance to say goodbye. This patient is going battle the virus alone and die alone.

There has always been a certain solace to being able to debrief “the bad” with your colleagues but that too is different. We are triaging patients in tents outside the hospital and we are covered in so much PPE that we
barely notice our colleagues anymore. All you see when you look at the frontline is their eyes. Eyes that have
now seen unprecedented amounts of suffering and death.

Shmuel Gajer
There was no heroism done. Just people reacting poorly in a time of confusion. Campaigns are being run to
publicize the job well done. A bunch of people at the top will pat themselves on the back. There are shoutouts
to others for a job well done to make everyone feels good about themselves. Articles and blogs are being posted
hourly to assist with people’s insecurities and make them feel as if they did or are doing something about the
situation. However we failed big time. In the end, not much will change. Next time around it will be the same.

There were constant changes in protocols and procedures that served just to increase our anxiety and confu-
sion. We watched the dying die alone. We watched other health care workers refuse to do their jobs and help
the sick appropriately. There was a lack of organization and communication by leadership, management, and
politicians. We didn’t know how to properly treat the sick and needy. We observed our coworkers getting really
sick. We didn’t have the ability to socialize and decompress with family and friends, which was especially rough
in these times. We didn’t transport some legitimately sick people because they weren’t sick enough. We didn’t
feel safe due to poor PPE or psychologically safe due to how one is perceived by those who are supposedly look-
ing out for us. We were fatigued and felt like we were operating blind in the darkness with no light in sight.
Although at times we were too busy to mentally process what was happening, we mostly felt that we were left
alone to fend for ourselves as we did procedures outside of our job description and scope of practice.

Henoch Junik
As a veteran soldier, I felt the call to action stronger than some. I felt like it was a time to put your head down
and push on no matter what. I worked every shift that was available and sometimes had only a few hours of
sleep between them. To me, the greater sacrifice was done by my family: my wife taking care of our six kids, all
under six, without me and my kids not seeing their dad for days in a row. When I was home, my mind was
really at work researching what news was real and what was fake so I could stay informed and safe. I feared for
my family’s safety more than my own.

As the call volume raised to obscene numbers, the NYC 911 system responded to more calls in one day than
some EMS systems respond to in a whole year. Like war does to other industries, we began to see some "ad-
vances" to prehospital care. We were finally given the authority to act like the clinicians that we were originally
trained to be. Working in a city that never sleeps, you’re bound to have some anxiety in your life. For many, that
anxiety came to the forefront when this pandemic began. Some EMS providers thought we were crossing ethical
lines by leaving patients at home that we would have transported before the pandemic. I fear that the anxiety will
stick with us for a long time and will change the way we interact going forward. However, I hope we focus more
on what worked and help expand prehospital care. My military experience prepared me for the exhaustion I felt.
It did not prepare me for my kids asking me when will I be home or for homeschooling and raising kids in this
uncertain world. During the pandemic, most people got to spend more time with family. However, frontline
workers lost time with their families and they will never get it back. And for me that is a great loss.

In addition to Ms. Eby, Ms. Vyater, Mr. Merl, Ms. Solomon, Mr. Gajer, and Mr. Junik, we’d like to
acknowledge the entire NYC EMS workforce, who worked at the front of the frontline, at the epicenter of what
we hope will be a once-in-a-lifetime viral pandemic.

Matt S. Friedman, MD
(Mattsfriedman@yahoo.com)
Reuben J. Strayer, MD
Department of Emergency Medicine,
Maimonides Medical Center,
Brooklyn, NY, USA

Supervising Editor: Brian J. Zink, MD
References


“Will buddies.” That was an unfortunate phrase a friend of mine and I coined together during a shift early in the pandemic. He was a resident in our program just a couple of years ago and is now one of my attendings. While we were working a shift in the ED swarming with critically ill COVID patients, it came up in the “doc box” that we were both drafting our wills. How’s that for a view into the dark places a lot of us have gone during the pandemic . . . and we’re both young and otherwise healthy. We have little kids and wanted our spouses to know we were looking out for them if anything happened to us, despite the little in savings I have still being a resident. One problem was that I couldn’t find anyone to notarize mine. He mentioned, “Don’t worry. I have someone. We can get them notarized together.” “Great, we can be will buddies,” I replied and we both laughed the cynical laugh a lot of ED docs use to cope in the difficult times we aren’t strangers to.

Just two weeks before this, COVID cases started to mount in New York City and the feel in the ED changed almost overnight. Within one or two days, the typical mix of urban ED patients—strokes, vasculopathies, heart attacks, gun shot wounds, chronic pain patients, psychiatric patients, GI bleeds—virtually disappeared and was replaced by a flood of hypoxic patients. Every corner of the ED was packed with patients hooked up to oxygen tanks, dispersed among the critically ill on ventilators or those about to be intubated surrounded by a swarm of people donning extra PPE for the high-risk procedure, and the patients just kept coming. As I hurried through the ED to evaluate yet the next patient, I would keep an eye on the pulse ox monitors, seeing confused patients, alone and scared, pulling off nonrebreathers and nasal cannulas with sat’s in the 70s and try to redirect them and get them settled in on their oxygen before getting to that next patient. We were all scrambling to titrate pressors, vent settings, oxygen levels, and call families for heartbreaking goals of care conversations or FaceTime them in to see their loved ones for what was often and likely the last time. The tide was unrelenting.

During that first week we started seeing COVID cases, I began what has since become a “new normal” decontamination routine when I got home. I’d undress in the garage, put all my clothes in a cloth laundry bag, bound it tightly, immediately shower in the basement and wipe down my glasses and cell phone a second time with Clorox. I was sleeping in that same basement and trying my best to distance myself from my two young boys, Finn (almost 5) and Roan (almost 3). That’s a hard thing to even think back on, distancing but still being present with my little guys. I’d sit in the living room or outside while they played but couldn’t hug, cuddle, tickle, kiss, wrestle with them, or even prepare their food. I was stricken with anxiety and guilt knowing they were potentially at risk being around me. The same with my husband. I was a pariah in my home. A vector for disease. So when I finished a string of five night shifts that first week, had slept a little and cried a lot, I helped pack up my kids, husband, and in-laws to head to the burbs in New Jersey to keep them safe from me and the explosion of COVID in the city. I drew big crayon notes for each of my kids telling them how much I love them that I tucked in their bags, made sure they had their favorite toys and lovies to sleep with, and gave them each a big hug while fighting back tears and telling them I’d see them real soon. I’ve never felt my heart break as it did that day because deep down I was fighting back the fear that could be the last time I ever see them again.

It feels like I’m being overly dramatic because that couldn’t possibly be true. But the sad reality is, it could. None of us know what straw we’ll draw if we get infected, and working in the thick of the storm day in and day out makes it difficult to see beyond that reality at times. We treat patient after patient
coming in critically ill with COVID: young, old, healthy, and those with any multitude of medical problems. We’ve watched colleagues fall ill, be intubated and linger in the ICU for weeks. Some get better and a lot don’t. And that has shaken me to my core. I think we, as physicians, often put up a mental divide between ourselves and our patients, never letting ourselves believe for too long that we too could fall prey to the sometimes awful things we see our patients afflicted by. But this disease reaffirms in me everyday just how human and vulnerable we all are. I don’t let myself linger on these thoughts long though as there is just too much work to do and too many patients to try to help, and I do feel a great privilege and responsibility being among those who can help.

This new disease levels the playing field so vastly in health care as we’re all learning about it for the first time together, residents and attendings. That’s changed the dynamic of the resident experience. I’ve never been asked to come to the bedside with an attending so often to troubleshoot and think things through together on such an even plane. Sometimes it’s crazy to think we both bring the same amount of experience with this disease to the table, having read as much as we can as fast as we can over the past couple of months everywhere we can, from our accrediting body websites to physician group chats on Facebook. We’ve seen the same numbers of patients with the disease. On second thought, given we spend so much more time in the ED as residents, we’ve likely seen more patients with this disease.

In my first couple of years of residency, I took solace in my training knowing a seasoned attending was often just an arms length away when a patient was crashing and burning and I just couldn’t figure out what to do next. I had that sense that thank God there’s someone here who really knows what’s going on, someone who has seen this or something like this before and has control of what to do when everyone else senses a totally out of control moment. But COVID is utterly different. I don’t have that ultimate backstop anymore. On the one hand that feels good and exciting as we all find our confidence in the practice of emergency medicine, and on the other hand it is terrifying knowing we’re all to some extent floundering, doing the best we can against a new enemy we admittedly know little about.

I chose emergency medicine excited about the thrill and challenge of not knowing what was going to come through the door on any given day or night and the challenge of having truly undifferentiated patients. And I accepted the risks that come with that: violent patients with weapons, patients with toxic or caustic exposures they bring into the ED with them, infectious diseases like TB or measles. But I never contemplated this new paradigm of putting my family in harm’s way. Becoming an emergency medicine doctor has been a lifelong dream and I’m humbled to be included in the ranks of such a group of BAFERDs, but my family didn’t sign up for this.

Liza Hartofilis, MD
(liza.hillenbrand@gmail.com)
Department of Emergency Medicine, NewYork-Presbyterian Hospital, New York, NY, USA
Supervising Editor: Brian J. Zink, MD
“Nassim, come here quickly and get dressed,” his young mother said in Daari as she handed him his white salwar kameez (a traditional South Asian loose pajama).

She needed Nassim, her only child, to get naan from the market near the American base. They were poor and lived in a traditional mud hut village in Eastern Afghanistan. With black tufted hair from days without bathing, Nassim had a boyish face, hazel eyes, and a soft smile. The details of who had killed Nassim’s father remained vague, and his name had been added to an ever-growing list of forgotten civilian casualties. Despite the daily hardships of poverty, Nassim still enjoyed simple pleasures such as playing soccer. His mother was nervous for some reason, fearful that he might be harassed at either the Afghan Army or Taliban checkpoints, and she wanted him to go to the market and return quickly. Grinning, Nassim squeezed her hand before running towards the market, waving goodbye as he came upon the main highway.

I felt the blast wave, like an unwanted soft touch on my cheek. Immediately, my senior medic and I peered in the direction of the explosion. From our small base in Eastern Afghanistan, over the defensive barriers, we saw a rising rat tail of black smoke, polluting the clear blue sky and the imposing snow-capped mountains in the background. I went back into my aid station, a pitting anxiety filling my stomach. Eight months into my first combat deployment as the physician for our base, I already knew what had happened.

Minutes later, a U.S. military advisor sprinted into my aid station, gasping for breath, and said, “there has been a VBIED (vehicle born improvised explosion device) outside of the base Dr. Shukla! There are reports of multiple Afghan casualties!”

“Ok,” I replied grimacing, “let me get my team spun up.” I felt like a cauldron of adrenaline and emotional exhaustion. Dealing with multiple casualties was not new to us. Some of the highest levels of violence had been enacted by both sides that year in hopes of leveraging a peace deal. It became such a daily routine that when our base did not get attacked, I would become anxious. Paradoxically, my anxiety would dissipate when enemy rockets flew in our direction.

We drove to the forward surgical team about 2 miles away where I saw the emergency medicine physician standing among the casualties as they were being laid out in rows, the semblance of organization in the midst of chaos. As I looked at the injured, at least 30 people, it seemed like one large Afghan family. The females wore elaborately tailored, bright Afghan dresses with decorative small mirrors sewn on; polka-dots of blood violated the symmetry of their fashion. The screaming children who had just been loaded off the MEDEVAC helicopters clung to their mothers or older siblings, disheveled and dusty. Here they were, Afghan civilians, the forgotten, unwilling victims of this conflict.

Over the basewide intercom, a metallic and distant voice rang out, “MASCAL, MASCAL, MASCAL,” alerting the residents of the base of a “mass casualty event” and requesting both medical and nonmedical personnel to help. As the whirlwind of casualty treatment slowed, disposition plans became formulated, and the other physicians and I determined that the patients would be delivered to the Afghan Army to obtain their follow-up care. As I had developed a working relationship with the Afghan Army physician, it was decided that I would escort the patients to the exchange point. With all the patients loaded onto a large flatbed truck, one of the physicians mentioned a
casualty that still needed to be picked up, the only one who had died: a child. The other Afghans had seen the boy’s body on the MEDEVAC helicopter but did not recognize him. To make matters worse, the child had no form of identification to assist in finding his parents.

“Right, let me hand off these causalities and then I will come back for the kid,” I told the group of physicians.

We drove to the exchange point where I met a different Afghan Army physician. Younger and dressed in a track suit, he seemed annoyed as I pointed to the strips of medical tape that had been placed on each casualty’s sleeve to help in the handoff. As the patients were moved into the waiting ambulances, I told the Afghan Army physician about the dead child and to wait while I fetched him.

We then drove to pick up the child.

I walked into the freezer, a dark shipping container outfitted with an industrial cooling system, and saw an ivory white body bag lying on a wooden pallet as the sole occupant. I picked up the body bag, walking with it draped over my arms like a curtain. Probably between 10 and 12 years old, the child was not heavy but his entire weight fell into my arms in a way that startled me. The white plastic body bag that canvassed him seemed foreign and cold, whether from death or from the refrigeration I was unsure. With the military advisors upfront, I rode alone in the back with the body laying at my feet.

I took deep, controlled breaths, my eyes wide open, yearning to see if there was something else required of me. I needed to know the direction of the head as I could not determine it by feeling the outside of the bag. I quickly unzipped it. The boy’s salwar kameez was saturated with his frozen and thawed blood, holding a distinct grip having just come from the freezer, but I could not find his head. He had been decapitated.

I quickly rezipped the bag and just stared. There was no use investigating the issue. I placed my hand on the bag and uttered a Hindu prayer, the only prayer I could remember at the time. As we drove, I fell into a meditation about the boy, his story, his family, his joys, and his name, as most likely his life would be forgotten as “just another” innocent civilian lost in this war.

I imagined that as his mother scurried to find Nassim that day after hearing about the VBIED, she probably looked off into the distance with a solemn gaze, watching as a rat tail of black smoke polluted the clear blue sky and the imposing snow-capped mountains in the background, feeling a light breeze . . .

. . . like an unwanted soft touch on her cheek.

This story is written in memory of an unknown child who was murdered in a VBIED attack during Dr. Anant Shukla’s deployment in support of Operation Freedom Sentinel 2019 in Eastern Afghanistan. The name Nassim, in Arabic, means breeze.

The views expressed are those of the author and do not reflect the official policy of the Department of the Army, the Department of Defense or the U.S. Government.

CPT Anant Shukla, MD FS
(anant.shukla.mil@mail.mil)
Mendoza Aviation Clinic, William Beaumont Army Medical Center, Fort Bliss, TX, USA

Supervising Editor: Brian J. Zink, MD