Academic Emergency Medicine

A GLOBAL JOURNAL OF EMERGENCY CARE

WILEY

Vol.29 No. 10
October 2022
ISSN 1069-6563

AEM
Academic Emergency Medicine

Acem_v29_i10_14028.indd 1 9/20/2022 2:13:37 PM

Academic Emergency Medicine

CONTENTS

Volume 29 · Number 10 · October 2022 · www.aemj.org

Elevating the human condition during times of emergency

SYSTEMATIC REVIEWS (WITH OR WITHOUT META-ANALYSES)

CME **Comparative efficacy of sedation or analgesia methods for reduction of anterior shoulder dislocation: A systematic review and network meta-analysis**
Minoru Hayashi, Kenichi Kano, Naoto Kuroda et al. 1160

ORIGINAL CONTRIBUTIONS

Restrictive fluids versus standard care in adults with sepsis in the emergency department (REFACED): A multicenter, randomized feasibility trial
Marie K. Jessen, Lars W. Andersen, Marie-Louise H. Thomsen et al. 1172

Electrocardiographic findings associated with early clinical deterioration in acute pulmonary embolism
Anthony J. Weekes, Jaron D. Raper, Alyssa M. Thomas et al. 1185

Outcomes of the National Heart, Lung, and Blood Institute K12 program in emergency care research: 7-year follow-up
Cynthia D. Morris, Jennifer N. B. Cook, Amber Lin et al. 1197

Evaluation of the "hemoptysis" item in clinical decision rules for the diagnosis of pulmonary embolism in the emergency department
Héloïse Bannelier, Judith Gorlicki, Andrea Penaloza et al. 1205

Periosteal block versus intravenous regional anesthesia for reduction of distal radius fractures: A randomized controlled trial
Sierra Beck, Alana Brunner-Parker, Rosemary Stamm et al. 1213

Changes in scientific characteristics of abstracts accepted to the Society for Academic Emergency Medicine Annual Meeting, 1990–2020
Carlos A. Camargo Jr., Krislyn M. Boggs, Rebecca E. Cash et al. 1221

SYSTEMATIC REVIEWS (WITH OR WITHOUT META-ANALYSES)

Screening tools to identify patients with unmet palliative care needs in the emergency department: A systematic review
Scott W. Kirkland, Esther H. Yang, Miriam Garrido Clua et al. 1229

SAEM
Society for Academic Emergency Medicine

Acem_v29_i10_14028.indd 2 9/20/2022 2:13:37 PM

Academic Emergency Medicine

CONTENTS

Volume 29 · Number 10 · October 2022 · www.aemj.org

RESEARCH METHODS AND STATISTICS

Pragmatic clinical trial design in emergency medicine: Study considerations and design types
Cameron J. Gettel, Maame Yaa A. B. Yiadom, Steven L. Bernstein et al. 1247

RESEARCH LETTERS

Laboratory trends in severe MIS-C
Curt Stankovic, Amy DeLaroche, Rajan Arora et al. 1258

Novel quality assessment methodology in focused cardiac ultrasound
Stephanie DeMasi, Lindsay A. Taylor, Adam Weltler et al. 1261

SPECIAL CONTRIBUTION

Global emergency medicine: A scoping review of the literature from 2021
Sean M. Kivlehan, Braden J. Hexom, Joseph Bonney et al. 1264

HOT OFF THE PRESS

Hot off the press: Low-risk, recurrent abdominal pain in the emergency department
Justin Morgenstern, Kirsty Challen, Corey Heitz et al. 1275

COMMENTARY - UNSOLICITED

The bridge
Rebekah Mannix, Amanda M. Stewart, Joel Hudgins et al. 1278

CORRESPONDENCE - UNSOLICITED LETTERS TO THE EDITOR

How to treat venous aneurysm in emergency department?
Zhoupeng Wu, Yukui Ma 1280

REFLECTION

Out of Time
Fabrizio Elia, Fabrizio Vallelonga 1281

Just another day
Richard J. Cunningham 1283

SAEM
Society for Academic Emergency Medicine

Acem_v29_i10_14028.indd 3 9/20/2022 2:13:38 PM

Academic Emergency Medicine

Official Journal of the Society for Academic Emergency Medicine | **saem.org**

JEFFREY A. KLINE, MD
Editor-In-Chief
Wayne State University School of Medicine
Detroit, Michigan
jkline@wayne.edu

CHRISTOPHER R. CARPENTER, MD, MSC
Deputy Editor
Washington University
St. Louis, Missouri

JAMES R. MINER, MD
Senior Associate Editor for Electronic Publishing
Hennepin County Medical Center
Minneapolis, Minnesota

KEN MILNE, MD, MSC
Senior Associate Editor for Social Media
South Huron Hospital Association
Exeter, Ontario, Canada

DEBORAH B. DIERCKS, MD, MSC
Senior Associate Editor for Cardiovascular Medicine
University of Texas Southwestern Medical Center
Dallas, Texas

ZACHARY F. MEISEL, MD, MPH, MSHP
Senior Associate Editor for Health Communication
Perelman School of Medicine, University of Pennsylvania
Philadelphia, Pennsylvania

TAMMIE E. QUEST, MD
Senior Associate Editor for Palliative Care
Emory University School of Medicine
Atlanta, Georgia

ASSOCIATE EDITORS

Beau Abar, PhD
University of Rochester Medical Center
Rochester, New York

Elizabeth Alpern, MD, MSCE
Lurie Children's Hospital,
Northwestern University, Feinberg School of Medicine
Chicago, Illinois

Harrison J. Alter, MD, MS
Highland Hospital - Alameda Health System Oakland, California

Kamna Balhara, MD, MS
Johns Hopkins Medical Institution
Baltimore, Maryland

Brigitte M. Baumann, MD, MSCE
Cooper University Hospital
Camden, New Jersey

Steven B. Bird, MD
University of Massachusetts Medical School
Worcester, Massachusetts

Bernard P. Chang MD, PhD
Columbia University Irving Medical Center New York, New York

Makini Chisholm-Straker, MD
Icahn School of Medicine at Mount Sinai
New York, New York

Robert Cloutier, MD, MCR
Oregon Health & Science University
Portland, Oregon

Wendy C. Coates, MD
Harbor-UCLA Medical Center
Torrance, California

D. Mark Courtney, MD, MSCI
University of Texas Southwestern Medical Center
Dallas, Texas

Latha Ganti, MD, MS, MBA
University of Central Florida
Orlando, Florida

Richard T. Griffey, MD, MPH
Washington University
St. Louis, Missouri

Kennon Heard, MD, PhD
University of Colorad o School of Medicine
Aurora, Colorado

Brian C. Hiestand, MD, MPH
Wake Forest University Health Sciences
Winston-Salem, North Carolina

Ula Hwang, MD, MPH
Mount Sinai School of Medicine
New York, New York

Timothy B. Jang, MD
Harbor-UCLA Medical Center
Torrance, California

Alan E. Jones, MD
University of Mississippi Medical Center
Jackson, Mississippi

Peter C. Jenkins, MD
Indiana University School of Medicine
Indianapolis, Indiana

Damon Kuehl, MD
Virginia Tech Carilion School of Medicine
Roanoke, Virginia

Michelle Lin, MD, MPH, MS
Icahn School of Medicine at Mount Sinai
New York, New York

Michelle L. Macy, MD, MS
Lurie Children's Hospital,
Northwestern University
Chicago, Illinois

Alice Mitchell, MD
Indiana University School of Medicine
Indianapolis, Indiana

Paul Musey, Jr., MD
Indiana University
Indianapolis, Indiana

Peter D. Panagos, MD
Washington University
St. Louis, Missouri

Jesse Pines, MD, MBA, MSCE
US Acute Care Solutions (USACS)
Canton, Ohio

Michael A. Puskarich, MD, MSCR
Hennepin County Medical Center
Minneapolis, Minnesota

Robert F. Reardon, MD
Hennepin County Medical Center
Minneapolis, Minnesota

Kristin L. Rising, MD, MSHP
Thomas Jefferson University
Philadelphia, Pennsylvania

Michael S. Runyon, MD, MPH
Carolinas Medical Center
Charlotte, North Carolina

Elizabeth Schoenfeld, MD
University of Massachusetts School of Medicine, Baystate Health
Worcester, Massachusetts

Nathan I. Shapiro, MD, MPH
Harvard Medical School,
Beth Israel Deaconess Medical Center
Boston, Massachusetts

Stephen W. Smith, MD
Hennepin County Medical Center
Minneapolis, Minnesota

Michelle D. Stevenson, MD, MS
University of Louisville
Louisville, Kentucky

Jill Stoltzfus, PhD
St. Luke's University Health Network,
Temple/St. Luke's School of Medicine
Bethlehem, PA

Andrew Taylor, MD, MHS
Yale School of Medicine
New Haven, Connecticut

Daniel L. Theodoro, MD, MSCI
Washington University
St. Louis, Missouri

Jody Vogel, MD, MSC, MSW
Stanford University
Palo Alto, California

Arjun Venkatesh, MD, MBA
Yale University School of Medicine
New Haven, Connecticut

Kabir Yadav, MDCM, MS, MSHS
Harbor-UCLA Medical Center
Torrance, California

Lalena M. Yarris, MD, MCR
Oregon Health & Science University
Portland, Oregon

Shahriar Zehtabchi, MD
State University of New York
Brooklyn, New York

Mark R. Zonfrillo, MD, MSCE
Alpert Medical School of Brown University
Providence, Rhode Island

RESIDENT EDITORS

Ian Ferguson, MD
Washington University
St. Louis, Missouri

Madeline Schwid, MD
Brigham & Womens Hospital/Harvard Medical School
Boston, Massachusetts

DEPARTMENT EDITORS

Associate Social Media Editor
Michael Gottlieb, MD, RDMS
Rush University Medical Center
Chicago, Illinois

The Biros Section on Research Ethics
Shellie L. Asher, MD
Albany Medical College
Albany, New York

Director, Resident Member of the Editorial Board Program
Mark B. Mycyk, MD
Cook County Hospital
Chicago, Illinois

CME Editor
Corey R. Heitz, MD
Lewis Gale Medical Center
Salem, Virginia

Editor of Infographics
Kirsty Challen, BSc, MBChB, MRes, PhD
Lancashire Teaching Hospitals
United Kingdom

Media Reviews
Peter E. Sokolove, MD
University of California San Francisco, San Francisco, California

Reflections and Resident Portfolios
Brian Zink, MD
Brown University
Providence, Rhode Island

EDITORIAL OFFICES

Peer Review Coordinator

Taylor Bowen
tbowen@saem.org
Tel 434.218.3366

Journal Manager

Stacey Roseen
sroseen@saem.org
Tel. 847-813-9823

SAEM

1111 E. Touhy Ave., Suite 540
Des Plaines, IL 60018
Tel 847.813.9823
Fax 847.813.5450
saem@saem.org

Copyright and Copying (in any format)

Copyright © 2022 Society for Academic Emergency Medicine. All rights reserved. No part of this publication may be reproduced, stored or transmitted in any form or by any means without the prior permission in writing from the copyright holder. Authorization to copy items for internal and personal use is granted by the copyright holder for libraries and other users registered with their local Reproduction Rights Organization (RRO), e.g. Copyright Clearance Center (CCC), 222 Rosewood Drive, Danvers, MA 01923, USA, www.copyright.com, provided the appropriate fee is paid directly to the RRO. This consent does not extend to other kinds of copying such as copying for general distribution, for advertising or promotional purposes, for republication, for creating new collective works or for resale. Permissions for such reuse can be obtained using the RightsLink "Request Permissions" link on Wiley Online Library. Special requests should be addressed to permissions@wiley.com

Disclaimer

The Publisher, the Society for Academic Emergency Medicine (SAEM) and Editors cannot be held responsible for errors or any consequences arising from the use of information contained in this journal; the views and opinions expressed do not necessarily reflect those of the Publisher, the Society for Academic Emergency Medicine (SAEM) and Editors, neither does the publication of advertisements constitute any endorsement by the Publisher, the Society for Academic Emergency Medicine (SAEM) and Editors of the products advertised.

For submission instructions, subscription and all other information visit http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1553-2712

Received: 20 May 2022 | Revised: 13 July 2022 | Accepted: 20 July 2022

DOI: 10.1111/acem.14568

SYSTEMATIC REVIEWS (WITH OR WITHOUT META-ANALYSES)

CME

Comparative efficacy of sedation or analgesia methods for reduction of anterior shoulder dislocation: A systematic review and network meta-analysis

Minoru Hayashi MD[1] | Kenichi Kano MD[1] | Naoto Kuroda MD[2,3,4] | Norio Yamamoto MD[4,5,6] | Akihiro Shiroshita MD, MPH[4,7] | Yuki Kataoka MD, MPH, DrPH[4,8,9,10]

[1]Department of Emergency Medicine, Fukui Prefectural Hospital, Fukui City, Fukui, Japan

[2]Department of Pediatrics, Wayne State University, Detroit, Michigan, USA

[3]Department of Epileptology, Tohoku University Graduate School of Medicine, Sendai, Japan

[4]Scientific Research Works Peer Support Group (SRWS-PSG), Osaka, Japan

[5]Department of Orthopedic Surgery, Miyamoto Orthopedic Hospital, Okayama, Okayama, Japan

[6]Department of Epidemiology, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Okayama, Japan

[7]Department of Respiratory Medicine, Ichinomiyanishi Hospital, Ichinomiya, Aichi, Japan

[8]Department of Internal Medicine, Kyoto Min-Iren Asukai Hospital, Kyoto, Japan

[9]Section of Clinical Epidemiology, Department of Community Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan

[10]Department of Healthcare Epidemiology, Kyoto University Graduate School of Medicine/Public Health, Kyoto, Japan

Correspondence
Norio Yamamoto, Department of Orthopedic Surgery, Miyamoto Orthopedic Hospital, 4-2-63, Kunitomi, Naka-ku, Okayama, Okayama 773-8236, Japan.
Email: lovescaffe@yahoo.co.jp

Abstract

Background: We performed a network meta-analysis (NMA) to compare the efficacy and safety of intravenous sedation (IVS), intraarticular anesthetic injection (IAA), and peripheral nerve block (PNB) as sedation or analgesia methods for the reduction of anterior shoulder dislocation.

Methods: We included randomized controlled trials (RCTs) comparing different sedation or analgesia methods for anterior shoulder dislocation reduction. The Cochrane Central Register of Controlled Trials, MEDLINE, EMBASE, ICTRP, ClinicalTrials.gov, and Google Scholar databases were searched in October 2021. We conducted a random-effects NMA within a frequentist framework. We evaluated the confidence in each outcome using the CINeMA tool.

Results: Sixteen RCTs (957 patients) were included. Regarding the primary outcomes, the three methods might result in little to no difference in the immediate success rate of reduction and patient satisfaction. The IAA method had a shorter emergency department length of stay than that of the IVS method (mean difference [MD] −107.88 min, 95% confidence interval [CI] −202.58 to −13.18). In the secondary outcomes, the IAA method had a lower pain score than that of the PNB method (standardized MD −1.83, 95% CI −3.64 to −0.02). The IAA and PNB methods might require a longer time for reduction than that of the IVS method (MD 5.3 min, 95% CI 2.4 to 10.36; MD 15.25, 95% CI 5.49 to 25.01). The three methods might result in little to no difference in the number of reduction attempts and total success rate of reduction. However, the confidence ratings for all treatment comparisons were very low. IAA and PNB had no adverse respiratory events.

Conclusions: The results of our NMA indicated that three sedation or analgesia methods (IVS, IAA, and PNB) might result in little to no difference in the success rate of reduction and patient satisfaction. IAA and PNB had no adverse respiratory events.

KEYWORDS
analgesia, conscious sedation, intraarticular injection, nerve block, shoulder dislocation

Supervising Editor: Dr. Paul Musey

 © 2022 Society for Academic Emergency Medicine.

INTRODUCTION

Shoulder dislocation accounts for 50% of all dislocations and is the most common type of dislocation.[1] Most shoulder dislocations are anterior (90%–98%).[1] Pain due to dislocation is severe, and the administration of sedatives and analgesics helps to relieve pain. However, pain also triggers muscle spasms around the shoulder. Therefore, for a successful shoulder dislocation reduction, it is important to relax the muscles surrounding the shoulder.[2]

Intravenous sedation (IVS) and intraarticular anesthetic injection (IAA) have been widely used as sedation or analgesia methods to reduce shoulder dislocation.[2–4] As for the comparative efficacy and safety between IVS and IAA, previous systematic reviews reported no significant differences in the success rate of reduction and patient satisfaction. These systematic reviews also revealed that IAA achieved fewer adverse events, shorter lengths of hospital stay, and lower medical costs than that of IVS.[2–4]

Recently, ultrasound-guided peripheral nerve blocks (PNB) have been used to reduce shoulder dislocations. Several randomized controlled trial (RCT) studies have compared the efficacy and safety of PNB and IVS, but the results have been controversial.[5–7] In addition, no systematic reviews have compared the efficacy and safety of these sedation or analgesia methods.

The best method of sedation and analgesia for the reduction of shoulder dislocation remains uncertain. Therefore, we conducted a systematic review and network meta-analysis (NMA) of RCTs to compare the efficacy and safety of IVS, IAA, and PNB for the reduction of anterior shoulder dislocation.

METHODS

Protocol and registration

We reported this systematic review and NMA of RCTs in accordance with the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) extension statement for reporting systematic reviews incorporating network meta-analyses of health care interventions, including the checklist and explanations[8] (Table S1). In addition, we registered the protocol before starting this review.[9]

Inclusion criteria of the articles for the review

Type of studies

We included RCTs that assessed sedation or analgesia methods for the reduction of anterior shoulder dislocation and excluded crossover trials, quasi-experimental studies, and quasi-randomized trials. In addition, we included all reports, including published and unpublished articles, conference abstracts, and letters.

Study participants

We included participants older than 15 years and had a diagnosis of anterior shoulder dislocation based on physical examination or radiography of the shoulder. In addition, we included the intervention and comparator as IVS, IAA, PNB, placebo, or no sedation or analgesia (no drug). We defined IVS as intravenous injection of sedatives. Any types and dosages of sedatives and analgesics were acceptable. We defined IAA as the injection of local anesthetics into the glenohumeral joint. All medications and doses were acceptable. We defined PNB as the injection of local anesthetics into the brachial plexus in the interscalene or suprascapular nerves. All medications and doses were acceptable. The exclusion criteria were patients who could not obtain informed consent, allergies to any study medications, multiple traumas, associated fractures of the humerus (except Hill-Sachs and Bankart lesions), hemodynamic instability, or respiratory distress.

Outcomes of interest

The primary outcomes were the immediate success rate of the reduction, patient satisfaction, and emergency department (ED) length of stay (min). We chose three outcomes because the ideal shoulder dislocation reduction should be fast and effective and have high patient satisfaction.[10] Immediate success and ED length of stay (min) were defined by the authors. In addition, we assessed patient satisfaction with the shoulder reduction procedure. The secondary outcomes were adverse events, pain score, the time required for reduction (min), number of reduction attempts, and total success rate of the reduction. Secondary outcomes were defined by the authors of the study.

Search strategy

We searched the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE via PubMed, EMBASE via Embase.com on October 31, 2021, and Google Scholar on November 12, 2021. We also searched the World Health Organization International Clinical Trials Platform Search Portal (ICTRP) and ClinicalTrials.gov for unpublished RCTs and ongoing RCTs on October 31, 2021 (Table S2).

Study selection and data extraction

Two independent reviewers (MH and KK) assessed all the identified publications for eligibility. If the study was abstract only and it was not clear whether it met the criteria for review, we contacted the original author. Two reviewers discussed and resolved any disagreements.

Data items

We extracted the following study characteristics:

1. Methods: study design, setting, and the number and country of study centers.
2. Participants: number, sex, age, and inclusion/exclusion criteria.
3. Interventions: sedative or analgesia methods and reduction techniques.
4. Outcomes: primary and secondary outcomes and the time points reported.

Geometry of the network

We demonstrated a network geometry that presented the nodes as direct comparisons as lines connecting these nodes. Nodes are used for sedation or analgesia. The numbers above the line represent the number of RCTs in direct comparisons.

Risk-of-bias assessment

Two reviewers (MH and KK) independently assessed the study level risk of bias using the risk-of-bias (ROB) 2 tool.[11] If necessary, the two reviewers discussed and resolved any disagreements with the third reviewer (NK).

Data synthesis and statistical analysis

Data synthesis

We pooled the relative risk ratios (RRs) for binary outcomes and the mean differences (MDs) or standardized mean differences (SMDs) for continuous outcomes. First, we conducted a pairwise meta-analysis for each treatment comparison using a random-effects model. Forest plots were used to visualize the effect of each treatment and assess heterogeneity. Second, we conducted a random-effects NMA within a frequentist framework. Because there were no direct comparisons between IAA and PNB, we only estimated the treatment effect comparison between IAA and PNB using NMA. In addition, we described a network plot to summarize the treatment effect and the study size of each comparison and league tables to summarize the results of pairwise meta-analyses and NMA. Third, we used the results from the intention-to-treat analysis and did not impute missing data based on the recommendation of the Cochrane Handbook.[12] Finally, we summarized adverse events based on the definition in the original article. However, we did not perform a meta-analysis because of nonstandardized definitions, inadequate monitoring, or possibly incomplete reporting.

Sensitivity analysis

We conducted the following sensitivity analysis for the primary outcomes.

1. Sensitivity analysis included only studies with a low overall assessment of ROB.
2. Sensitivity analysis included only studies that met the ideal outcome: at first attempt or first reduction technique for the immediate success rate of the reduction, most satisfied patient for patient satisfaction, and time from the beginning of the procedure to hospital discharge for the ED length of stay.

Statistical analysis

We used Review Manager software (RevMan V.5.4.1, The Nordic Cochrane Center, The Cochrane Collaboration) for pairwise meta-analyses and MetaInsight for frequentist NMA.[13] In addition, we used R software version 4.0.2 (R Foundation for Statistical Computing) and the packages "meta" (version 5. 2-0) and "pwr" (version 1. 3-0) for sample size calculations in future trials.

Assessment of the confidence for each outcome

Two reviewers (MH and NK) evaluated the confidence for each outcome using the CINeMA tool.[14,15] The CINeMA framework includes the following domains: within-study bias, across-studies bias, indirectness, imprecision, heterogeneity, and incoherence. For the domains of within-study bias and indirectness, we assessed the contribution of each study in each network estimate and combined these contributions with the study-specific ratings (very low, low, moderate, or high) to assess the relative effect for each comparison in the network. For the domains of imprecision, heterogeneity, and incoherence, we assessed major concerns, some concerns, and no concerns about how far the 95% confidence interval (CI) extend on both sides of the no effect line equal to the point estimate between the two interventions.

Sample size calculation for future trials

As there was no direct comparison between IAA and PNB, we performed a sample size calculation for a future trial targeting patients with anterior shoulder dislocation. We simulated a 1:1 RCT and set the following parameters: alpha of 0.05 and power of 0.8. We calculated the pooled immediate success rate of IAA using a random-effects model and then calculated the PNB using the RR from the NMA. Based on the estimated immediate success rates of IAA and PNB, we calculated the sample size.

RESULTS

Results of the search

We identified 1855 records during the search conducted in October 2021. Of these reports, 29 were considered for inclusion after reviewing their titles and abstracts. After a full-text review, 16 studies (957 patients) were included (Figure 1). Table S3 shows the list of studies excluded from this review and the reasons for exclusion. Table 1 presents the characteristics of the included studies. While lidocaine was administered using the IAA and PNB methods in most studies, various drugs were administered in the IVS group. In addition, the reduction techniques varied among the included studies. Eleven studies compared the efficacy and safety between IAA and IVS.[16–26] Four studies compared the efficacy and safety between PNB and IVS.[5–7,28] One study compared the efficacy and safety of IAA and no drugs.[27] To our knowledge, no study has compared the efficacy and safety of IAA and PNB. The ROB assessment is summarized in Figure S1. For the primary outcomes, six of the 10 included studies were assessed as having low ROB and the other four were assessed as having high ROB in the immediate success rate. However, patient satisfaction and ED length of stay were assessed as all high ROB. For the secondary outcomes, many outcomes were assessed as high or of some concerns for ROB. Most of the elevation in ROB was due to the unclear allocation sequence generation in Domain 1 and the lack of blinding of the outcome assessor, which might have influenced the outcomes in Domain 4.

Primary outcomes

Immediate success rate of the reduction

We included 10 studies[7,18–24,26,27] (Figure 2). The pairwise analysis is shown in Figure S3a. The IAA method might result in little to no difference in immediate success rate compared with the IVS method (eight studies, 408 patients; RR 0.93, 95% CI 0.84 to 1.02).[18–24,26] The evidence was very uncertain regarding the effect of the PNB method on the immediate success rate compared with the IVS method (one study, 41 patients; RR 1.13, 95% CI 0.84 to 1.52[7]; Figure 2). Other comparisons are shown in Figure 2 and Table S4a. The confidence ratings for all treatment comparisons are presented in Table S5a.

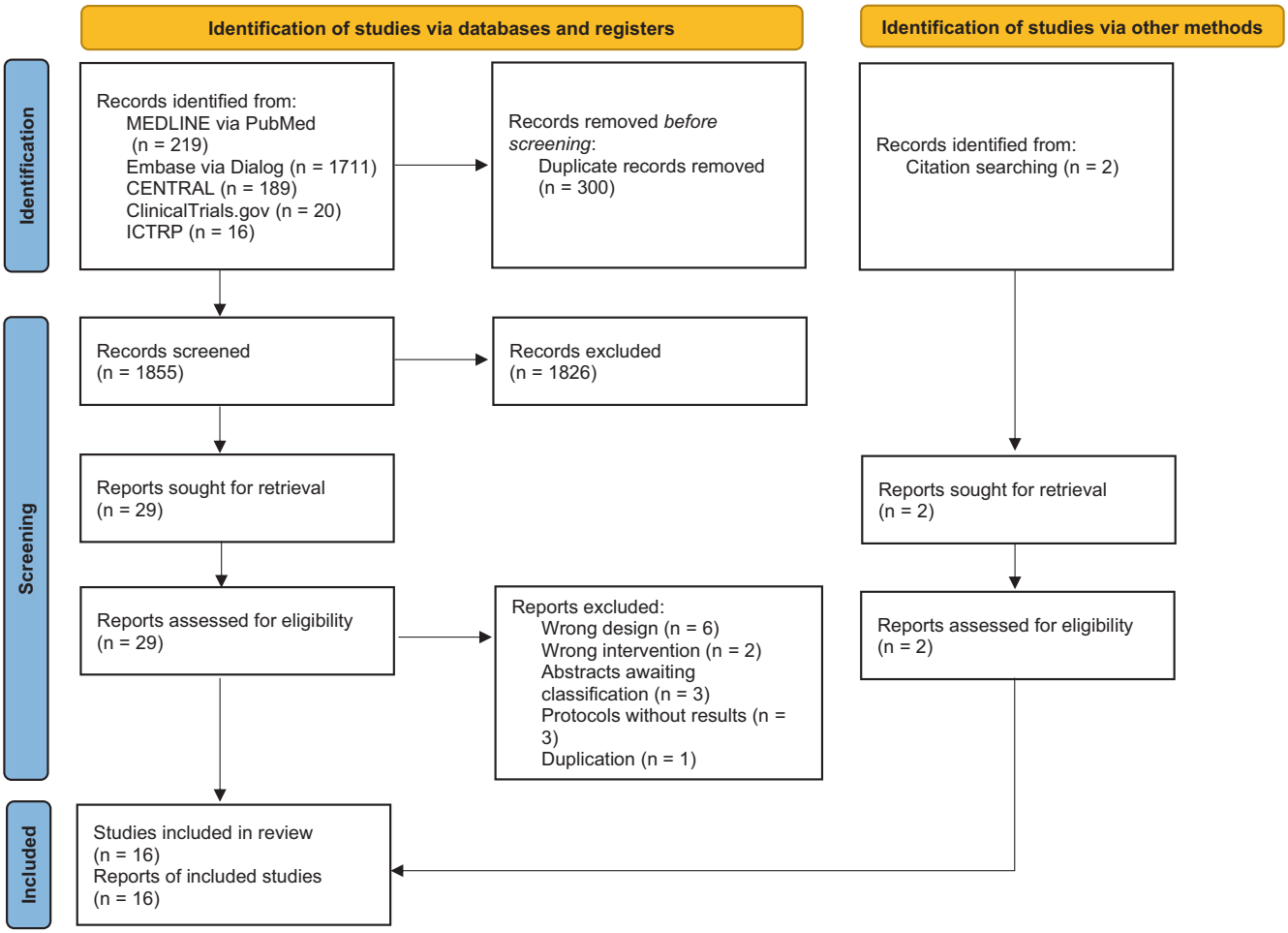


FIGURE 1 PRISMA flow diagram of the literature search results. We identified 1855 records during the search conducted in October 2021. Of these reports, 29 were considered for inclusion after review of titles and abstracts. After full-text review, 16 studies were included.

TABLE 1 Characteristics of included studies

Author, year, country	Intervention (I)	Control (C)	Patient number	Mean age (years)	Reduction technique
			I/C	I/C	
	IAA	IVS			
Suder, 1994, Denmark[16]	1% lidocaine 20 mL	Diazepam or pethidine discretion of the treating physician	33/35	48 (15–79)[a]	Kocher or Hippocratic method
Suder, 1995, Denmark[17]	1% lidocaine 20 mL	Diazepam or pethidine discretion of the treating physician	26/26	47 (18–89)[a]	Kocher or Hippocratic method
Matthews, 1995, USA[19]	1% lidocaine 20 mL	Midazolam 2 mg and morphine sulfate 10 mg	15/15	38/35.4	Traction–countertraction or Milch Or Scapular manipulation method
Kosnik, 1999, USA[20]	4 mg/kg (max 200 mg) of 1% lidocaine	Diazepam 5 mg and morphine sulfate 10 mg	29/20	41/41	Traction–countertraction method
Miller, 2002, USA[21]	1% lidocaine 20 mL	Midazolam 2 mg and fentanyl 100 μg	16/14	33/35	Modified Stimson or Scapular manipulation method
Orlinsky, 2002, USA[22]	1% lidocaine 20 mL	Diazepam 5–10 mg and meperidine 1–2 mg/kg	29/25	36/39	External rotation or traction–countertraction method
Moharari, 2008, Iran[23]	1% lidocaine 20 mL	Diazepam 5 mg and meperidine 25 mg	24/24	31.7/38.9	Traction–countertraction method
Cheok, 2011, Malaysia[24]	1% lidocaine 20 mL	Diazepam 0.1 mg/kg and meperidine 1 mg/kg	32/31	35.2/29.7	Traction–countertraction method
Hames, 2011, Canada[25]	4 mg/kg (max 200 mg) of 1% lidocaine	Discretion of the treating physician	25/19	43/27	Not stated
Ahmed, 2011, Iraq[26]	1% lidocaine 20 mL	Midazolam 2 mg and fentanyl 100 μg	15/15	32/35[b]	Modified Stimson or scapular manipulation method
Kashani, 2016, Iran[18]	1% lidocaine 20 mL	Midazolam 0.05 mg/kg and fentanyl 1 μg/kg	52/52	28.18/29.9	Leidelmeyer method
	IAA	No drug (no sedation or analgesia)			
Tamaoki, 2012, Brazil[27]	1% lidocaine 20 mL	No drug	22/20	39.1/32.6	Traction–countertraction method
	PNB (Interscalene nerve block)	IVS			
Blaivas, 2011, USA[5]	Lidocaine with epinephrine 20–30 mL	Etomidate. Dose was not stated	21/21	39/35.9	Not stated
Raeyat Doost, 2017, Iran[6]	1% lidocaine with epinephrine 15–25 mL	Propofol 1 mg/kg and fentanyl 2 μg/kg	30/30	28.5/28.9	Leidelmeyer or Milch or traction–countertraction method
	PNB (suprascapular nerve block)	IVS			
Tezel, 2014, Turkey[7]	2% prilocaine 5 mL	Ketamine 1–2 mg/kg	21/20	24/23.5	Modified Kocher method
Abbasi, 2020, Iran[28]	1% lidocaine 5–10 mL	Discretion of the treating physician	100/100	46.55 (20–91)[a]	Milch or traction–countertraction method

Abbreviations: IAA, intraarticular anesthetic injection; IVS, intravenous sedation; PNB, peripheral nerve block.

[a]Mean age, years (range).

[b]Median age.

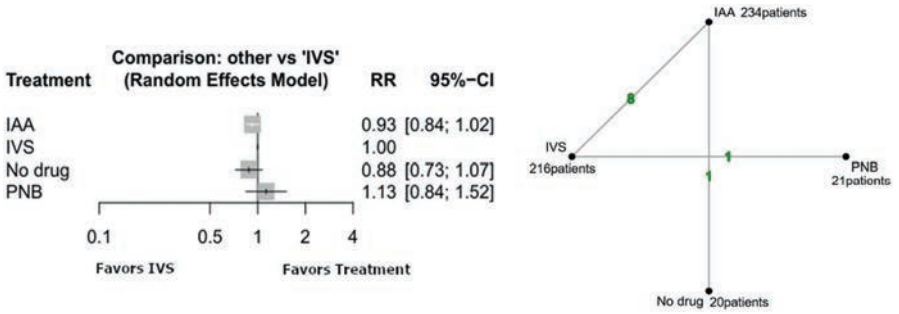


FIGURE 2 Forest plot and network plot for all interventions compared with IVS in the immediate success rate of reduction. The three methods might result in little to no difference in the success rate of reduction. CIs are slightly different in Figure 2 and S3a. This is because of the different statistical analysis software used. IAA, intraarticular anesthetic injection; IVS, intravenous sedation; no drug, no sedation or analgesia; PNB, peripheral nerve block; RR, risk ratio.

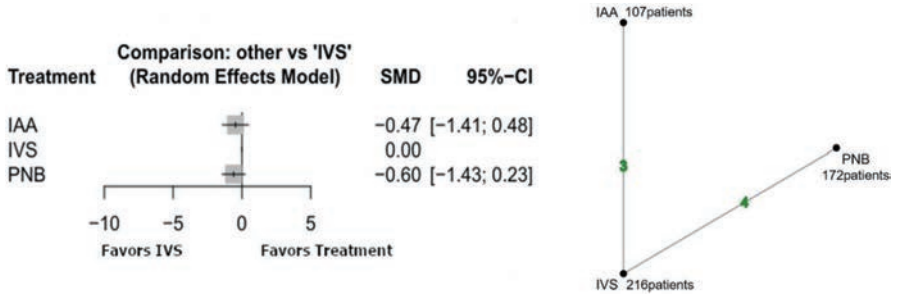


FIGURE 3 Forest plot and network plot for all interventions compared with IVS in patient satisfaction. The three methods might result in little to no difference in patient satisfaction. CIs are slightly different in Figure 3 and S3b. This is because of the different statistical analysis software used. IAA, intraarticular anesthetic injection; IVS, intravenous sedation; PNB, peripheral nerve block; SMD, standardized mean difference.

Patient satisfaction

We included seven studies[5–7,16,18,24,28] (Figure 3). The units of measurement for patient satisfaction were differed across studies; therefore, we converted the patient satisfaction reported in the survey scale to scores. Two studies reported patient satisfaction as yes or no; thus, we defined yes as 2 and no as 1.[16,24] In three studies, we changed the scores from 1 to 5 in the order of high patient satisfaction reported on a 5-level scale.[7,18,28] For example, quite satisfied 5 and complete dissatisfaction 1.[18] We pooled these scores using SMD according to the Cochrane Handbook.[12] In a study by Suder et al.,[17] the table of patient satisfaction may be incorrect. The total number of satisfied and unsatisfied patients in both groups differed from the number of included patients, but we could not contact the authors. Therefore, we excluded this study from the meta-analysis. The pairwise analysis is shown in Figure S3b. The evidence was very uncertain regarding the effect of the IAA method on patient satisfaction compared with the IVS method (three studies, 222 patients; SMD −0.47, 95% CI −1.41 to 0.48).[16,18,24] The evidence was uncertain regarding the effect of the PNB method on patient satisfaction compared with the IVS method (four studies, 273 patients; SMD −0.60, 95% CI−1.43 to 0.23)[5–7,28]; Figure 3. Other comparisons are presented in Table S4b. The confidence ratings for all treatment comparisons are presented in Table S5b.

ED length of stay (min)

Eleven studies were included[5–7,19,21–26,28] (Figure 4). Two studies assessed the time from entry into a room in the ED to discharge.[5,19] Two studies assessed the time from initial physician assessment to discharge.[24,25] Two studies assessed the time from the start of reduction to discharge.[7,21] Two studies assessed the time from the start of sedation or analgesia methods to discharge.[6,22] One study assessed the duration from reduction to discharge.[26] We could not determine the definition of the ED length of stay in two studies.[23,28] The pairwise analysis is shown in Figure S3c. The IAA method might have a shorter ED length of stay than that of the IVS method; however, the evidence was very uncertain (seven studies, 299 patients; MD −107.88 min, 95% CI −202.58 to −13.18).[19,21–26] The evidence was very uncertain about the effect of the PNB method on the ED length of stay compared with the IVS method (four studies, 343 patients; MD −26.23 min, 95% CI, −149.02 to 96.57)[5–7,28]; Figure 4. Other comparisons are presented

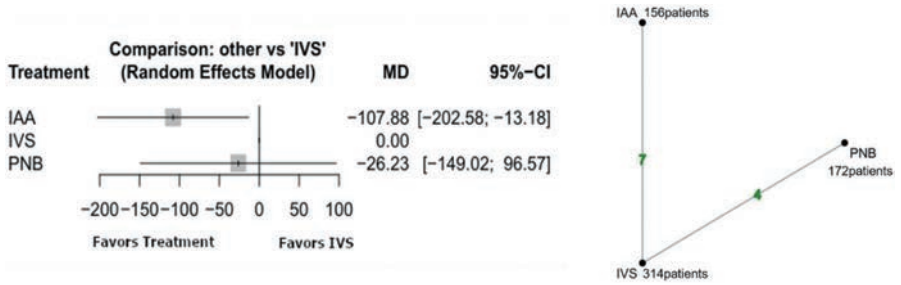


FIGURE 4 Forest plot and network plot for all interventions compared with IVS on the ED length of stay (min). The IAA method had a shorter ED length of stay than that of the IVS method (MD −107.88 min, 95% CI −202.58 to −13.18). IAA, intraarticular anesthetic injection; IVS, intravenous sedation; MD, mean difference; PNB, peripheral nerve block.

in Table S4c. The confidence ratings for all treatment comparisons are listed in Table S5c.

Secondary outcomes

All adverse events

Sixteen studies (957 patients) assessed the adverse events (Table 2). Two studies reported no adverse events.[25,27] Respiratory adverse events (hypoxia, hypoventilation, apnea, and respiratory depression) were common in the IVS group.[5–7,16–24,26,28] Nausea,[18,19,22,23,28] vomiting,[7,20,24] hypotension,[5,6,23] and headache[18,23] were also reported in the IVS group. Psychological agitation and drowsiness were reported in the IAA group,[22,23] and mild local anesthetic systemic toxicity was reported in the PNB group.[6]

Pain score

Eleven studies were included[5,6,17–21,23,24,26,27] (Figure S2a). Two studies assessed the pain scores after sedation or analgesia methods before reduction.[21,23] Two studies assessed the pain scores after reduction.[5,27] Six studies assessed the pain scores during reduction.[6,18–20,24,26] In the study by Ahmed et al.,[26] the standard deviation (SD) of the pain score was not reported. Hence, we adopted the SD as a substitute from the study by Miller et al.[21] because they used the same pain scale of 1 to 10, based on the Cochrane Handbook.[12] The pairwise analysis is shown in Figure S3d. The evidence was uncertain about the effect of the IAA method on the pain score compared with the IVS method (eight studies, 406 patients; SMD −0.56, 95% CI −1.36 to 0.25).[17–21,23,24,26] The evidence was very uncertain about the effect of the PNB method on the pain score compared with the IVS method. The IAA method might have a lower pain score than that of the PNB method; however, the evidence was very uncertain (no study, indirect comparison; SMD −1.83, 95% CI −3.64 to −0.02; Table S4d). The confidence ratings for all treatment comparisons are presented in Table S5d.

Time required for reduction (min)

We included six studies[17,21,23,26–28] (Figure S2b). Two studies assessed the period from the start of sedation or analgesia methods to the end of reduction.[17,28] Four studies assessed the period from the start to the end of reduction.[21,23,26,27] The pairwise analysis is shown in Figure S3e. The IAA method might require a longer time for reduction than that of the IVS method; however, the evidence was very uncertain (four studies, 160 patients; MD 5.3 min, 95% CI 0.24 to 10.36).[17,21,23,26] The PNB method might require a longer time for reduction than that of the IVS method; however, the evidence was very uncertain (one study, 200 patients; MD 15.25 min, 95% CI 5.49 to 25.01[28]; Figure S2b). Other comparisons are presented in Table S4e. The confidence ratings for all treatment comparisons are presented in Table S5e.

Number of reduction attempts

We included five studies[6,7,18,22,23] (Figure S2c). The pairwise analysis is shown in Figure S3f. The IAA method may result in little to no difference in the number of reduction attempts compared with the IVS method (3 studies, 172 patients) (MD: 0.18; 95% CI: −0.09 to 0.45).[18,22,23] The evidence was very uncertain regarding the effect of the PNB method on the number of reduction attempts compared with the IVS method (two studies, 101 patients; MD 0.07, 95% CI −0.23 to 0.36; Figure S2c).[6,7] Other comparisons are presented in Table S4f. The confidence ratings for all treatment comparisons are listed in Table S5f.

Total success rate of the reduction

Fifteen studies were included[5–7,16–27] (Figure S2d). The pairwise analysis is shown in Figure S3g. The IAA method may result in little to no difference in the total success rate compared with the IVS method (11 studies, 572 patients; RR 0.99, 95% CI 0.94 to 1.04).[16–26] The evidence was very uncertain about the effect of the PNB method on the total success rate compared with the IVS method (three studies, 143 patients; RR 1.00, 95% CI 0.92

TABLE 2 All adverse events

	Intervention (I)	Control (C)	Patient number	Adverse events (number)	
			I/C	I	C
Author, year	IAA	IVS			
Suder, 1994, Denmark[16]	1% lidocaine 20 mL	Diazepam or pethidine Discretion of the treating physician	33/35	0	10 respiratory depression
Suder, 1995, Denmark[17]	1% lidocaine 20 mL	Diazepam or pethidine Discretion of the treating physician	26/26	0	3 respiratory depression
Matthews, 1995, USA[19]	1% lidocaine 20 mL	Midazolam 2 mg and morphine sulfate 10 mg	15/15	0	2 required reversal agents, 3 nausea
Kosnik, 1999, USA[20]	4 mg/kg (max 200 mg) of 1% lidocaine	Diazepam 5 mg and morphine sulfate 10 mg	29/20	0	1 vomiting
Miller, 2002, USA[21]	1% lidocaine 20 mL	Midazolam 2 mg and fentanyl 100 μg	16/14	0	3 respiratory depression
Orlinsky, 2002, USA[22]	1% lidocaine 20 mL	Diazepam 5–10 mg and meperidine 1–2 mg/kg	29/25	1 psychological agitation	1 required reversal agents and nausea
Moharari, 2008, Iran[23]	1% lidocaine 20 mL	Diazepam 5 mg and meperidine 25 mg	24/24	3 drowsiness	14 (5 respiratory depression, 5 hypotension, 5 drowsiness, 2 headache, 2 nausea, 1 localized paresthesia)
Cheok, 2011, Malaysia[24]	1% lidocaine 20 mL	Diazepam 0.1 mg/kg and meperidine 1 mg/kg	32/31	0	9 (13 respiratory depression, 6 vomiting, 6 allergy, 20 thrombophlebitis)
Hames, 2011, Canada[25]	4 mg/kg (max 200 mg) of 1% lidocaine	Discretion of the treating physician	25/19	0	0
Ahmed, 2011, Iraq[26]	1% lidocaine 20 mL	Midazolam 2 mg and fentanyl 100 μg	15/15	0	1 serious cardiorespiratory complication
Kashani, 2016, Iran[18]	1% lidocaine 20 mL	Midazolam 0.05 mg/kg and fentanyl 1 μg/kg	52/52	0	11 (6 apnea, 5 hypoxia, 2 nausea, 2 headache)
	IAA	No drug (no sedation or analgesia)			
Tamaoki, 2012, Brazil[27]	1% lidocaine 20 mL	No drug	22/20	0	0
	PNB (interscalene nerve block)	IVS			
Blaivas, 2011, USA[5]	Lidocaine with epinephrine 20–30 mL	Etomidate. Dose was not stated	21/21	0	2 O_2 saturation < 95%, 2 hypotension
Raeyat Doost, 2017, Iran[6]	1% lidocaine with epinephrine 15–25 mL	Propofol 1 mg/kg and fentanyl 2 μg/kg	30/30	1 systemic toxicity	3 hypoventilation, 1 hypotension
	PNB (suprascapular nerve block)	IVS			
Tezel, 2014, Turkey[7]	2% prilocaine 5 mL	Ketamine 1–2 mg/kg	21/20	0	2 hypoxia, 3 vomit, 3 agitation
Abbasi, 2020, Iran[28]	1% lidocaine 5–10 mL	Discretion of the treating physician	100/100	1 blood pressure fluctuation, 1 heartbeat fluctuation, 4 hematoma	21 respiratory distress, 15 apnea, 4 blood pressure fluctuation, 4 heartbeat fluctuation, 1 nausea

Abbreviations: IAA, intraarticular anesthetic injection; IVS, intravenous sedation; PNB, peripheral nerve block.

to 1.08[5–7]; Figure S2d). Other comparisons are presented in Table S4g. The confidence ratings for all treatment comparisons are presented in Table S5g.

Additional analysis

We performed the NMA for the immediate success rate of the reduction by restricting the studies of the overall assessment of ROB to low,[18,20,22–24,27] one technique,[7,18,21–24,26,27] or first attempt.[7,18,22,23] The IAA method might result in a slightly reduced immediate success rate compared with the IVS method in the low ROB studies (RR 0.85, 95% CI 0.77 to 0.94; Figures S2e and S3h, Tables S4h and S5h).

The NMA for patient satisfaction by restricting the patients with the highest satisfaction scores in the studies[7,18,24,28] was similar to the original analysis (Figures S2f and S3i, Tables S4i and S5i). Two studies restricted the patients from the beginning of the procedure to ED discharge compared with the PNB and IVS groups.[6,7] The results of the comparisons were similar to those of the original analysis (Figures S2g and S3j, Tables S4j and S5j).

Sample size calculation for future trials

The random-effects model showed that the point estimate of the immediate success rate of IAA was 0.78. Based on the risk ratio (IAA vs. PNB) of 0.82 NMA, we estimated that the immediate success rate of PNB was 0.95. Our sample size calculation revealed that a total of 118 patients will be needed in a future RCT.

Difference between protocol and review

We could not perform subgroup analysis because none of the studies reported anterior shoulder dislocations separately for the first time or recurrent anterior shoulder dislocations. In addition, we could not perform a sensitivity analysis for double-blind studies and for exclusion of studies using imputed statistics because none of the studies were applicable. We did not present a summary of the findings regarding the primary outcomes.

DISCUSSION

This frequentist NMA included 16 RCTs that compared the efficacy and safety of IVS, IAA, and PNB as sedation or analgesia methods for reduction of anterior shoulder dislocation.[5–7,16–28] The three methods might result in little to no difference in the immediate success rate of reduction and patient satisfaction. The IAA method might result in a shorter ED length of stay than that of the IVS method; however, the evidence is very uncertain. The IAA method may have a lower pain score than that of the PNB method; however, the evidence is very uncertain. The IAA and PNB methods may require a longer time for reduction than that of the IVS method; however, the evidence is very uncertain. The three methods might result in little to no difference in the number of reduction attempts and the total success rate of the reduction. The confidence for each relative treatment effect in NMA was low or very low, and the results were uncertain.

Our results were in line with the previous systematic reviews comparing the IAA method with the IVS method.[2,4] To our knowledge, no systematic review has compared the PNB method with other methods. The three methods might result in little to no difference in the immediate success rate of reduction. As for the number of reduction attempts related to the immediate success rate, our results might also result in little or no difference among the three methods. The three methods might also result in little to no difference in the total success rate of the reduction. Previous studies have reported that the reduction techniques and experience of the operator are relevant in the success rate,[29] and it might be more relevant than the sedation or analgesic methods in the success rate and number of reduction attempts.

Lack of difference in patient satisfaction among the three sedation or analgesia methods might be interpreted as patient satisfaction being more affected by waiting time, the attitude of medical staff, and appropriate explanations than by sedation or analgesia methods.[30]

The IAA method might have a shorter ED length of stay than that of the IVS method (MD −107.88 min, 95% CI −202.58 to −13.18). Evidence for the effect of the PNB method on the ED length of stay was compared with the IVS method. One of the four included studies compared the PNB method with the IVS method. Abbasi et al.[28] reported that the PNB method succeeded in only 34% of the patients and administered sedatives because inadequate pain relief accounted for 66% of all cases using the PNB method. Thus, the PNB method had a longer ED stay than that of the IVS method.[28] Contrary to the results of the study by Abbasi et al., the PNB method had a shorter ED length of stay than that of the IVS method in three of the four included studies.[5–7] Therefore, the PNB method has the potential to achieve a shorter ED length of stay than that of the IVS method under adequate pain relief. The ROB was high in all included RCTs because there was a subjective decision of discharge by medical staff or there was no description of the discharge decision in the studies.[5–7,28] In the future, RCTs with a low ROB using appropriate discharge criteria are required. The results of this study suggest that IAA is a better method in terms of shorter ED length of stay than the IVS.

The IAA method may have a lower pain score than that of the PNB method. However, comparisons between the IAA and PNB methods were only indirect, and the confidence rating was very low. Hence, IAA may be a better method for those unfamiliar with PNB because operators need training for the success of PNB.[6] The evidence was very uncertain about the effect of the IAA method on the pain score compared with the IVS method. All RCTs included in our study were conducted using landmark-guided IAA, and pain scores varied from

0.29 (SD 0.67) to 7.1 (SD 2.6) in the IAA method.[18,26] This might be because IAA may not have been administered appropriately. A previous study reported that landmark-guided IAA administered local anesthetics to inappropriate points in 41.1% of cases, as assessed using ultrasound.[31] Moreover, ultrasound-guided IAA can be used to administer local anesthetics to the appropriate point.[31] Therefore, under conditions of adequate pain relief, the IAA method has the potential to achieve a lower pain score than that of the IVS method.

The IAA and PNB methods may require a longer time for reduction than that of the IVS method (MD 5.3 min, 95% CI 2.4 to 10.36; and MD 15.25 min, 95% CI 5.49 to 25.01). The IAA method might require little to no difference in the time for reduction compared with the PNB method, but the IAA method tended to require fewer minutes than that of the PNB method (MD −9.95 min, 95% CI −1.04 to 20.94). The IVS may have the advantage of being sedated, which allows operators to perform the procedure smoothly. In contrast, patients administered with IAA or PNB were awake during the shoulder reduction technique. Awake states may prolong the time because of hesitation or interruption of the procedure due to complaints of pain. A longer reduction time may decrease patient satisfaction; however, there was no difference among the three methods in this NMA.

IAA and PNB are safer than IVS in terms of fewer respiratory adverse events. Respiratory depression was defined as $SpO_2 < 92\%$ or end-tidal $CO_2 > 40$ in one study[23] and as a respiratory rate < 12 breaths/min in another study[24]; however, respiratory depression definition was unclear in four studies.[16,17,21,28] Therefore, respiratory depression assessment in the IVS group may be incorrect due to the assessor's subjective assessment and variation. In previous systematic reviews, the number of adverse events was less common in the IAA group than those in the IVS group (RR 0.15, 95% CI 0.07 to 0.32[2]; OR 0.16, 95% CI 0.06 to 0.43).[4] However, we did not perform a meta-analysis of all adverse events because the definitions of adverse events differed between the included studies. Patients sedated with IVS need to be monitored until they are awake for discharge[32]; IAA and PNB are good alternatives because these require less manpower although all of the evidence in this review is low to very low confidence.

From the safety and efficacy perspective, IAA may be a better choice when the patient cooperates during shoulder dislocation reduction. In contrast, IVS is a good alternative for uncooperative patients, as sedation facilitates the procedure for the operators.

STRENGTHS

To the best of our knowledge, this is the first NMA to compare the efficacy of sedation or analgesia methods for the reduction of anterior shoulder dislocation among IVS, IAA, and PNB. The strength of this NMA is the indirect comparison between IAA and PNB since there are no direct comparisons. Another strength is that we evaluated confidence in the evidence from NMA using CINeMA, which was described in the Cochrane Handbook.[12,14] Another strength of this study is the inclusion of unpublished data. Other strengths of this study are the inclusion of unpublished data and the sample size calculation for future RCTs with direct comparisons between IAA and PNB. The total number of participants required was 118.

LIMITATIONS

The NMA has some limitations. First, the number of RCTs that met the inclusion criteria was small, and the included RCTs were single-center studies with small numbers of patients. Therefore, the confidence for each relative treatment effect in NMA was very low, and the results were uncertain. Second, no studies have directly compared the IAA and PNB methods. RCTs with direct comparisons are needed in the future. At least one ongoing RCT is being conducted between the IAA and IVS methods[33] and two between the IVS and PNB methods.[34,35] In the future, updating the present systematic review and meta-analysis will improve the amount of evidence available and thus our confidence in the estimates. Third, benzodiazepines were used in most of the included studies as the IVS method compared to that in the IAA method. Short-acting sedatives such as propofol, etomidate, or propofol and ketamine (ketofol) were not used as sedatives, which may not reflect the current practice.[36] The use of short-acting sedatives may shorten ED length of stay and reduce respiratory adverse events in the IVS group. Fourth, depth of sedation was not included in many RCTs. For this reason, standardization by the depth of sedation and subgroup analysis could not be performed. Fifth, the quality of the reduction techniques and physicians varied among the studies and could not be standardized. Sixth, the definitions of the outcome periods varied among the studies. In particular, various outcome periods influence the ED length of stay and the time required for reduction. Finally, these results apply to patients with similar characteristics of the included studies because most patients were young adults.

CONCLUSIONS

The results of our network meta-analysis indicated that three sedation or analgesia methods (intravenous sedation, intraarticular anesthetic injection, and peripheral nerve block) might result in little to no difference in the success rate of reduction and patient satisfaction. Intraarticular anesthetic injection and peripheral nerve block had no adverse respiratory events. No randomized controlled trials compared the intraarticular anesthetic injection method with the peripheral nerve block method; therefore, high-quality randomized controlled trials are needed.

AUTHOR CONTRIBUTIONS

Minoru Hayashi, Naoto Kuroda, Norio Yamamoto, Akihiro Shiroshita, and Yuki Kataoka participated in the conception and design of the study. Minoru Hayashi, Kenichi Kano, and Naoto Kuroda conceived the study and acquired and interpreted the data. Minoru Hayashi

and Akihiro Shiroshita performed the analyses. Minoru Hayashi drafted the manuscript. All authors revised and approved the final version of the manuscript.

ACKNOWLEDGMENTS

The authors thank Dr. Tunku Sara Ahmad for providing us with the definition of the length of hospital stay in their study. This work was supported by the Systematic Review Workshop Peer Support Group (SRWS-PSG). We also appreciate Editage (www.editage.com) for English language editing.

CONFLICT OF INTEREST

The authors report no conflict of interest.

ORCID

Minoru Hayashi https://orcid.org/0000-0002-0160-6288
Naoto Kuroda https://orcid.org/0000-0003-1977-4893
Norio Yamamoto https://orcid.org/0000-0002-7902-9994
Akihiro Shiroshita https://orcid.org/0000-0003-0262-459X
Yuki Kataoka https://orcid.org/0000-0001-7982-5213

REFERENCES

1. Guler O, Ekinci S, Akyildiz F, et al. Comparison of four different reduction methods for anterior dislocation of the shoulder. *J Orthop Surg Res*. 2015;10:80.
2. Jiang N, Hu YJ, Zhang KR, Zhang S, Bin Y. Intra-articular lidocaine versus intravenous analgesia and sedation for manual closed reduction of acute anterior shoulder dislocation: an updated meta-analysis. *J Clin Anesth*. 2014;26:350-359.
3. Fitch RW, Kuhn JE. Intraarticular lidocaine versus intravenous procedural sedation with narcotics and benzodiazepines for reduction of the dislocated shoulder: a systematic review. *Acad Emerg Med*. 2008;15:703-708.
4. Wakai A, O'Sullivan R, McCabe A. Intra-articular lignocaine versus intravenous analgesia with or without sedation for manual reduction of acute anterior shoulder dislocation in adults. *Cochrane Database Syst Rev*. 2011;4:CD004919.
5. Blaivas M, Adhikari S, Lander L. A prospective comparison of procedural sedation and ultrasound-guided interscalene nerve block for shoulder reduction in the emergency department. *Acad Emerg Med*. 2011;18:922-927.
6. Raeyat Doost E, Heiran MM, Movahedi M, Mirafzal A. Ultrasound-guided interscalene nerve block vs procedural sedation by propofol and fentanyl for anterior shoulder dislocations. *Am J Emerg Med*. 2017;35:1435-1439.
7. Tezel O, Kaldirim U, Bilgic S, et al. A comparison of suprascapular nerve block and procedural sedation analgesia in shoulder dislocation reduction. *Am J Emerg Med*. 2014;32:549-552.
8. Hutton B, Salanti G, Caldwell DM, et al. The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations. *Ann Intern Med*. 2015;162:777-784.
9. Hayashi M, Kano K, Kuroda N, Yamamoto N, Shiroshita A, Kataoka Y. Protocol: comparative efficacy of sedation or analgesia methods for reduction of anterior shoulder dislocation: a systematic review and network meta-analysis. 2021; https://www.protocols.io/view/protocol-comparative-efficacy-of-sedation-or-analg-81wgb7jbnvpk/v3. Accessed 21 Apr 2022.
10. Baykal B, Sener S, Turkan H. Scapular manipulation technique for reduction of traumatic anterior shoulder dislocations: experiences of an academic emergency department. *Emerg Med J*. 2005;22:336-338.
11. Sterne JAC, Savović J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898.
12. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.3 (updated February 2022). Cochrane, 2022. Available from www.training.cochrane.org/handbook. Accessed 1 Aug 2022.
13. Owen RK, Bradbury N, Xin Y, Cooper N, Sutton A. MetaInsight: an interactive web-based tool for analyzing, interrogating, and visualizing network meta-analyses using R-shiny and netmeta. *Res Synth Methods*. 2019;10(4):569-581.
14. Nikolakopoulou A, Higgins JPT, Papakonstantinou T, et al. Cinema: an approach for assessing confidence in the results of a network meta-analysis. *PLoS Med*. 2020;17:1-19.
15. Papakonstantinou T, Nikolakopoulou A, Higgins JPT, Egger M, Salanti G. CINeMA: software for semiautomated assessment of the confidence in the results of network meta-analysis. *Campbell Syst Rev*. 2020;16:1-15.
16. Suder PA, Mikkelsen JB, Hougaard K, Jensen PE. Reduction of traumatic, primary anterior shoulder dislocations with local anesthesia. *J Shoulder Elbow Surg*. 1994;3:288-294.
17. Suder PA, Mikkelsen JB, Hougaard K, Jensen PE. Reduction of traumatic secondary shoulder dislocations with lidocaine. *Arch Orthop Trauma Surg*. 1995;114:233-236.
18. Kashani P, Asayesh Zarchi F, Hatamabadi HR, Afshar A, Amiri M. Intra-articular lidocaine versus intravenous sedative and analgesic for reduction of anterior shoulder dislocation. *Turk J Emerg Med*. 2016;16:60-64.
19. Matthews DE, Roberts T. Intraarticular lidocaine versus intravenous analgesic for reduction of acute anterior shoulder dislocations. A prospective randomized study. *Am J Sports Med*. 1995;23:54-58.
20. Kosnik J, Shamsa F, Raphael E, Huang R, Malachias Z, Georgiadis GM. Anesthetic methods for reduction of acute shoulder dislocations: a prospective randomized study comparing intraarticular lidocaine with intravenous analgesia and sedation. *Am J Emerg Med*. 1999;17:566-570.
21. Miller SL, Cleeman E, Auerbach J, Flatow EL. Comparison of intra-articular lidocaine and intravenous sedation for reduction of shoulder dislocations: a randomized, prospective study. *J Bone Joint Surg Am*. 2002;84:2135-2139.
22. Orlinsky M, Shon S, Chiang C, Chan L, Carter P. Comparative study of intra-articular lidocaine and intravenous meperidine/diazepam for shoulder dislocations. *J Emerg Med*. 2002;22:241-245.
23. Moharari RS, Khademhosseini P, Espandar R, et al. Intra-articular lidocaine versus intravenous meperidine/diazepam in anterior shoulder dislocation: a randomised clinical trial. *Emerg Med J*. 2008;25:262-264.
24. Cheok CY, Mohamad JA, Ahmad TS. Pain relief for reduction of acute anterior shoulder dislocations: a prospective randomized study comparing intravenous sedation with intra-articular lidocaine. *J Orthop Trauma*. 2011;25:5-10.
25. Hames H, McLeod S, Millard W. Intra-articular lidocaine versus intravenous sedation for the reduction of anterior shoulder dislocations in the emergency department. *CJEM*. 2011;13:378-383.
26. Ahmed ZABMBCH, Surgery FO. Intra-articular lidocaine compared to intravenous sedation for reduction of shoulder dislocation. *Tikrit Med J*. 2011;17:138-45.
27. Tamaoki MJ, Faloppa F, Wajnsztejn A, et al. Effectiveness of intra-articular lidocaine injection for reduction of anterior shoulder dislocation: randomized clinical trial. *Sao Paulo Med J*. 2012;130:367-372.
28. Abbasi S, Karbor R, Farsi D, et al. A comparison of the procedural sedation and analgesia (PSA) and ultrasound-guided suprascapular nerve block techniques in shoulder reduction. *J Crit Rev*. 2020;7:599-602.

29. Alkaduhimi H, van der Linde JA, Willigenburg NW, van Deurzen DFP, van den Bekerom MPJ. A systematic comparison of the closed shoulder reduction techniques. *Arch Orthop Trauma Surg*. 2017;137:589-599.
30. Driesen BEJM, van Riet BHG, Verkerk L, Bonjer HJ, Merten H, Nanayakkara PWB. Long length of stay at the emergency department is mostly caused by organisational factors outside the influence of the emergency department: a root cause analysis. *PLoS One*. 2018;13:e0202751.
31. Omer T, Perez M, Berona K, et al. Accuracy of landmark-guided glenohumeral joint injections as assessed by ultrasound in anterior shoulder dislocations. *West J Emerg Med*. 2021;22:1335-1340.
32. Green SM, Roback MG, Krauss BS, et al. Unscheduled procedural sedation: a multidisciplinary consensus practice guideline. *Ann Emerg Med*. 2019;73:e51-e65.
33. Kendall Healthcare Group, Ltd.IAA vs PS for Anterior Shoulder Dislocations. Accessed Apr 27, 2022. https://clinicaltrials.gov/ct2/show/record/NCT03625076
34. Tel-Aviv Sourasky Medical Center. Interscalene nerve block vs. Sedation for Shoulder Dislocation Reduction Accessed Apr 27, 2022. https://clinicaltrials.gov/ct2/show/NCT03041506
35. Iran University of Medical Sciences. Shoulder dislocation treatment under the anesthetic with sedation for performing the procedure and suprascapular nerve block. Accessed Apr 27, 2022. https://trialsearch.who.int/Trial2.aspx?TrialID=IRCT20151003024317N5
36. Godwin SA, Burton JH, Gerardo CJ, et al. American College of Emergency Physicians. Clinical policy: procedural sedation and analgesia in the emergency department. *Ann Emerg Med*. 2014;63:247-258.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Hayashi M, Kano K, Kuroda N, Yamamoto N, Shiroshita A, Kataoka Y. Comparative efficacy of sedation or analgesia methods for reduction of anterior shoulder dislocation: A systematic review and network meta-analysis. *Acad Emerg Med*. 2022;29:1160-1171. doi: 10.1111/acem.14568

Received: 4 April 2022 | Revised: 22 May 2022 | Accepted: 27 May 2022

DOI: 10.1111/acem.14546

ORIGINAL CONTRIBUTION

Restrictive fluids versus standard care in adults with sepsis in the emergency department (REFACED): A multicenter, randomized feasibility trial

Marie K. Jessen MD[1,2] | Lars W. Andersen DMSc[1,3,4] | Marie-Louise H. Thomsen RN[1,2] | Peter Kristensen MD[5] | Wazhma Hayeri MD[6] | Ranva E. Hassel MD[2] | Tina G. Messerschmidt RN[6] | Christoffer G. Sølling PhD[7] | Anders Perner PhD[8] | Jens Aage K. Petersen PhD[3] | Hans Kirkegaard DMSc[1,2,4]

[1]Department of Clinical Medicine, Research Center for Emergency Medicine, Aarhus University and Aarhus University Hospital, Aarhus, Denmark

[2]Department of Emergency Medicine, Aarhus University Hospital, Aarhus, Denmark

[3]Department of Anesthesiology and Intensive Care, Aarhus University Hospital, Aarhus, Denmark

[4]Prehospital Emergency Medical Services, Central Denmark Region, Aarhus, Denmark

[5]Department of Emergency Medicine, Regional Hospital Viborg, Viborg, Denmark

[6]Department of Emergency Medicine, Regional Hospital Randers, Randers, Denmark

[7]Department of Intensive Care, Regional Hospital Viborg, Viborg, Denmark

[8]Department of Intensive Care, Copenhagen University Hospital, Rigshospitalet, Copenhagen, Denmark

Correspondence
Marie K. Jessen, MD, Research Center for Emergency Medicine, Department of Clinical Medicine, Aarhus University and Aarhus University Hospital, Palle Juul-Jensens Boulevard 99, J103, 8200 Aarhus, Denmark.
Email: marie.jessen@rm.dk

Funding information Funding for the trial was provided by Carl and Ellen Hertz foundation (DKK 15,000), Frimodt-Heineke Foundation (DKK 25,000), Ruth & Holger Hesses Memorial Fund (DKK 60,000), Health Research Foundation of Central Denmark Region (DKK 215,000), "Akutpuljen" Central Denmark Region (DKK 582,000), and Aarhus University (salary for primary investigator, DKK 1.5 mio). The funding agencies had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication.

Abstract

Background: Fluid treatment in sepsis is a challenge and clinical equipoise exists regarding intravenous (IV) volumes. We aimed to determine whether a 24-h protocol restricting IV fluid was feasible in adult patients with sepsis without shock presenting to the emergency department (ED).

Methods: The REFACED Sepsis trial is an investigator-initiated, multicenter, randomized, open-label, feasibility trial, assigning sepsis patients without shock to 24 h of restrictive, crystal IV fluid administration or standard care. In the IV fluid restriction group fluid boluses were only permitted if predefined criteria for hypoperfusion occurred. Standard care was at the discretion of the treating team. The primary outcome was total IV crystalloid fluid volumes at 24 h after randomization. Secondary outcomes included total fluid volumes, feasibility measures, and patient-centered outcomes.

Results: We included 123 patients (restrictive 61 patients and standard care 62 patients) in the primary analysis. A total of 32% (95% confidence interval [CI] 28%–37%) of eligible patients meeting all inclusion criteria and no exclusion criteria were included. At 24 h, the mean (±SD) IV crystalloid fluid volumes were 562 (±1076) ml versus 1370 (±1438) ml in the restrictive versus standard care group (mean difference −801 ml,

Supervising Editor: Dr. Michael Puskarich

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2022 The Authors. *Academic Emergency Medicine* published by Wiley Periodicals LLC on behalf of Society for Academic Emergency Medicine.

95% CI −1257 to −345 ml, p = 0.001). Protocol violations occurred in 21 (34%) patients in the fluid-restrictive group. There were no differences between groups in adverse events, use of mechanical ventilation or vasopressors, acute kidney failure, length of stay, or mortality.

Conclusions: A protocol restricting IV crystalloid fluids in ED patients with sepsis reduced 24-h fluid volumes compared to standard care. A future trial powered toward patient-centered outcomes appears feasible.

INTRODUCTION

Sepsis is a global health burden, estimated to cause 11 million yearly deaths. Even in survivors, sepsis can cause permanent organ dysfunction and impaired health-related quality of life.[1,2] Infections are common in emergency department (ED) patients accounting for up to one in four admissions, among these up to one in four with sepsis.[1,3,4]

The treatment of sepsis includes intravenous (IV) antibiotics and fluids, source control, and supportive care.[5] The effect of IV fluids in sepsis is debated; strict fluid restriction may lead to impaired circulation and perfusion whereas liberal administration may lead to fluid overload resulting in edema and capillary leakage, and both too little and too much has been associated with organ dysfunction.[6–17] Although sepsis without hypotension is more common than sepsis associated hypotension and septic shock,[3] the Surviving Sepsis Campaign (SSC) only gives recommendations for fluid treatment of the latter conditions with a recommendation to give 30 ml/kg within the first 3 h.[5] However, the evidence supporting this recommendation is of low quality, the use of fluid varies, and better evidence has been requested.[4,9,18–23]

Recent observational studies and interventional trials investigating fluid volumes in adult patients with primarily septic shock in the intensive care unit (ICU) have shown either no difference or indicated benefit with fluid restriction.[7,24] IV fluids given without clear indication may be harmful.[25] Two ED-based trials from Africa in patients with sepsis and sepsis-associated hypotension found a higher mortality rate among patients in the intervention group where patients received early, aggressive fluid therapy.[26,27] Whether these results are generalizable to ED sepsis patients without shock is unknown. Although trials are currently exploring fluid strategies in patients with hypotension and septic shock,[28–30] there appear to be no trials on fluid administration in patients with early sepsis without shock or hypotension.

The aim of the Restrictive Fluid Administration vs. Standard of Care in Emergency Department Sepsis Patients (REFACED Sepsis) feasibility trial was to test if a restrictive IV fluid protocol in ED patients with sepsis without shock is feasible and could decrease the volume of IV fluids administered compared to standard care.

METHODS

Trial registration and protocol

The REFACED Sepsis trial was registered at the EU Clinical Trials Register (EudraCT number 2021–000224-35 [May 3, 2021]) and at ClinicalTrials.gov (identifier NCT05076435 [October 10, 2021]). The trial protocol was approved by the Committee on Health Research Ethics—Central Denmark Region (identifier 1-10-72-163-21 [June 28, 2021]). The protocol has previously been published[31] and is provided as a supplement to this paper. The trial was funded but the funding agencies had no role in the design, conduct or interpretation of the study, nor the decision to submit the manuscript for publication.

Trial design and setting

The REFACED Sepsis trial was an investigator-initiated, multicenter, randomized, parallel-group, open-label, feasibility trial, assigning patients with sepsis without shock to a 24-h restrictive fluid administration protocol or standard care. Participants were recruited in the EDs at Aarhus University Hospital, the Regional Hospital Randers, and the Regional Hospital Viborg. The three EDs serve a mixed rural–urban population of 0.9 million people and provide 24-h emergency care to all adult acute patients except those transferred directly to catheterization laboratories, cardiology wards and stroke units, and women in labor. ED patients are either referred by a general practitioner or brought in by ambulance after an emergency call. In the three EDs, patient contacts vary between 15,000 and 63,000 per year. Emergency health care in Denmark is publicly funded.

Selection of participants

We included patients fulfilling all of the following inclusion criteria: (1) unplanned ED admission; (2) age ≥ 18 years; (3) sepsis defined as (a) infection suspected by the treating clinician, (b) blood cultures drawn, (c) IV antibiotics administered or planned, and (d) an infection-related increase in the SOFA score ≥ 2[32]; and (4) expected hospital stay > 24 h as deemed by the treating clinician. We excluded patients

who fulfilled any of the following: (1) received ≥ 500 ml of IV fluids, (2) vasopressors or invasive ventilation started prior to screening, (3) known or suspected severe bleeding judged by the treating clinician, (4) known or suspected pregnancy, (5) prior enrollment in the trial, or (6) patients who the clinician expected not to survive the next 24 h. SOFA score was calculated automatically on the randomization website when entering laboratory values during the randomization process. A patient could be randomized as soon as the infection-related SOFA score was 2, without awaiting all laboratory values to be available for a total SOFA score at enrollment. All laboratory blood tests for a total SOFA score calculation—except an arterial blood gas analysis—were performed prior to enrollment, results were available within a maximum of 2 h, and a total SOFA score was calculated based on these post hoc. If an arterial blood gas analysis was not performed the respiratory component of the SOFA score was assumed normal, that is, giving 0 points. Known organ dysfunction was accounted for, as described in SEPSIS-3.[32] If there was any uncertainty about the impact of known organ dysfunction, this could be discussed with the primary investigator around the clock per telephone. Regarding exclusion criteria (1) and (2), we assumed that patients who had not received ≥500 ml of IV fluids and who had not received vasopressors at the time of inclusion were not in septic shock.

According to Danish law, patients with acute illness can only be included in a trial if all patients can provide written, informed consent *or* if none of the patients can provide written, informed consent, a combination is not possible. Since most sepsis patients are not able to provide informed consent, we only included patients who were unable to provide written, informed consent. As such, patients who were fully awake, oriented, and/or in no apparent distress were excluded.

Before enrollment, consent for inclusion was obtained from an independent physician followed by consent from a next of kin and/or the patient as soon as possible when they regained the capacity to provide consent. More details are presented in the protocol (Appendix S1).

Randomization

Eligible patients fulfilling all inclusion criteria and no exclusion criteria were randomly assigned in a 1:1 ratio to one of the two intervention groups. The randomization was stratified by site. Randomization was performed using a centralized Web-based system according to a computer-generated allocation sequence list with varying block sizes (4, 6, or 8), stratified by site. The allocation sequence list and block sizes were only known by the data manager at Trial Partner, Aarhus University, who was not otherwise involved in the trial.

Intervention

Patients were assigned to either restrictive IV fluid administration or standard care for 24 h. The assigned treatment protocol was followed in the ED as well as wards or ICUs if the patient was transferred within the 24-h period. The intervention protocol targeted IV crystalloid fluid administration. An overview of the trial, including the restrictive fluid algorithm, is provided in Figure 1.

In the restrictive fluid group, IV crystalloid fluids should not be given unless one of the below mentioned hypoperfusion criteria were met.

- Lactate concentration ≥ 4 mmoL/L (arterial or venous);
- Hypotension (systolic blood pressure < 90 mm Hg);
- Mottling beyond edge of kneecap (i.e., Mottling score > 2)[33];
- Severe oliguria, that is, diuresis < 0.1 ml/kg/h, during the first 4 h of admission.

If one or more of these criteria were met, a fluid bolus of 250 ml of isotonic crystalloid (isotonic saline or Ringer's acetate/lactate) could be administered per protocol. It was not a requirement that a fluid bolus was administered.

The treating physician could at any time violate the protocol by giving additional fluid if judged necessary. The physician had to state the reason for violating the protocol.

IV fluids could be given as carrier for medications, but the volume should be reduced if possible. In case of documented overt fluid loss (e.g., vomiting, large aspirates, diarrhea, drain losses, or ascites drainage) IV fluid could be given to correct for the loss. In case the oral/enteral route for water or electrolyte solutions was contraindicated or failed as judged by the clinical team, IV fluids could be given to correct significant electrolyte deficiencies or to ensure a total fluid input of 1 L per 24 h (counting all fluids including medications and nutrition). If a patient underwent surgery during the 24-h inclusion period, they temporarily paused the protocol, but clinicians were encouraged to continue restrictive fluid therapy, and all intraoperative fluids were registered.

In the standard care group, fluids were administered by clinicians' choice. Aside from the fluid administration, all management of the patient care was at the discretion of the treating team. In both groups, patients were allowed to drink as much as desired or allowed by the treating team. It was not possible to blind the intervention for neither the treating team, patients, nor relatives.

Measurements

In both trial groups, oral and IV fluids were registered on a paper case report form for 24 h (Figures S1 and S2) from randomization by the treating team with help from research assistants. Data were obtained from the electronic medical record by the research team on baseline characteristics, vital signs, blood tests, use of vasopressors, mechanical ventilation and dialysis, in-hospital course, and death and entered in an electronic case report form in REDCap.

Outcomes

The primary outcome was the total volume of IV crystalloid fluids administered during the first 24 h after randomization. The

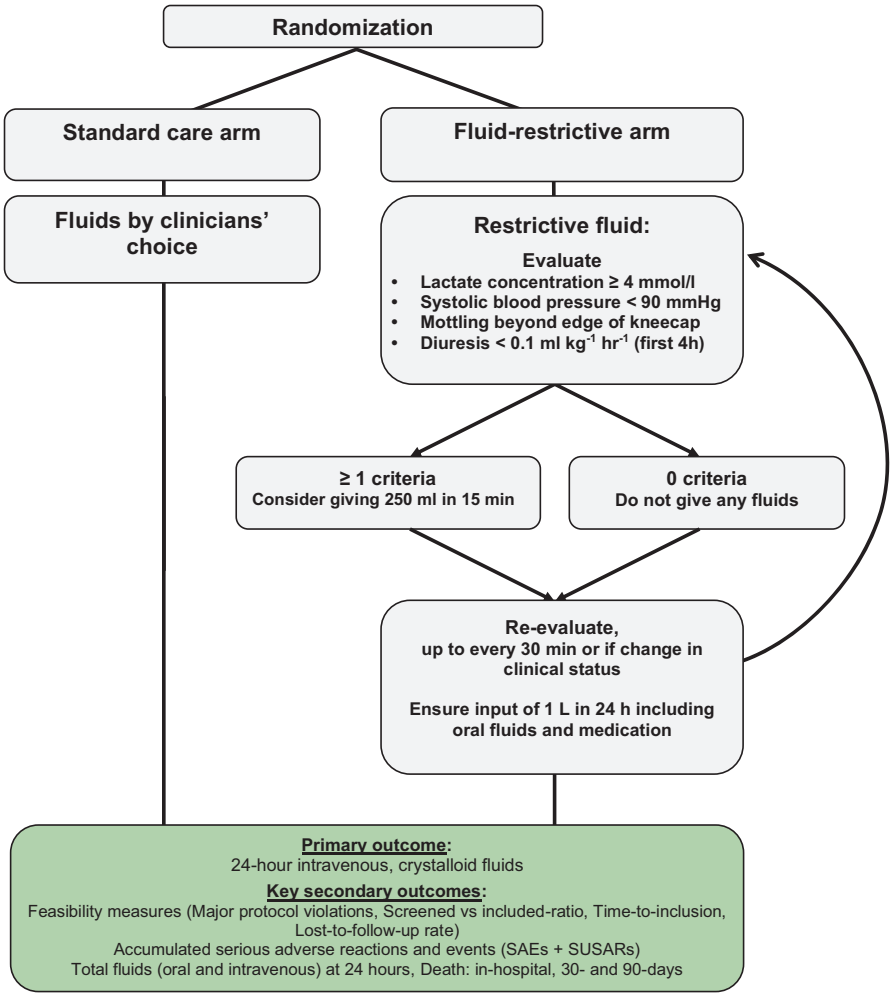


FIGURE 1 Treatment algorithms: summary of trial interventions.

secondary outcomes were feasibility measures (number of patients randomized vs. screened positive, i.e., with all inclusion criteria fulfilled and no exclusion criteria fulfilled); time from admission to inclusion; number of patients with major protocol violations; number of patients with incomplete data on the primary outcome (e.g., due to discharge or death within 24 h); serious adverse reactions and events within 7 days; total fluid volume (oral, IV, and combined) at 24 h; mechanical ventilation within 7 days; vasopressor use within 7 days; development or worsening of acute kidney failure according to the KDIGO3 criteria[34] within 7 days of randomization; hospital length of stay; and in-hospital, 30-day, and 90-day mortality.

Sample size

The sample size calculation was based on data from an observational study conducted in the Central Denmark Region in which sepsis patients meeting inclusion criteria for the current trial received a mean (±SD) of 2670 (±1695) ml IV fluids in 24 h from admission.[4] We therefore estimated that the mean (±SD) total amount of crystalloid IV fluid in the standard care group would be 2650 (±1700) ml. We considered a mean (±SD) difference of 1 L to be of clinical relevance and therefore estimated 1650 (±1700) ml in the restrictive fluid group. Based on these estimates, an alpha of 5%, a power of 90%, and a two-sample t-test, a sample size of 124 patients was required.

Data analysis

All analyses were conducted in a modified intention-to-treat population defined as all randomized patients for whom consent to use data was obtained. Baseline characteristics were compared using descriptive statistics. We used linear regression to estimate the mean difference in IV crystalloid fluid volume between the allocated groups with adjustment for the stratification variable site. As the data were not normally distributed, we performed an additional post hoc analysis using median regression to estimate the difference in medians.[35] Other continuous variables were analyzed similarly. For binary outcomes, we used logistic regression adjusted for site with differences between groups presented as odds ratios (ORs). We used summary statistics for the feasibility measures. We performed all analyses using Stata version 17 (StataCorp LP) and considered p-values of <0.05 as statistically significant.

RESULTS

Characteristics of trial participants

From November 3, 2021, to December 18, 2021, we screened 2412 unique patients with suspected infection. Of these, 383 unique patients met all inclusion criteria and no exclusion criteria, and 124 patients were randomized (Figure 2, Table S1); 62 patients were assigned to the fluid restriction group and 62 were allocated to the standard care group. One patient in the restrictive group withdrew consent for the use of data; we thus analyzed data from 123 patients (99%). One patient was inadvertently included twice. Both admissions were included in the analyses.

Patient characteristics are presented in Table 1 and Table S1. Overall, patients had a median (IQR) age of 76 (67–84) years and 58% were male. Most patients had not received IV fluids before randomization (Table 1).

Feasibility measures

Feasibility measures are shown in Table 2. Overall, 32% (95% CI 28% to 37%) of patients meeting all inclusion criteria and no exclusion criteria were included (Regional Hospital Viborg 43%, Aarhus University Hospital 41%, and Regional Hospital Randers 17% [Table S2]). Randomized patients and nonrandomized patients were similar in characteristics at admission, but nonrandomized patients more often presented with abdominal complaints whereas randomized patients more of the had respiratory complaints (Table S3). The median (IQR) time from ED admission to randomization was 140 (90–194) min. One patient died and five patients were discharged within 24 h.

Fluid results

Fluid administration during the 24-h period in both groups is presented in Table 3, Figure 3, and Table S4 and S5. At 24 h, the mean (±SD) IV crystalloid volumes were 562 (±1076) ml vs. 1370 (±1438) ml in the restrictive versus standard care group and the mean (95% CI) difference was −801 ml (−1257 to −345; p = 0.001, corresponding to a relative decrease in fluid volume of 58%. The difference in medians was −1000 ml (95% CI −1392 to −607), using median regression. Thirty-eight out of 61 (62%) patients in the restrictive group and 15

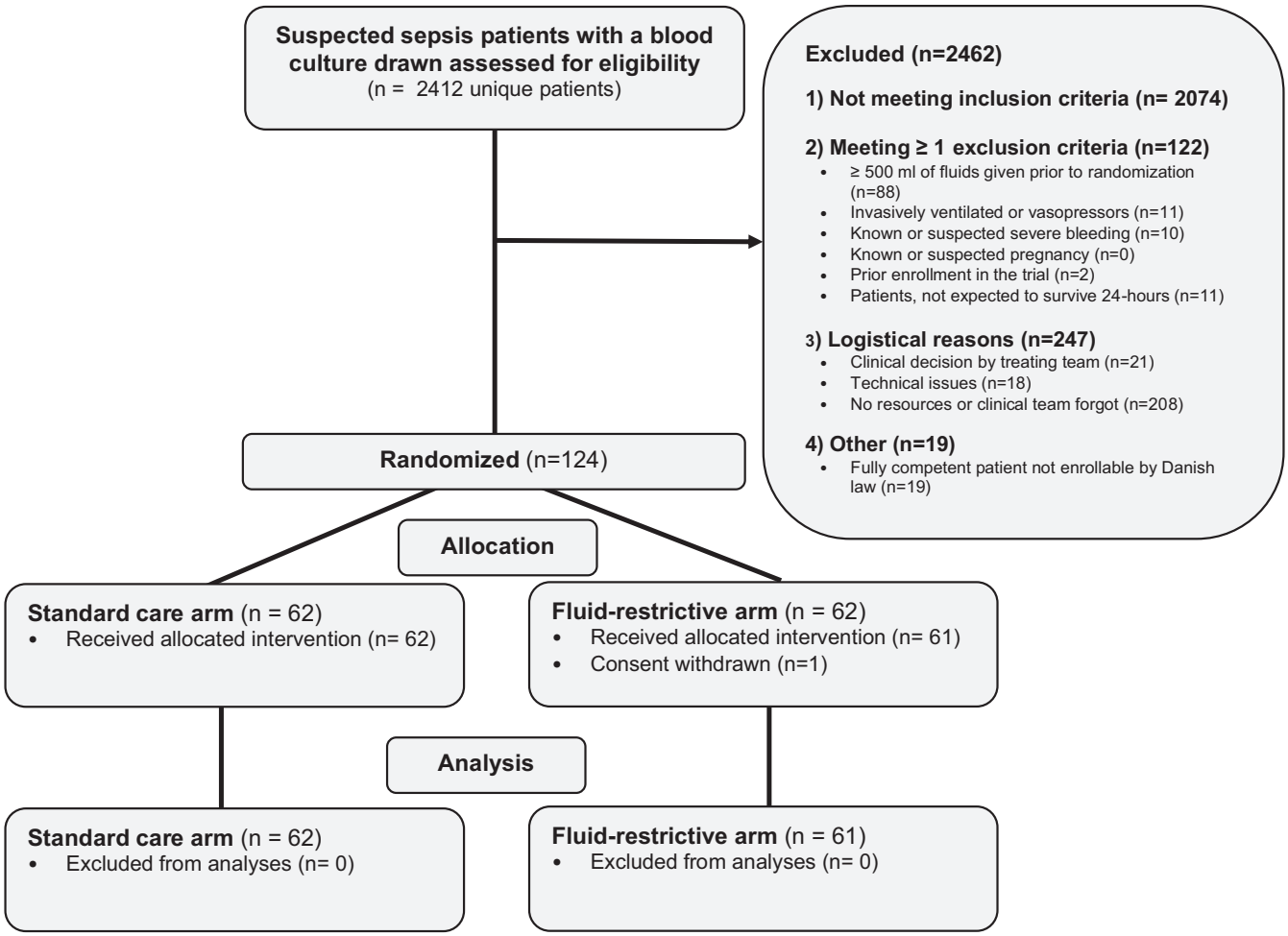


FIGURE 2 Screening, randomization, and follow-up in the REFACED Sepsis feasibility trial. REFACED Sepsis, Restrictive Fluid Administration vs. Standard of Care in Emergency Department Sepsis Patients.

TABLE 1 Baseline characteristics according to group allocation

Variable	Restrictive fluids ($n = 61$)	Standard care ($n = 62$)
Age (years)	75 (67–85)	76 (68–83)
Male sex	37 (61)	34 (55)
Weight (kg)	75 (64–92)	77 (69–90)
Prior history of comorbidities		
Kidney failure[a]	5 (8)	9 (15)
Diabetes[b]	11 (18)	9 (15)
Heart failure[c]	9 (15)	13 (21)
DNI/DNAR[d]	29 (48)	18 (29)
Vital signs at randomization		
Systolic blood pressure (mm Hg)	130 (107–144)	137 (126–147)
Diastolic blood pressure (mm Hg)	72 (62–79)	71 (62–83)
Mean arterial pressure (mm Hg)	88 (81–101)	94 (85–103)
Respiratory rate (breaths/min)	24 (20–28)	23 (20–28)
Oxygen saturation (%)	94 (91–96)	96 (93–97)
Heart rate (beats/min)	97 (80–115)	96 (88–110)
Temperature (°C)	38.1 (37.5–38.8)	38.6 (37.9–39.3)
GCS score	15 (15–15)	15 (15–15)
Blood tests before randomization		
Creatinine (μmol/L)	93 (65–136)	91 (65–132)
Platelet count ($\times 10^9$/L)	247 (179–299)	230 (157–323)
Bilirubin (μmol/L)	11 (7–21)	12 (8–18)
Leukocytes ($\times 10^9$/L)	14.2 (10.7–17.4)	13.5 (9.6:17.9)
C-reactive protein (mg/L)	117 (47–194)	125 (55–235)
Lactate (mmol/L)	1.2 (1.0–1.8)	1.4 (1.0–2.1)
Total SOFA score at randomization[e]	3 (2–3)	3 (2–3)
Suspected infectious source[f]		
Respiratory [*n* with COVID-19]	45 (74) [4]	43 (69) [3]
Urinary	9 (15)	12 (19)
Skin/soft tissue	3 (5)	1 (2)
Abdominal	3 (5)	5 (8)
Other/unknown	3 (5)	4 (6)
Time to IV antibiotics from admission (h)	2.8 (1.6–3.9)	2.9 (1.6–3.9)
IV fluids given prior to randomization (ml)	0 [0–200]	0 [0–100]

Note: All data are presented as median (IQR) or *n* (%) unless otherwise stated.Abbreviations: DNI/DNAR, do not intubate or do not attempt resuscitation orders; GCS, Glasgow Coma Scale.

[a]Renal failure defined according to KDIGO criteria (see supplemental material).

[b]Diabetes requiring chronic oral or injection treatment.

[c]Heart failure with history of ejection fraction ≤ 40%.

[d]DNI and/or DNAR documented prior to or within 6 h of admission.

[e]For SOFA subscores, see Table S1.

[f]Some patients had more than one infectious source, why the total sum is >100%.

of 62 (24%) patients in the standard care group received no IV crystalloid fluids in the first 24 h (Figure 3, Table S4). The mean (±SD) of combined oral and IV fluids in the first 24 h was 2881 (±1295) ml in the restrictive group versus 3720 (±1623) ml in the standard care group with a mean difference of −840 ml (95% CI −1364 to −317, p = 0.002). Further details of fluid administration, type of fluid, and time intervals are shown in Table 2, Figure 3, Tables S3–S6, and Figures S4–S7.

In the restrictive fluid group, hypotension was the most frequently used hypoperfusion criterion for administering fluids per protocol. Twenty-one of 61 patients had fluid administered despite no criteria fulfilled, that is, had a protocol violation. High or rising creatinine/impaired renal function was the most frequently used specific reason for giving IV fluids outside the protocol (Table 4). One patient (standard care group) underwent surgery and had the fluid resuscitation protocol temporarily suspended during surgery in accordance with the protocol and had a total 2750 ml of IV fluid and medication administered during surgery, included in total volumes but not in IV crystalloid fluid volumes.

Secondary outcomes

There were no significant differences between groups in use of mechanical ventilation or vasopressors or new-onset acute kidney failure at 7 days, nor in-hospital length of stay or in-hospital, 30-day, or 90-day mortality (Table 5).

Adverse events

There were 17 (28%) and 18 (29%) patients experiencing any of the predefined adverse events or reactions in the restrictive and standard care groups, respectively, with acute myocardial infarction, death, new-onset acute kidney injury, and hypervolemia accounting for all events (Table 2 and Tables S7 and S8).

DISCUSSION

We conducted a randomized, multicenter trial to examine the feasibility of restricting 24-h IV, crystalloid fluid volumes in sepsis patients without shock in three EDs. The restrictive protocol significantly reduced 24-h IV fluid volumes and total fluid volumes.

Despite a low randomized-to-screened ratio, we included 124 patients within 6 weeks at three sites. Although randomization required drawing of blood cultures and results from laboratory values prior to randomization, patients were randomized within a median of 140 min of arrival to the ED and most patients did not receive IV fluids prior to randomization. Based on these feasibility measures, we consider a larger trial feasible.

Fluid volumes in the current trial were lower than those in our previous cohort study.[4] In our sample size estimation, we assumed

TABLE 2 Feasibility measures and secondary effect parameters stratified by group allocation

	Restrictive fluids ($n = 61$)	Standard care ($n = 62$)	Overall ($n = 123$)
Screened eligible/included ratio (%)	—	—	124/383 = 32% (95% CI 28%–37%)[a]
Time from ED admission to inclusion (min)			
Mean (±SD)	149 (±76)	161 (±106)	155 (±92)
Median (IQR)	140 (90–197)	139 (92–179)	140 (90–194)
Patients with incomplete data on primary outcome	2 (3)	4 (7)	6 (5)
Reasons for lost to follow-up within 24 h	1 discharge 1 death	4 discharges	5 discharges 1 death
Patients with protocol violations	21 (34)[b]	—	—
Patients who received no crystalloid fluid within 24 h of enrollment	38 (62)	15 (24)	53 (43)
Accumulated adverse reactions and events within 7 days	17 3 deaths, 1 myocardial infarction, 4 hypervolemia, 9 acute kidney injury	18 1 death, 1 heart failure, 2 myocardial infarctions, 4 hypervolemia, 10 acute kidney injury	35

Note: All data are presented as *n* (%) unless otherwise stated.

[a]For site-specific screening/included ratio and explanations, see Table S2.

[b]IV fluids given if none of the following was true: (a) one or more hypoperfusion criteria fulfilled; (b) to correct documented fluid loss; (c) to correct significant electrolyte deficiencies; (d) fluid administered as carrier for medication (e.g., antibiotics); (e) ensure a total fluid input of 1 L per 24 h (for the specific reasons, see Table S6).

that the total mean (±SD) amount of IV fluid in the standard care group would be 2650 (±1700) ml as it was for similar patients in our descriptive study.[4,36] However, the standard care group only received 2067 (±1655) ml total IV fluids in the present trial. This may represent the Hawthorne effect and/or a change in current practice toward more restrictive fluid administration in general. However, there is limited evidence on patient-centered outcomes to support this change of practice yet.[28,30,37,38] On the other hand, patients in the REFACED Sepsis trial received approximately 300 ml more oral fluids than in our previous descriptive study (1650 ml compared to 1319 ml), and in general patients had a large proportion of the total 24-h fluids through the enteral route.

The trial protocol was, in line with recent trials,[29,38,39] able to substantially reduce IV fluid volumes. Although the mean difference (801 ml) was slightly lower than the estimate used in our sample size calculation (1000 ml), the median difference was 1000 ml, and the relative reduction was large (58%). Given the relative low volume of fluid in the control group, we consider a separation of 801 ml satisfactory and the protocol successful. In ED patients with sepsis-associated hypotension, the REFRESH trial reduced 24-h fluids from 4250 to 3543 ml in the restrictive group, with a relative reduction of 30%. Although the absolute reduction in fluid volume administered were similar between this and the current trial, we almost doubled the relative fluid volume reduction in REFACED Sepsis (58%). The 58% reduction is more in line with the RIFTS pilot trial, where ICU patients with sepsis or septic shock received 665 ± 1119 ml in the restrictive group and 1251 ± 1588 ml in the usual care group with a mean difference of 586 (62–1109) ml in the first 24 h postrandomization and a relative reduction of 47%[38] and the ICU-based CLASSIC septic shock feasibility trial, although the intervention in CLASSIC lasted for up to 5 days.[39] However, in all the above-mentioned trials, patients received large fluid volumes prior to randomization in opposition to this current trial resulting in total fluids exceeding our totals but all three trials were also conducted in more severely ill patients. The REFACED Sepsis, REFRESH, and CLASSIC trials use patient specific hypoperfusion criteria for administering fluid in contrary to a "one-size-fits-all" strategy for example with a fixed fluid volume for all patients.[29,39]

The REFACED Sepsis study and its use of hypoperfusion criteria was inspired by the CLASSIC trials.[28,39] The four hypoperfusion criteria were chosen to represent central (systolic blood pressure), general (lactate), peripheral (mottling), and renal (oliguria) circulation and perfusion status. The cutoff value of lactate was chosen based on the former SSC guideline (2016),[40] and their 1-h bundle[41] and

TABLE 3 Fluid volumes in the first 24 h stratified by group allocation

	Restrictive fluids ($n = 61$)	Standard care ($n = 62$)	Mean difference (95% CI) or difference in medians [95% CI][a]	p-value for mean difference
Primary outcome				
24-h IV crystalloid fluid volumes (ml)				
Mean (±SD)	562 (±1076)	1370 (±1438)	−801 (−1257 to −345)	0.001
Median [IQR]	0 [0–600]	1000 [80–2000]	−1000 [−1392 to −607][a]	
24-h IV crystalloid fluid volumes per kg bodyweight (ml/kg)				
Mean (±SD)	9 (±16)	17 (±19)	−9 (−15 to −2)	0.007
Median [IQR]	0 [0–11]	12.5 [1–26]		
Secondary outcomes				
24-h oral and IV fluid volumes (ml)				
Mean (SD)	2881 (1295)	3720 (1623)	−840 (−1364 to −317)	0.002
Median [IQR]	2820 [1900–3500]	3498 [2800–4450]	−660 [−1116 to −204][a]	
24-h oral and IV fluid volumes per kg bodyweight (ml/kg)				
Mean (±SD)	38 (±20)	48 (±22)	−9 (−17 to −2)	0.18
Median [IQR]	36 [22–49]	45 [32–56]		
24-h other IV fluids[b] (ml)				
Mean (±SD)	667 (±500)	697 (±705)	−35 (−252 to 182)	0.75
Median [IQR]	500 [400–800]	416 [300–800]		
24-h total IV fluid volume (ml)				
Mean (±SD)	1229 (±1292)	2067 (±1678)	−837 (−1374 to −298)	0.003
Median [IQR]	792 [400–1400]	1625 [1200–2650]		
24-h total oral fluid volume (ml)				
Mean (SD)	1651 (888)	1653 (816)	−4 (−310; 302)	0.98
Median [IQR]	1750 [1100–2225]	1600 [950–2150]		

Note: This table shows fluid volumes in the restrictive fluid group and in the standard care group. Mean differences and differences in medians as well as p-values are derived from the regression analyses. All mean and median differences are estimated with the standard care group as reference.

[a]Adjusted for site. Median regression was only performed for the predefined primary and secondary outcomes.

[b]Other IV fluids accounts for dissolved IV administered medication, glucose, plasma, albumin, blood, etc. (for further information, see Table S5).

data indicating that the marked increase in mortality occur at lactate values > 4 mmoL/L.[42,43] The mottling trigger was based on mottling score of ≥2 as described by Ait-Oufella et al.[33] and validated in a pre-hospital setting.[44] Recently, some trials have used capillary refill time as a marker of peripheral perfusion,[45,46] which could have been used instead of mottling but it was chosen to align with the CLASSIC criteria. Severe oliguria was defined as urine output ≤ 0.1 ml/kg/h and the criterion was only to be used within the first 4 h of admission. In the REFACED Sepsis trial, a total of 29 boli of 250 ml crystalloid were given per protocol in the 61 patients in the restrictive fluid group (Table 4). The protocol was violated (i.e., giving fluids although no hypoperfusion criteria were fulfilled) in 35% in the fluid restrictive group and 24% did not receive IV, crystalloid fluids in 24 h in the standard care group, in comparison to 45% and 30%, respectively, in the CLASSIC feasibility trial.[39] Overall, the fact, that 38/61 (62%) patients in the restrictive group did not receive crystalloid fluids unless as carrier for medication, to correct electrolytes or to replace fluid loss, shows that clinicians are able to restrict fluids in a large proportion of sepsis patients.

Interestingly, the most frequently used specific reason for giving IV fluids outside the protocol in the fluid-restrictive group was high or rising creatinine/impaired renal function. This may be due to a general perception that a low degree of prerenal kidney failure should be treated with fluids. However, the evidence for this to our knowledge is limited, and descriptive studies show improvement with less fluids.[47–50]

The strengths of our trial include the multicenter inclusion, recruitment in both university and regional hospitals, and a short inclusion period. The fast inclusion and completion of the trial underlines the importance of the trial; sepsis patients account for a large proportion of ED patients. We believe, this patient population (i.e., older, high do not intubate/do not attempt resuscitation [DNI/

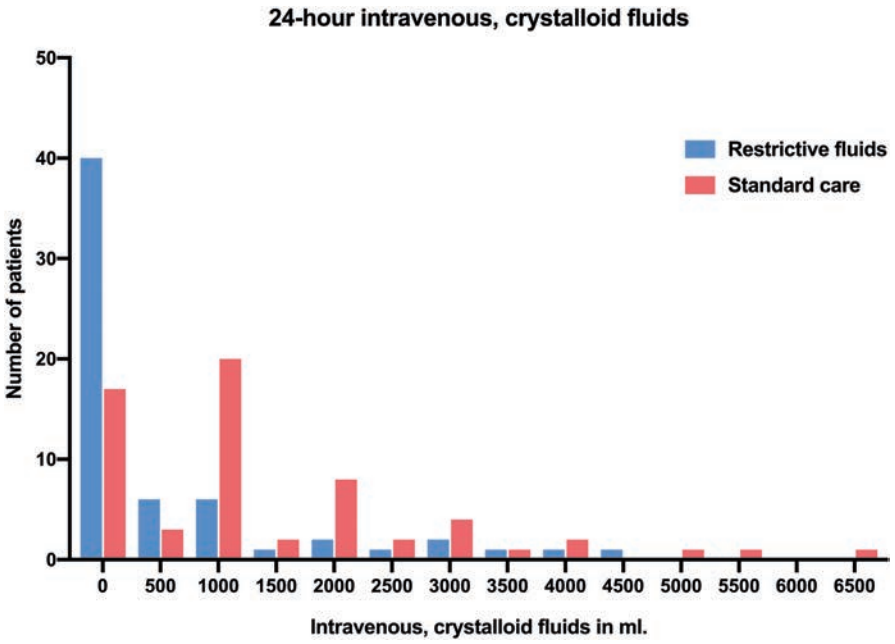


FIGURE 3 Distributions of 24-h IV crystalloid fluids by group allocation. Histogram showing distributions of 24-h IV, crystalloid fluids in ml by group allocation. The y-axis represents the number of patients with the given fluid volume from each group.

TABLE 4 Reasons for fluid administration and protocol violations in the restrictive fluid group

Description of fluid indication	Number of patients with bolus/boli given for the fluid indication, *N*/total (%)	Number of 250 ml crystalloid boli given for the fluid indication, *n*
Hypoperfusion criteria		
Lactate concentration ≥ 4 mmol/L[a]	2/61 (3.3%)	2
Hypotension (sBP < 90 mm Hg)	9/61(14.8%)	24
Mottling beyond edge of kneecap[b]	1/61 (1.6%)	1
Severe oliguria, i.e., diuresis < 0.1 ml/kg/h[c]	2/61 (3.3%)	2
Other allowed reasons for fluid administration		
Correct significant electrolyte deficiencies	3/61 (4.9%)	4
Replace fluid loss	0	0
Ensure a total fluid input of 1 L per 24 h[d]	0	0
Protocol violations		
Improve circulation or low blood pressure (but sBP ≥90 mm Hg)	5/61 (8.2%)	9
High or rising creatinine or impaired kidney function	7/61 (11.5%)	12
Dehydration indicated by treating physician	6/61 (9.8%)	9
Other reasons or administration by mistake[e]	12/61 (19.7%)	16

Abbreviation: sBP, systolic blood pressure.

[a]Lactate measurement from an arterial or venous blood gas/blood sample.

[b]Mottling score > 2 as described by Ait-Oufella et al.[33]

[c]Criteria only possible to use within first 4 h after randomization.

[d]Total fluid input included oral fluids and fluids given with medication.

[e]Other reasons included administering more fluid than allowed by study protocol, sparse urine output > 4 h from randomization, administration of fluid by mistake outside of protocol.

DNAR] rate, unable to consent) represents a very important patient group in the ED, which have traditionally not been included in clinical trials. We consider the inclusion of this patient population a strength as it increases the generalizability of the results.

There are some important considerations for a possible future large-scale trial. It could be of interest to include patients slightly sicker but still without septic shock at arrival, that is, including more patients with low blood pressure. Since these patients often rapidly have fluids administered prehospital or in hospital and thereby fulfill the exclusion criteria of receiving >500 ml before they could possibly have been included, it would require even closer contact to the prehospital services and first-line in-hospital treating team to limit IV

TABLE 5 Secondary outcomes stratified by group allocation.

Variable	Restrictive fluids ($n = 61$)	Standard Care ($n = 62$)	Effect estimate[a] Mean difference (95% CI) and median difference [95% CI][b]	*p*-value for effect estimate
In-hospital length of stay (days)				
Mean (SD)	7.5 (4.9)	6.2 (5.9)	1.2 (−0.8; 3.1)	0.24[c]
Median [IQR]	5.9 [4.0; 10.0]	4.9 [3.0; 7.3]	0.8 [−0.7; 2.4]	
			Odds ratio (95% CI)	
Mechanical ventilation within 7 days	2 (3.3%)	2 (3.2%)	1.01 (0.14–7.43)	0.99
Vasopressors within 7 days	2 (3.3%)	4 (6.5%)	0.49 (0.09–2.79)	0.42
New onset or worsening acute kidney failure within 7 days[d]	9 (14.8%)	10 (16.1%)	0.90 (0.34–2.39)	0.83
Mortality, in-hospital	7 (11.5%)	6 (9.7%)	1.19 (0.37–3.83)	0.80
Mortality, 30 days	9 (14.8%)	10 (16.1%)	0.82 (0.34–2.39)	0.83
Mortality, 90 days[e]	12 (19.7%)	15 (25.0%)	0.73 (0.31–1.74)	0.48

Note: All data are reported as numbers (%) if not otherwise stated.

Abbreviations: CI, confidence interval; IQR, interquartile range; SD, standard deviation.

[a]All analyses of effect are adjusted for site.

[b]Median regression adjusted for site.

[c]*p* value for mean difference.

[d]Any development or worsening of acute kidney injury, defined as the KDIGO[34] creatinine score >0 compared to at randomization.

[e]Two patients withdrew consent to obtain 90-day mortality status, both from the standard care group.

fluid administration prior to and at arrival and thereby increase the chance of randomizing the patient. This may be appropriate since the evidence is still very sparse both prehospital and in hospital. Including the sickest sepsis patients, but still without shock, would probably increase the chance of finding a difference in outcomes between treatment arms. Also, it would increase inclusion rates and the generalizability to include more patients with abdominal infections. To ensure an even greater separation between the groups, even stricter criteria for fluid administration in the restrictive fluid group should be ensured, maybe focusing even more on changing from IV to oral fluid administration in this group. Also, the intervention period could be extended.

We found a high prevalence of DNI/DNAR orders within the sepsis population in the REFACED Sepsis trial, and the population was in general older with a median age of 76 years and a high mortality (30-day mortality of 15%) in comparison to other sepsis studies, but similar to our descriptive study leading up to this study.[4,29,38,51] If an effect of fluid restriction on patient-centered outcomes, for example, mortality or days alive at home, will be found in a future large-scale trial, there is a potential of improving outcomes for a large patient group with significant mortality and high burden on health care systems.

LIMITATIONS

The trial was designed to show differences in IV fluid volumes and the sample size was therefore inadequate to assess clinical outcomes such as mortality. We did not prespecify any benchmarks for feasibility. It was not possible to blind the allocated intervention for neither patients, the treating team, nor investigators, which may have affected the results and potentially caused fluid in the standard care group to be quite restrictive in comparison to our previous cohort study.[4] Since patients were excluded if more than 500 ml of IV fluids had been administered prior to randomization, we may have missed patients presenting with more severe illness prehospitally or at ED admission. Also, patients who fulfilled all inclusion criteria later in their ED course may have been missed. Both above-mentioned limitations could affect the generalizability or the results.

The proportion of elderly patients, and patients with DNI/DNAR orders was high in the REFACED Sepsis trial, resulting in a high mortality rate compared to other sepsis studies, although it was similar to the ones found in a cohort study conducted at two of the sites in REFACED Sepsis.[4] Also, the fact that 19 patients were not included since they were actually able to provide consent and there for not includable due to Danish regulations may have caused inclusion of sicker patients. The trial was conducted during autumn and winter season, during the COVID-19-pandemic, which likely resulted in inclusion of more patients with respiratory symptoms. At two sites, patients with a high risk of surgery within the 24 h were not enrolled (described in supplemental material), resulting in few patients with abdominal symptoms in comparison to other trials in sepsis and fluids affecting the generalizability.[29,38,39] These local conditions, as well as challenges related to COVID-19, contributed to a low included-to-screened positive ratio. To keep the in-hospital patient flow, multiple

departments and clinical personnel were involved in this trial at the three hospitals causing organizational challenges and obstacles, and thus presence of investigators, research nurses, and assistants was necessary. No differences in adverse events were seen in the two groups; however, the study was not powered to show certain differences in these.

CONCLUSIONS

The results indicate that it is feasible to protocolize and restrict 24-h intravenous fluid volumes in sepsis patients without shock in EDs. The mean difference (801 ml) was slightly lower than the estimate used in our sample size calculation (1000 ml), but the median difference was 1000 ml, and the relative reduction was large. A large-scale trial to investigate the effect of restrictive fluids on patient-centered outcomes appears feasible; however, modifications to the protocol may increase the separation in intravenous fluid volumes between the two intervention groups.

AUTHOR CONTRIBUTIONS

Marie K. Jessen, Lars W. Andersen, Anders Perner, Jens Aage K. Petersen, and Hans Kirkegaard conceived the study and designed the trial. Marie K. Jessen and Hans Kirkegaard obtained research funding. Lars W. Andersen, Anders Perner, Jens Aage K. Petersen, and Hans Kirkegaard supervised the conduct of the trial. Marie K. Jessen, Marie-Louise H. Thomsen, Peter Kristensen, Wazhma Hayeri, Tina G. Messerschmidt, Christoffer G. Sølling, and Ranva E. Hassel undertook recruitment of patients. Peter Kristensen, Wazhma Hayeri, Christoffer G. Sølling, and Ranva E. Hassel were site investigators. Marie-Louise H. Thomsen and Marie K. Jessen collected data. Marie K. Jessen managed the data, including quality control. Lars W. Andersen provided statistical advice. Marie K. Jessen drafted the manuscript, and all authors contributed substantially to its revision. Marie K. Jessen takes responsibility for the paper as a whole.

ACKNOWLEDGMENTS

We are very grateful to the clinical staff of doctors and nurses at the participating departments for their important contribution and to patients and relatives for their consent to participate.

CONFLICT OF INTEREST

The Department of Intensive Care, Rigshospitalet, where AP is affiliated, receives funding for research from the Novo Nordisk Foundation, Pfizer, Fresenius Kabi, AK Pharma, and Sygeforsikringen "danmark" outside the submitted work. AP is also the sponsor of the CLASSIC trial. The authors declare no potential conflict of interest.

TRIAL REGISTRATION

EudraCT number 2021-000224-35 (2021-05-03), ClinicalTrials.gov number NCT05076435 (date 2021-10-13), Committee on Health Research Ethics—Central Denmark Region number 1-10-72-163-21 (date 2021-06-28).

ORCID

Marie K. Jessen https://orcid.org/0000-0001-9445-7690
Lars W. Andersen https://orcid.org/0000-0001-5752-8082
Christoffer G. Sølling https://orcid.org/0000-0003-0721-926X
Anders Perner https://orcid.org/0000-0002-4668-0123
Jens Aage K. Petersen https://orcid.org/0000-0002-6174-4527
Hans Kirkegaard https://orcid.org/0000-0003-4853-8152

REFERENCES

1. Rudd KE, Johnson SC, Agesa KM, et al. Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the global burden of disease study. *Lancet*. 2020;395(10219):200-211. doi:10.1016/s0140-6736(19)32989-7
2. Winters BD, Eberlein M, Leung J, Needham DM, Pronovost PJ, Sevransky JE. Long-term mortality and quality of life in sepsis: a systematic review. *Crit Care Med*. 2010;38(5):1276-1283. doi:10.1097/CCM.0b013e3181d8cc1d
3. Henriksen DP, Laursen CB, Jensen TG, Hallas J, Pedersen C, Lassen AT. Incidence rate of community-acquired sepsis among hospitalized acute medical patients-a population-based survey. *Crit Care Med*. 2015;43(1):13-21. doi:10.1097/ccm.0000000000000611
4. Jessen MK, Andersen LW, Thomsen MH, et al. Twenty-four-h fluid administration in emergency department patients with suspected infection: a multicenter, prospective, observational study. *Acta Anaesthesiol Scand*. 2021;65(8):1122-1142. doi:10.1111/aas.13848
5. Evans L, Rhodes A, Alhazzani W, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock 2021. *Crit Care Med*. 2021;49(11):e1063-e1143. doi:10.1097/ccm.0000000000005337
6. Ueyama H, Kiyonaka S. Predicting the need for fluid therapy-does fluid responsiveness work? *J Intensive Care*. 2017;5:34. doi:10.1186/s40560-017-0210-7
7. Malbrain ML, Marik PE, Witters I, et al. Fluid overload, de-resuscitation, and outcomes in critically ill or injured patients: a systematic review with suggestions for clinical practice. *Anaesthesiol Intensive Ther*. 2014;46(5):361-380. doi:10.5603/ait.2014.0060
8. Seymour CW, Gesten F, Prescott HC, et al. Time to treatment and mortality during mandated emergency care for Sepsis. *N Engl J Med*. 2017;376(23):2235-2244.
9. Marik PE, Linde-Zwirble WT, Bittner EA, Sahatjian J, Hansell D. Fluid administration in severe sepsis and septic shock, patterns and outcomes: an analysis of a large national database. *Intensive Care Med*. 2017;43(5):625-632. doi:10.1007/s00134-016-4675-y
10. Boyd JH, Forbes J, Nakada TA, Walley KR, Russell JA. Fluid resuscitation in septic shock: a positive fluid balance and elevated central venous pressure are associated with increased mortality. *Crit Care Med*. 2011;39(2):259-265. doi:10.1097/CCM.0b013e3181feeb15
11. Kelm DJ, Perrin JT, Cartin-Ceba R, Gajic O, Schenck L, Kennedy CC. Fluid overload in patients with severe sepsis and septic shock treated with early goal-directed therapy is associated with increased acute need for fluid-related medical interventions and hospital death. *Shock*. 2015;43(1):68-73. doi:10.1097/shk.0000000000000268
12. Wu X, Hu Z, Yuan H, Chen L, Li Y, Zhao C. Fluid resuscitation and markers of glycocalyx degradation in severe sepsis. *Open Med (Wars)*. 2017;12:409-416.
13. Malbrain M, Van Regenmortel N, Saugel B, et al. Principles of fluid management and stewardship in septic shock: it is time to consider

the four D's and the four phases of fluid therapy. *Ann Intensive Care*. 2018;8(1):66. doi:10.1186/s13613-018-0402-x

14. Sethi M, Owyang CG, Meyers C, Parekh R, Shah KH, Manini AF. Choice of resuscitative fluids and mortality in emergency department patients with sepsis. *Am J Emerg Med*. 2018;36(4):625-629. doi:10.1016/j.ajem.2017.09.042
15. Shaw AD, Raghunathan K, Peyerl FW, Munson SH, Paluszkiewicz SM, Schermer CR. Association between intravenous chloride load during resuscitation and in-hospital mortality among patients with SIRS. *Intensive Care Med*. 2014;40(12):1897-1905. doi:10.1007/s00134-014-3505-3
16. Byrne L, Obonyo NG, Diab SD, et al. Unintended consequences: fluid resuscitation worsens shock in an ovine model of endotoxemia. *Am J Respir Crit Care Med*. 2018;198(8):1043-1054. doi:10.1164/rccm.201801-0064OC
17. Messmer AS, Zingg C, Müller M, Gerber JL, Schefold JC, Pfortmueller CA. Fluid overload and mortality in adult critical care patients-a systematic review and meta-analysis of observational studies. *Crit Care Med*. 2020;48(12):1862-1870. doi:10.1097/ccm.0000000000004617
18. Hjortrup PB, Haase N, Wetterslev J, Perner A. Associations of hospital and patient characteristics with fluid resuscitation volumes in patients with severe sepsis: post hoc analyses of data from a multicentre randomised clinical trial. *PLoS One*. 2016;11(5):e0155767.
19. Angus DC, Barnato AE, Bell D, et al. A systematic review and meta-analysis of early goal-directed therapy for septic shock: the ARISE, ProCESS and ProMISe investigators. *Intensive Care Med*. 2015;41(9):1549-1560. doi:10.1007/s00134-015-3822-1
20. Keijzers G, Macdonald SP, Udy AA, et al. The Australasian Resuscitation in SEPSIS Evaluation: Fluids or Vasopressors in Emergency Department Sepsis (ARISE FLUIDS), a multi-Centre observational study describing current practice in Australia and New Zealand. *Emerg Med Australas*. 2020;32(4):586-598.
21. Alhazzani W, Moller MH, Arabi YM, et al. Surviving sepsis campaign: guidelines on the management of critically ill adults with coronavirus disease 2019 (COVID-19). *Crit Care Med*. 2020;48:e440-e469. doi:10.1097/ccm.0000000000004363
22. Harris T, Coats TJ, Elwan MH. Fluid therapy in the emergency department: an expert practice review. *Emerg Med J*. 2018;35(8):511-515. doi:10.1136/emermed-2017-207245
23. Perner A, Gordon AC, Angus DC, et al. The intensive care medicine research agenda on septic shock. *Intensive Care Med*. 2017;43(9):1294-1305. doi:10.1007/s00134-017-4821-1
24. Meyhoff TS, Møller MH, Hjortrup PB, Cronhjort M, Perner A, Wetterslev J. Lower vs higher fluid volumes during initial management of sepsis: a systematic review with meta-analysis and trial sequential analysis. *Chest*. 2020;157(6):1478-1496. doi:10.1016/j.chest.2019.11.050
25. Nijssen EC, Rennenberg RJ, Nelemans PJ, et al. Prophylactic hydration to protect renal function from intravascular iodinated contrast material in patients at high risk of contrast-induced nephropathy (AMACING): a prospective, randomised, phase 3, controlled, open-label, non-inferiority trial. *Lancet*. 2017;389(10076):1312-1322. doi:10.1016/s0140-6736(17)30057-0
26. Andrews B, Semler MW, Muchemwa L, et al. Effect of an early resuscitation protocol on in-hospital mortality among adults with sepsis and hypotension: a randomized clinical trial. *JAMA*. 2017;318(13):1233-1240. doi:10.1001/jama.2017.10913
27. Andrews B, Muchemwa L, Kelly P, Lakhi S, Heimburger DC, Bernard GR. Simplified severe sepsis protocol: a randomized controlled trial of modified early goal-directed therapy in Zambia. *Crit Care Med*. 2014;42(11):2315-2324. doi:10.1097/ccm.0000000000000541
28. Meyhoff TS, Hjortrup PB, Møller MH, et al. Conservative vs liberal fluid therapy in septic shock (CLASSIC) trial-protocol and statistical analysis plan. *Acta Anaesthesiol Scand*. 2019;63(9):1262-1271. doi:10.1111/aas.13434
29. Macdonald SPJ, Keijzers G, Taylor DM, et al. Restricted fluid resuscitation in suspected sepsis associated hypotension (REFRESH): a pilot randomised controlled trial. *Intensive Care Med*. 2018;44(12):2070-2078. doi:10.1007/s00134-018-5433-0
30. Self WH, Semler MW, Bellomo R, et al. Liberal versus restrictive intravenous fluid therapy for early septic shock: rationale for a randomized trial. *Ann Emerg Med*. 2018;72:457-466. doi:10.1016/j.annemergmed.2018.03.039
31. Jessen MK, Andersen LW, Thomsen MH, et al. Restrictive Fluid Administration vs. Standard of Care in Emergency Department Sepsis Patients (REFACED Sepsis)—protocol for a multicenter, randomized, clinical, proof-of-concept trial. *Pilot Feasibility Stud*. 2022;8(1):75. doi:10.1186/s40814-022-01034-y
32. Singer M, Deutschman CS, Seymour CW, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA*. 2016;315(8):801-810. doi:10.1001/jama.2016.0287
33. Ait-Oufella H, Lemoinne S, Boelle PY, et al. Mottling score predicts survival in septic shock. *Intensive Care Med*. 2011;37(5):801-807. doi:10.1007/s00134-011-2163-y
34. KDIGO clinical practice guideline for acute kidney injury. *Kidney International Supplements*. 2012;2:1.
35. Staffa SJ, Zurakowski D. Calculation of confidence intervals for differences in medians between groups and comparison of methods. *Anesth Analg*. 2020;130(2):542-546. doi:10.1213/ane.0000000000004535
36. Jessen MK, Andersen LW, Thomsen MH, et al. Restrictive Fluid Administration vs. Standard of Care in Emergency Department Sepsis Patients (REFACED Sepsis)—protocol for a multicenter, randomized, clinical, proof-of-concept trial. Pilot and Feasibility Studies 2022; Accepted for publication. doi:10.1186/s40814-022-01034-y
37. Australasian Resuscitation In Sepsis Evaluation: FLUid or Vasopressors In Emergency Department Sepsis (ARISE FLUIDS). ClinicalTrials.gov NCT04569942. First posted September 3, 2020. Access date: November 21, 2021. https://clinicaltrials.gov/ct2/show/record/NCT04569942.
38. Corl KA, Prodromou M, Merchant RC, et al. The restrictive IV fluid trial in severe sepsis and septic shock (RIFTS): a randomized pilot study. *Crit Care Med*. 2019;47(7):951-959. doi:10.1097/ccm.0000000000003779
39. Hjortrup PB, Haase N, Bundgaard H, et al. Restricting volumes of resuscitation fluid in adults with septic shock after initial management: the CLASSIC randomised, parallel-group, multicentre feasibility trial. *Intensive Care Med*. 2016;42(11):1695-1705. doi:10.1007/s00134-016-4500-7
40. Rhodes A, Evans LE, Alhazzani W, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock: 2016. *Crit Care Med*. 2017;45(3):486-552. doi:10.1097/ccm.0000000000002255
41. Levy MM, Evans LE, Rhodes A. The surviving sepsis campaign bundle: 2018 update. *Intensive Care Med*. 2018;44(6):925-928. doi:10.1007/s00134-018-5085-0
42. Wacharasint P, Nakada TA, Boyd JH, Russell JA, Walley KR. Normal-range blood lactate concentration in septic shock is prognostic and predictive. *Shock*. 2012;38(1):4-10. doi:10.1097/SHK.0b013e318254d41a
43. Webb AL, Kramer N, Rosario J, et al. Delta lactate (three-hour lactate minus initial lactate) prediction of in-hospital death in sepsis patients. *Cureus*. 2020;12(4):e7863.
44. Jouffroy R, Saade A, Tourtier JP, et al. Skin mottling score and capillary refill time to assess mortality of septic shock since pre-hospital setting. *Am J Emerg Med*. 2019;37(4):664-671. doi:10.1016/j.ajem.2018.07.010
45. Castro R, Kattan E, Ferri G, et al. Effects of capillary refill time-vs. lactate-targeted fluid resuscitation on regional, microcirculatory

and hypoxia-related perfusion parameters in septic shock: a randomized controlled trial. *Ann Intensive Care*. 2020;10(1):150.
46. Hernández G, Ospina-Tascón GA, Damiani LP, et al. Effect of a resuscitation strategy targeting peripheral perfusion status vs serum lactate levels on 28-day mortality among patients with septic shock: the ANDROMEDA-SHOCK randomized clinical trial. *Jama*. 2019;321(7):654-664.
47. Manzoor H, Bhatt H. *Prerenal Kidney Failure*. StatPearls Publishing LLC; 2022.
48. Montomoli J, Donati A, Ince C. Acute kidney injury and fluid resuscitation in septic patients: are we protecting the kidney? *Nephron*. 2019;143:1-4. doi:10.1159/000501748
49. Rice DM, Ratliff PD, Judd WR, Kseibi SA, Eberwein KA. Assessing the impact of CKD on outcomes in septic shock patients receiving standard vs reduced initial fluid volume. *Am J Emerg Med*. 2020;38(10):2147-2150. doi:10.1016/j.ajem.2020.07.055
50. Truong TN, Dunn AS, McCardle K, et al. Adherence to fluid resuscitation guidelines and outcomes in patients with septic shock: reassessing the "one-size-fits-all" approach. *J Crit Care*. 2019;51:94-98. doi:10.1016/j.jcrc.2019.02.006
51. Douglas IS, Alapat PM, Corl KA, et al. Fluid response evaluation in sepsis hypotension and shock: a randomized clinical trial. *Chest*. 2020;158(4):1431-1445. doi:10.1016/j.chest.2020.04.025

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Jessen MK, Andersen LW, Thomsen M-L, et al. Restrictive fluids versus standard care in adults with sepsis in the emergency department (REFACED): A multicenter, randomized feasibility trial. *Acad Emerg Med*. 2022;29:1172-1184. doi:10.1111/acem.14546

Received: 25 March 2022 | Revised: 21 June 2022 | Accepted: 21 June 2022

DOI: 10.1111/acem.14554

ORIGINAL CONTRIBUTION

Electrocardiographic findings associated with early clinical deterioration in acute pulmonary embolism

Anthony J. Weekes MD, MSc[1] | Jaron D. Raper MD[1] | Alyssa M. Thomas MD, MPH[1] | Kathryn Lupez MD[1] | Carly A. Cox MD[1] | Dasia Esener MD, MS[2] | Jeremy S. Boyd MD[3] | Jason T. Nomura MD[4] | Jillian Davison MD[5] | Patrick M. Ockerse MD[6] | Stephen Leech MD, MPH[5] | Eric Abrams MD[2] | Christopher Kelly MD[6] | Nathaniel S. O'Connell PhD[7]

[1]Department of Emergency Medicine, Atrium Health's Carolinas Medical Center (Carolinas Medical Center is the Central Site of the Pulmonary Embolism Short-term Outcomes Registry (PESCOR) consortium), Charlotte, North Carolina, USA

[2]Department of Emergency Medicine Kaiser Permanente, San Diego, California, USA

[3]Department of Emergency Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA

[4]Department of Emergency Medicine, Christiana Care, Newark, Delaware, USA

[5]Department of Emergency Medicine, Orlando Health, Orlando, Florida, USA

[6]Division of Emergency Medicine, University of Utah Health, Salt Lake City, Utah, USA

[7]Department of Biostatistics and Data Science, Wake Forest School of Medicine, Winston-Salem, North Carolina, USA

Correspondence
Anthony J. Weekes MD, MSc, 1000 Blythe Blvd, Medical Education Building, 3rd Floor, Charlotte, North Carolina 28203.
Email: anthony.weekes@atriumhealth.org

Present address
Jaron D. Raper, Department of Emergency Medicine, University of Alabama at Birmingham, Birmingham, Alabama, USA

Alyssa M. Thomas, Emergency Department, Houston Methodist Baytown Hospital, Houston, Texas, USA

Kathryn Lupez, Department of Emergency Medicine, Tufts Medical Center, Boston, Massachusetts, USA

Carly A. Cox, Emergency Medicine of Idaho, Meridian, Idaho, USA

Funding information
Agency for Healthcare Research and Quality, Grant/Award Number: R01HS025979

Abstract

Objectives: We sought to determine associations of early electrocardiogram (ECG) patterns with clinical deterioration (CD) within 5 days and with RV abnormality (abn-IRV) by echocardiography in pulmonary embolism (PE).

Methods: In this prospective, multicenter study of newly confirmed PE patients, early echocardiography and initial ECG were examined. Initial ECG patterns included lead-specific ST-segment elevation (STE) or depression (STD), T-wave inversion (TWI), supraventricular tachycardia (SVT), sinus tachycardia, and right bundle branch block as complete (cRBBB) or incomplete (iRBBB). We defined CD as respiratory failure, hypotension, dysrhythmia, cardiac arrest, escalated PE intervention, or death within

Supervising Editor: Dr. Stephen Smith

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2022 The Authors. *Academic Emergency Medicine* published by Wiley Periodicals LLC on behalf of Society for Academic Emergency Medicine.

5 days. We calculated odds ratios (ORs) for CD and abnlRV with univariate and full multivariate models in the presence of other variables.

Results: Of 1676 patients, 1629 (97.2%) had both ECG and GDE; 415/1676 (24.7%) had CD, and 529/1629 (32.4%) had abnlRV. AbnlRV had an OR for CD of 4.25 (3.35, 5.38). By univariable analysis, the absence of abnormal ECG patterns had OR for CD and abnlRV of 0.34 (0.26, 0.44; $p < 0.001$) and 0.24 (0.18, 0.31; $p < 0.001$), respectively. By multivariable analyses, one ECG pattern had a significant OR for CD: SVT 2.87 (1.66, 5.00). Significant ORS for abnlRV were: TWI V_{2-4} 4.0 (2.64, 6.12), iRBBB 2.63 (1.59, 4.38), STE aVR 2.42 (1.58, 3.74), S1-Q3-T3 2.42 (1.70, 3.47), and sinus tachycardia 1.68 (1.14, 2.49).

Conclusions: SVT was an independent predictor of CD. TWI V_{2-4}, iRBBB, STE aVR, sinus tachycardia, and S1-Q3-T3 were independent predictors of abnlRV. Finding one or more of these ECG patterns may increase considerations for performance of echocardiography to look for RV abnormalities and, if present, inform concerns for early clinical deterioration.

INTRODUCTION

The electrocardiogram (ECG) is one of the first tests performed in the emergency department (ED) on patients with symptoms of chest discomfort or dyspnea. Abnormal ECG findings can be found in cardiopulmonary conditions outside of acute coronary syndrome. Although no specific ECG pattern is diagnostic of acute PE, abrupt pulmonary arterial occlusion may provoke abnormalities of the right ventricle (RV) due to increased pressure, dilatation, or myocardial injury. Abnormal ECG patterns may also occur due to left ventricular (LV) ischemia from combinations of hypoxia or hypotension (decreased oxygen supply) and due to tachycardia (increased oxygen demand). Abnormal ECG patterns associated with RV abnormality (abnlRV) in PE include rightward axis shifts, repolarization abnormalities, ST-segment myocardial injury patterns, conduction delays, and rhythm disturbances.[1,2]

There is a lack of agreement in the literature on ECG changes in PE and their relation to abnlRV, and one study found no relationship whatsoever.[3] Prior research in this area has been limited due to retrospective analyses, heterogeneous cohorts of PE patients, small sample sizes, and ECGs not immediately performed at time of PE presentation.[4–12] Few reports involve immediate and contemporaneous evaluations of ECG and echocardiography at PE presentation. Still fewer focus on clinical outcomes that require hospital-based monitoring or support.

The primary objective of this study was to prospectively determine differences in proportions of predefined ECG patterns between ED patients with PE who experience early clinical deterioration and those who do not. The null hypothesis was there is no difference in proportions of ECG patterns between these two groups of patients. Our secondary objective was to determine the proportions of specific ECG patterns in PE patients with and without abnlRV by goal-directed echocardiography (GDE). The secondary null hypothesis was there is no difference in the proportions of ECG patterns between those with and without abnlRV. For exploratory objectives, we evaluated (1) the association of mean number of ECG patterns with classifications by PE triaging strategies, and (2) the association of specific ECG patterns with abnormal heart features (LV dysfunction, abnlRV by computed tomography [CT], and elevated biomarkers).

METHODS

Study design and setting

This study was a preplanned analysis from a previously reported registry database of ED patients with confirmed acute PE (clinicaltrials.gov NCT02883491 and NCT03915925).[13] All components of this registry were collected at six urban, academic EDs with emergency medicine residency and advanced emergency ultrasound fellowship programs. A central institutional review board (IRB) approved this federally funded multi-site study.

Study population

Adult ED patients (18 years or older) with image confirmed acute PE diagnosed within 12 hours of ED presentation were eligible for enrollment. PE was confirmed by the presence of filling defects in the pulmonary arteries with contrast-enhanced chest CT or detection of high probability on nuclear ventilation perfusion scan. We excluded patients who refused consent for clinical follow-up and those who had a ventricular paced rhythm or poor quality ECG tracings that were considered uninterpretable.

Study protocol

As previously reported, the primary goal of the PE registry was to include early imaging, ECGs, and cardiac biomarkers of ED patients with confirmed PE in the database and determine their association with clinical deterioration.[13] Research associates at each study site maintain awareness of all ED patients with positive PE studies. In addition, physicians caring for these patients routinely notified the research team and performed GDE. Clinical characteristics and demographic information were collected on each patient. ECGs were performed during the index PE ED visit as part of routine care. Emergency physicians performed GDE during the index PE ED visit, which was subsequently interpreted by faculty of the emergency ultrasound program. ECG and echocardiography were performed independently of each other. The clinical course of patients was monitored by thorough review of the electronic medical record (EMR) or direct communication with inpatient health care providers.

The 12-lead ECG was completed during the ED management phase. ECGs were performed at the discretion of the treating health care team in response to the patient's symptoms, clinical course, or PE diagnosis. Site investigators interpreted the 12-lead ECGs using predefined criteria discussed in pre-enrollment training modules. The 12-lead ECG was evaluated for each of the following previously reported patterns in PE: sinus tachycardia (≥100 beats per minute); incomplete right bundle branch block (iRBBB) if QRS complex ≥0.10 and <0.12 seconds and complete right bundle branch block (cRBBB) if QRS complex width ≥0.12 seconds; compilation of S wave in lead I, Q wave in lead III, and T wave inversion (TWI) in lead III, referred to as S1-Q3-T3 (the McGinne-White sign); TWIs in precordial leads V_{2-4}; TWI in leads II, III, and aVF; supraventricular tachycardia (SVT; including atrioventricular [AV] nodal reentry tachycardia, AV reentrant tachycardia, atrial tachycardia [unifocal or multifocal], junctional tachycardia, sinus nodal reentry tachycardia, atrial fibrillation, and atrial flutter); ST-segment depressions (STDs) in V_{4-6}; ST-segment elevation (STE) in lead aVR, left bundle branch block (LBBB); and LV hypertrophy (LVH) with associated TWI.[2,4,5,8] The ECG was considered positive for abnormal patterns if one or more of the above patterns were present. A TWI wave was defined as a negative deflection of the T wave greater than 0.5 mV below the T–P segment, which represents the isoelectric line (except in aVR or coexisting RBBB). TWI or STE associated with LBBB or RBBB patterns were considered normal. We used predefined criteria and guidelines. STE was determined 1 mm above the J point relative to the TP segment in the absence of early repolarization signs, LVH, or LBBB. ST depression was determined >0.5 mm below the J point relative to the TP segment. TWI or STE in precordial leads associated with LBBB or RBBB patterns were not considered independent ECG abnormalities for this study. In addition, we reported the total number of mutually exclusive abnormal ECG patterns in the patient's initial ECG as a discrete predictor variable.

Key outcome measures

The primary outcome was the incidence of one or more discrete clinical deterioration events that were considered PE-related complications, including death. Clinical deterioration included any episode that required hospital-based monitoring or hospital-based support or intervention. As previously reported, clinical deterioration endpoints included respiratory failure, cardiac arrest, new dysrhythmia, sustained hypotension requiring intravenous volume expansion or adrenergic medication, and escalated PE interventions.

In this prospective study, each site had institutional guidelines on PE management. We defined respiratory failure as respiratory distress associated with emergent or unplanned mechanical ventilation (intubation, non-invasive positive pressure ventilation, or cricothyrotomy). We defined cardiac arrest as new episodes of unstable cardiac rhythm or absent electrical activity with advanced cardiac life support for asystole, pulseless electrical activity, ventricular fibrillation, or unstable ventricular tachycardia. We defined new dysrhythmia as the onset of atrial fibrillation with rapid ventricular response, atrial flutter, SVT, stable ventricular tachycardia, or bradycardia that was not present at the ED presentation. We defined sustained hypotension as systolic blood pressure <90 mm Hg (or >40 mm Hg decrease from baseline) or shock index greater than 1.0 for >15 minutes with administration of >499 ml of intravenous fluids bolus for volume expansion or an infusion of norepinephrine, dopamine, or epinephrine. Escalated PE interventions included reperfusion interventions (systemic or catheter-directed thrombolysis, mechanical catheter directed aspiration, and surgical thrombectomy) and extracorporeal membrane oxygenation interventions.[13] Clinical deterioration episodes were considered as single discrete episodes if separated by 12 hours or more. Individual patients could have more than one episode of clinical deterioration within the 5-day window. We followed patients via electronic medical record review during the hospital stay and up to 30 days later, as previously published.[13]

The secondary outcome was the presence of abnlRV as determined by GDE. GDE was performed during the ED visit by emergency medicine physicians trained in GDE image acquisition and subsequently interpreted by site investigators (advanced emergency ultrasound fellowship directors). As previously reported, abnlRV included the presence of one or more of the following echocardiography findings: severe RV dilatation detected as qualitative assessment of RV:LV ratio ≥1:0 and RV apex blunting in two or more views; severe RV systolic dysfunction as qualitatively estimated by diminished longitudinal contraction of the RV (estimated tricuspid annular plane systolic excursion [TAPSE] ≤ 1.0 cm in subcostal or apical four-chamber views); and flattening or leftward bowing of the interventricular septum.[13,14] (Rationale for using TAPSE <1.0 cm as cut-off: The American Society of Echocardiography provides a normal reference range for RV measurements but no stratified classification for mild, moderate, or severe RV dilatation or systolic dysfunction.[15–17] In previous reports, using TAPSE <1.0 cm to

estimate severe RV systolic dysfunction had strong agreement with comprehensive echocardiography and strong inter- and intra-rater agreement.)

Site investigators (advanced emergency ultrasound fellowship directors) reviewed and interpreted the images according to interpretation guidelines discussed in training modules before the start of the study. GDE for abnlRV in PE was scored on a 0–3 scale, where 0 = no severe RV dilatation, 1 = severe RV dilatation only, 2 = severe RV dilatation with either septal deviation or severe RV systolic dysfunction, and 3 = severe RV dilatation with both septal deviation and severe RV systolic dysfunction. The presence of abnlRV was defined by a score of 1, 2, or 3. The frequency of ECG patterns was compared to abnlRV responses as an ordinal variable (determined by the GDE severity score). As previously reported in the main manuscript from the PE registry database, the ordinal scale showed increased odds of clinical deterioration as scores increased.[13] The GDE scale showed high intra- and inter-rater agreement and accuracy compared to comprehensive echocardiography.[14,18]

We compared the mean number of ECG patterns between assigned risk groups within a PE triaging strategy. We stratified each PE triaging strategy into low-risk and not low-risk groups. We used three PE triaging strategies: simplified pulmonary embolism severity index (sPESI), modified European Society of Cardiology (ESC), and the pulmonary embolism short-term clinical outcomes estimator (PE-SCORE).[13,19,20] The sPESI was dichotomized as low-risk (0 points) versus not low-risk (>0 points). ESC was dichotomized as low-risk versus not low-risk. Low-risk ESC was defined as sPESI = 0 points and being without abnormal cardiac biomarkers or abnormal RV by CT or echocardiography. Not low-risk PE involves abnormal vital signs, RV abnormality by cardiac biomarkers, or abnormal RV by imaging (CT or echocardiography). PE-SCORE, which has a scale of 0 to 10 total points, was stratified as low-risk (0 points), intermediate-risk (1 to 4 points), and high-risk (>4 points) groups.

We compared proportions of the specific ECG patterns to the following abnormal heart features: abnormal LV systolic function (<30%), RV abnormality as measured by CT RV:LV ratio ≥1.0, and elevated cardiac biomarkers (definitions follow). As a surrogate of myocardial stretch, brain natriuretic peptide (BNP) elevation was defined by BNP ≥90 pg/ml or N-terminal BNP ≥500 pg/ml. We used elevated troponin levels as a surrogate of myocardial injury. Toward the last quarter of the enrollment phase, the central site had an institution-wide change from troponin I to high sensitivity troponin. As previously reported in the first manuscript of the PE registry, at each participating institution, an elevated troponin I was defined as ≥0.07 ng/ml (99% upper reference limit ≥0.028 ng/ml). An elevated high sensitivity troponin was defined as ≥20 pg/ml for males and 12 pg/ml for females (99% upper reference limit range between 24 and 30 pg/ml). We used i-STAT BNP test cartridge measured in pg/mL and i-STAT cardiac troponin test cartridge measured in pg/ml (both from Abbott Point of Care, Abbott Park, IL) for troponin I or high sensitivity troponin assay.

Data analysis

We conducted univariate analyses for the primary and secondary outcomes. For the secondary outcome, GDE score (ordinal from 0–3) was dichotomized to 0 vs. >0 for ease of interpretation and presentation of results. In addition, we stratified ECG patterns by abnormal LV systolic function, CT RV:LV ratio elevation, troponin elevation, and BNP elevation as binary dependent variables. We stratified the number of abnormal ECG patterns by the three PE triaging strategy classifications as dependent variables. Descriptive statistics included means, standard deviations, median, and interquartile ranges (IQRs) for continuous variables, as well as frequencies and percentages for binary variables. For univariable statistical comparisons, we used two-sample t-tests or analysis of variance (ANOVA) for continuous variables and chi-square tests for categorical variables.

For each primary and secondary outcome, we fit a multivariable logistic regression model with all ECG variables included as primary main effects of interest, and with the following as confounders: race, gender, ethnicity, age, initial heart rate, shock index, initial respiratory rate, initial oxygen saturation, preceding episode of syncope, prior diagnosis of PE, and whether patients had one or more abnormal ECG patterns or not. In addition, in the model for primary composite outcome, we included RV:LV ratio ≥1 as determined by CT, troponin elevation, echocardiography showing abnlRV, and natriuretic peptide elevation as additional predictors. We displayed results for each full regression model in tables with odds ratios (ORs) with 95% confidence intervals (CIs). All p values <0.05 were considered statistically significant.

RESULTS

As shown in Figure 1, 1736 PE patients met inclusion criteria for the registry databases between August 2016 and November 2020. Of these, 1680 (96.8%) patients had an initial ECG performed in the ED. Four of these patients had external cardiac devices leading to ventricularly paced rhythms and were excluded from this analysis. We therefore had 1676 patients with initial native ECG tracings available for statistical analyses for the primary outcome. ECGs were performed within a mean of 1.90 ±3.38 hours of measuring initial vital signs. Enrollment numbers for the six clinical sites were: Site 1 (695, 41.5%), Site 2 (248, 14.8%), Site 3 (206, 12.3%), Site 4 (156, 9.3%), Site 5 (185, 11.0%), and Site 6 (186, 11.1%).

All ECGs and GDE were performed during the immediate ED course and before anticoagulation or escalated PE intervention. Anticoagulation or escalated PE intervention was initiated either during the ED course or immediately on arrival to the hospital floor in 97.2% (1630/1676) of the PE patients. We followed patients via EMR review during the hospital stay and up to 30 days, as previously published.[13] All patients included in this report had 5-day follow up information completed.

Demographics (Table 1) included the mean age of 59.6 ±16.6 years, 48.2% female, 65.5% Caucasian, 29.1% African

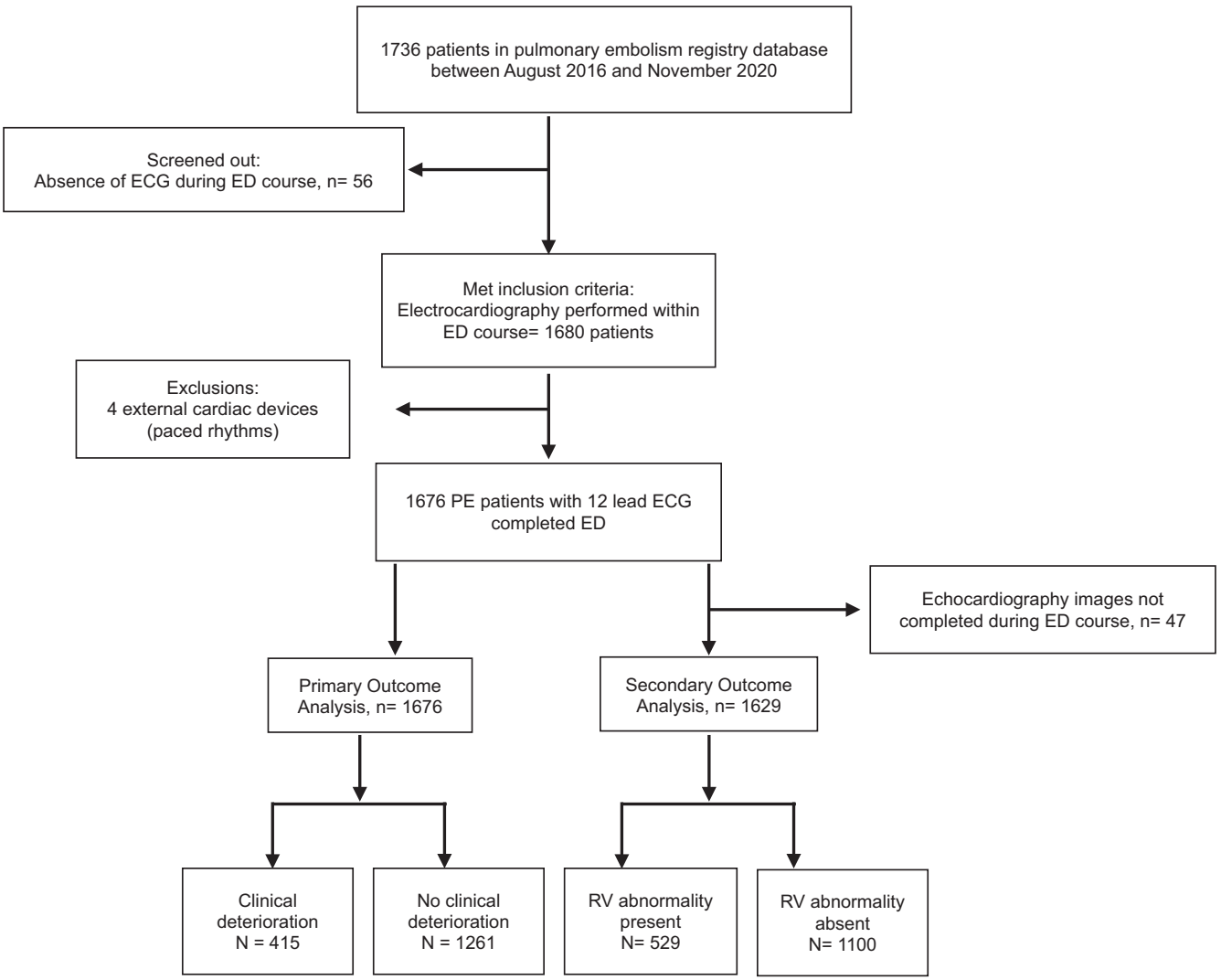


FIGURE 1 Flow diagram for patients and key outcomes

American, and 1% Asian. The mean shock index was 0.77 ±0.25 beats per minute/mm Hg, respiratory rate 19.9 ±4.6 breaths per minute, and oxygen saturation 95.4 ±4.6%. Of the 1676 patients, 160 (9.5%) had a preceding episode of syncope and 420 (25.1%) had prior diagnosis of PE or deep venous thrombosis (DVT). For each PE triaging strategy, clinical deterioration was experienced by a lower proportion of patients classified as low-risk compared to those classified as not low-risk.

Of the 1676 patients, 415 (24.7%) had one or more clinical deterioration events or hospital-based intervention within 5 days of PE diagnosis. Forty-seven of the 1676 patients did not have GDE. Of the 1629 patients with ECG and GDE, 529 (32.4%) had abnlRV by GDE. Table 1 shows numbers and proportions with abnlRV by GDE stratified by clinical deterioration, leading to an OR of 4.25 (3.35, 5.38, $p < 0.0001$).

Table 2 shows the prevalence of each specific ECG pattern. By univariate analysis, the absence of abnormal ECG patterns had an OR for clinical deterioration of 0.34 (0.25, 0.44; $p < 0.001$). The most common ECG patterns were sinus tachycardia (38.9%), S1-Q3-T3 pattern (16.3%), incomplete or complete RBBB (15.1%), and TWI in V_{2-4} (14%). In comparison to patients without the specific ECG pattern, those with the following patterns were significantly more likely to experience clinical deterioration: sinus tachycardia, TWI in V_{2-4}, STE in aVR, STE in V_1, SVT, TWI in II and III and aVF, and STD in V_{4-6}. LVH with TWI (3%) and LBBB were uncommon and not significantly different between outcome groups.

The results of full model multivariable logistic regression analysis are shown in Table 3. SVT was the only ECG pattern that was an independent predictor of CD. Absence of abnormal ECG patterns was not an independent predictor. Abnormal RV by GDE, elevated troponin measurements, elevated natriuretic peptide measurements, CT derived RV:LV ratio ≥1.0, shock index >1.0 and preceding episode of syncope were independent predictors of clinical deterioration. Table 4 shows univariable analysis results for all ECG patterns, except LBBB and LVH, had significant ORs for abnlRV by GDE (our secondary outcome). Sinus tachycardia, TWI in II III, aVF and V_{2-4}, S1–Q3-T3, iRBBB, and STE in aVR had ORs with the lower boundary of 95% CI above 1.0. For presence and absence of abnlRV by

TABLE 1 Patient characteristics

	No clinical deterioration (*N* = 1261)	Clinical deterioration (*N* = 415)	Overall (*N* = 1676)
Gender			
Female	601 (47.7%)	206 (49.6%)	807 (48.2%)
Male	660 (52.3%)	209 (50.4%)	869 (51.8%)
Race			
Caucasian	819 (64.9%)	279 (67.2%)	1098 (65.5%)
African/American	376 (29.8%)	112 (27.0%)	488 (29.1%)
American Indian/ Alaskan Native	10 (0.8%)	2 (0.5%)	12 (0.7%)
Asian	12 (1.0%)	4 (1.0%)	16 (1.0%)
Ethnicity			
Hispanic	91 (7.2%)	25 (6.0%)	116 (6.9%)
Mean Age, years			
Mean	58.76 ± 16.8	62.2 ± 15.7	59.6 ± 16.6
Systolic blood pressure, mm Hg			
Mean	135.5 ± 23.1	122.2 ± 25.5	132.2 ± 24.4
Heart rate, beats per minute			
Mean	95.4 ± 20.05	106.2 ± 23.6	98.1 ± 21.5
Shock index			
Mean	0.73 ± 0.22	0.91 ± 0.29	0.77 ± 0.25
Respiratory rate, breaths per minute			
Mean	19.6 ± 4.2	21.0 ± 5.54	19.9 ± 4.60
Oxygen saturation, %			
Mean	95.8 ± 4.2	94.2 ± 5.6	95.4 ± 4.63
Cancer			
Present	302 (23.9%)	109 (26.3%)	411 (24.5%)
COPD			
Present	176 (14.0%)	75 (18.1%)	251 (15.0%)
Previous PE/DVT			
Present	313 (24.8%)	107 (25.8%)	420 (25.1%)
Abnormal RV by GDE (*N* = 1629)			
Present	296 (24.2%)	233 (57.5%)	529 (43.2%)
Absent	928 (75.8%)	172 (42.5%)	1100 (89.9%)
PE-SCORE			
Low-risk (0 points)	294 (23.3%)	15 (3.6%)	309 (18.4%)
Intermediate-risk (1–4 points)	864 (68.5%)	295 (71.1%)	1159 (69.2%)
High-risk (>4 points)	26 (2.1%)	75 (18.1%)	101 (6.0%)
Missing	77 (6.1%)	30 (7.2%)	107 (6.4%)
ESC			
Low-risk	166 (13.2%)	5 (1.2%)	171 (10.2%)
Not low-risk	1095 (86.8%)	410(98.8%)	1505 (89.8%)
Missing	0(0%)	0 (0%)	0 (0%)
sPESI			
Low-risk	518 (41.1%)	71 (17.1%)	589 (35.1%)
Not low-risk	743 (58.9%)	344 (82.9%)	1087 (64.9%)
Missing	0 (0%)	0 (0%)	0 (0%)

Abbreviations: COPD, chronic obstructive pulmonary disease; DVT, deep venous thrombosis; ESC, European Society of Cardiology PE dichotomized into low-risk versus not low-risk classifications; GDE, goal-directed echocardiography; PE, pulmonary embolism; PE-SCORE, pulmonary embolism short-term clinical outcomes risk estimator; RV, right ventricle; sPESI, simplified Pulmonary Embolism Severity Index.

GDE, the mean number of abnormal ECG patterns was 2.20 ± 1.69 and 0.93 ± 1.13, respectively ($p < 0.0001$). The absence of abnormal ECG patterns had an OR for abnlRV by GDE of 0.24 (0.18, 0.31; $p < 0.001$). The results of full multivariable logistic regression analysis are shown in Table 5. Independent ECG predictors of abnlRV in order of ORs (high to low) were: STE in aVR, TWI V_{2-4}, incomplete RBBB, S1-Q3-T3, and sinus tachycardia. LVH with associated TWI lowered the odds of abnlRV. With multivariable analysis, the absence of any of the ECG patterns was not an independent predictor of clinical deterioration. Clinical predictors were age, shock index, initial heart rate, initial respiratory rate, and preceding episode of syncope.

Of the 1676 patients, 1090 (65.0%) had at least one of the abnormal ECG patterns. The mean number of abnormal ECG patterns was 1.33 ± 1.4. For patients with and without the primary outcome, the mean numbers of abnormal ECG patterns were 1.97 ± 1.65 and 1.20 ± 1.30, respectively (Table 2). Supporting Information Table S1, which is available online, shows patients with low-risk classifications by all three PE triaging strategies had a lower number of abnormal ECG patterns compared to those in the higher risk classifications.

Supporting Information Tables S2–S9 show results of analyses of specific ECG patterns stratified by LV systolic dysfunction, RV dilatation by CT, and elevated cardiac biomarkers. Of 1629 with GDE, 1620 had a report on LV systolic function, with 127 (7.8%) having severe LV dysfunction. SVT, LBBB, and LVH with TWI were significantly associated with severe LV systolic dysfunction by univariable analysis (Supporting Information Table S2), whereas only SVT and LBBB were significant by multivariable analysis (Supporting Information Table S3). Supporting Information Table S4 shows 450 of 1659 patients (27.1%) had troponin elevations. All ECG patterns, except SVT, LBBB, and LVH with TWI, were associated with troponin elevation by univariable analysis. However, Supporting Information Table S5 shows sinus tachycardia, S1-Q3-T3, TWI V_{2-4}, and STE aVR were independent predictors of troponin elevations by multivariable analysis. Supporting Information Table S6 shows 642 of 1608 patients (39.9%) with BNP measurements had BNP elevations. All ECG patterns, except iRBBB, were associated with natriuretic peptide elevation by univariable analysis. Supporting Information Table S7 shows S1-Q3-T3, STE V_1, TWI V_{2-4}, and SVT were significant by multivariable analysis. Supporting Information

TABLE 2 ECG patterns stratified by primary clinical outcomes[a]

	No clinical deterioration (*N* = 1261)	Clinical deterioration (*N* = 415)	Overall (*N* = 1676)	OR	*p*-Value
Mean number of abnormal ECG patterns	1.20 ± 1.31	1.99 ± 1.67	1.34 ± 1.44		<0.0001
Number of abnormal ECG patterns					
Zero abnormal patterns	509 (40.4%)	77 (18.6%)	586 (35.0%)	0.34 (0.26, 0.44)	<0.0001
>0 abnormal patterns	751 (59.6%)	338 (81.4%)	1089 (65.0%)		
Missing	1 (0.1%)	0 (0%)	1 (0.1%)		
Complete RBBB					
Absent	1180 (93.6%)	369 (88.9%)	1549 (92.4%)	1.82 (1.24, 2.66)	0.003
Present	81 (6.4%)	46 (11.1%)	127 (7.6%)		
Incomplete RBBB					
Absent	1175 (93.2%)	376 (90.6%)	1551 (92.5%)	1.42 (0.95, 2.11)	0.104
Present	86 (6.8%)	39 (9.4%)	125 (7.5%)		
Sinus tachycardia					
Absent	835 (66.2%)	189 (45.5%)	1024 (61.1%)	2.34 (1.87, 2.94)	<0.001
Present	426 (33.8%)	226 (54.5%)	652 (38.9%)		
S1-Q3-T3 pattern					
Absent	1080 (85.6%)	322 (77.6%)	1402 (83.7%)	1.72 (1.30, 2.28)	<0.001
Present	181 (14.4%)	93 (22.4%)	274 (16.3%)		
STE V_1					
Absent	1162 (92.1%)	362 (87.2%)	1524 (90.9%)	1.72 (1.21, 2.45)	0.003
Present	99 (7.9%)	53 (12.8%)	152 (9.1%)		
T wave inversions V_{2-4}[b]					
Absent	1112 (88.2%)	330 (79.5%)	1442 (86.0%)	1.92 (1.43, 2.58)	<0.001
Present	149 (11.8%)	85 (20.5%)	234 (14.0%)		
T wave inversions II, III, aVF					
Absent	1153 (91.4%)	354 (85.3%)	1507 (89.9%)	1.84 (1.32, 2.57)	<0.001
Present	108 (8.6%)	61 (14.7%)	169 (10.1%)		
ST-segment depression V_{4-6}					
Absent	1183 (93.8%)	354 (85.3%)	1537 (91.7%)	2.61 (1.83, 3.73)	<0.001
Present	78 (6.2%)	61 (14.7%)	139 (8.3%)		
STE aVR					
Absent	1146 (90.9%)	324 (78.1%)	1470 (87.7%)	2.82 (2.09, 3.82)	<0.001
Present	114 (9.0%)	91 (21.9%)	205 (12.2%)		
Missing	1 (0.1%)	0 (0%)	1 (0.1%)		
SVT (including atrial fibrillation with rapid ventricular response)					
Absent	1218 (96.6%)	361 (87.0%)	1579 (94.2%)	4.24 (2.80, 6.43)	<0.001
Present	43 (3.4%)	54 (13.0%)	97 (5.8%)		
LBBB associated with TWI					
Absent	1238 (98.2%)	412 (99.3%)	1650 (98.4%)	0.39 (0.11, 1.31)	0.179
Present	23 (1.8%)	3 (0.7%)	26 (1.6%)		
LVH with TWI					
Absent	1233 (97.8%)	402 (96.9%)	1635 (97.6%)	1.42 (0.73, 2.78)	0.39
Present	28 (2.2%)	13 (3.1%)	41 (2.4%)		

Abbreviations: ECG, electrocardiography; LBBB, left bundle branch block; LVH, left ventricular hypertrophy; OR, odds ratio; STE, ST-segment elevation; SVT, supraventricular tachycardia; RBBB, right bundle branch block; TWI, T-wave inversion (0.5 mV negative deflection).

[a]The total number of ECG patterns present was greater than the total number of unique patients because any single patient could have more than one of the ECG patterns.

[b]For those with either iRBBB/cRBBB without TWI in V_{2-4} (n = 168) versus those with TWI V_{2-4} (n = 82), 56 (33.3%) and 29 (35.2%) had clinical deterioration, respectively. The OR for iRBBB without TWI was 0.91 (0.52, 1.59; p = 0.75).

TABLE 3 Multivariable analysis stratified by clinical deterioration within 5 days

	Primary composite outcome (one or more clinical deterioration events within 5 days)		
Predictors	**OR**	**CI**	***p-Value***
(Intercept)	0.56	0.03–13.59	0.722
Complete RBBB	1.06	0.64–1.73	0.825
Incomplete RBBB	1.03	0.61–1.69	0.918
Sinus tachycardia	1.43	0.96–2.15	0.081
S1-Q3-T3 pattern	0.85	0.58–1.22	0.375
ST elevation V_1	0.90	0.55–1.45	0.670
T wave inversions V2-4	1.01	0.66–1.52	0.981
T wave inversions II, III, aVF	1.08	0.67–1.71	0.754
ST depression in V 4–6	1.14	0.69–1.85	0.614
STE aVR	1.16	0.75–1.78	0.494
SVT	2.87	1.66–5.00	<0.001
LBBB with associated TWI	0.21	0.03–0.86	0.060
LVH with associated TWI	1.49	0.60–3.47	0.369
Echocardiography showing abnormal RV	1.68	1.17–2.41	0.005
Elevated troponin level	1.47	1.08–2.00	0.014
RV:LV ratio 1.0 or greater by CT	1.49	1.07–2.08	0.019
Elevated natriuretic peptide level	1.26	0.93–1.70	0.130
Male	1.01	0.77–1.32	0.952
African American/ Black	0.66	0.36–1.23	0.182
White	0.81	0.46–1.48	0.486
Ethnicity	1.00	1.00–1.00	0.322
Age	1.01	1.00–1.02	0.025
Initial heart rate	0.99	0.97–1.00	0.008
Initial shock index	13.07	5.69–30.65	<0.001
Initial respiratory rate	1.01	0.98–1.04	0.534
Initial oxygen saturation on room air	0.97	0.95–1.00	0.060
Preceding episode of syncope	1.75	1.14–2.66	0.010
Prior history of PE or DVT	1.02	0.75–1.39	0.885
No abnormal ECG pattern	0.84	0.54–1.32	0.460
Observations	1472		
R^2 Tjur	0.197		

Abbreviations: CI, confidence interval; DVT, deep venous thrombosis; ECG, electrocardiogram; LBBB, left bundle branch block; LVH, left ventricular hypertrophy; OR, odds ratio; PE, pulmonary embolism; STE, ST-segment elevation; SVT, supraventricular tachycardia; TWI = T-wave inversion (0.5 mV negative deflection).

Table S8 shows elevated CT RV:LV ratios in 552 of 1640 patients (33.7%) who had CT performed. All ECG patterns, except SVT, LBBB, and LVH with TWI, were associated with elevated CT RV:LV ratio by univariable analysis. Supporting Information Table S9 shows complete and incomplete RBBB, S1-Q3-T3, TWI V_{2-4}, STE aVR, and SVT were significant by multivariable analysis. Finally, Supporting Information Table S10 shows the prognostic metrics of the following categorical variables: abnormal ECG patterns, elevated troponin, elevated natriuretic peptide, initial shock index (>1.0), CT RV:LV ratio, hypotension (< 100 mm Hg), hypoxia (<92%), and preceding syncope. Predictor positive likelihood ratios ordered high to low were: hypotension, initial shock index, preceding syncope, hypoxia, elevated troponin, increased RV:LV ratio, elevated natriuretic peptide, and abnormal ECG pattern. The number of abnormal ECG patterns and the CT RV:LV ratio demonstrated the highest negative predictive values, followed by elevated natriuretic peptide, initial shock index, elevated troponin, hypotension, hypoxia, and preceding syncope.

LIMITATIONS

This study had several limitations. First, preexisting ECG findings were not reviewed to determine if index PE-associated ECG abnormalities were acute or preexisting. Second, there are multiple causes of tachycardia that are unrelated to and not directly provoked by the pathophysiology of PE. Third, TWIs may be associated with normal variants and can be associated with RBBB. Furthermore, the RBBB pattern can exist in the asymptomatic population without PE.[21] Fourth, we reported predetermined ECG patterns that were fully recorded and part of the medical record. It is plausible that ECG abnormalities were in evolution. STE was defined as the ST segment being ≥1 mm above the isoelectric line (TP segment). However, use of the J point relative to PQ junction, instead of TP segment, is currently advised for determining ST depression.[22]

Several investigators witnessed important but transient changes to ECG patterns in higher acuity acute PE patients that were not fully recorded or available in the EMR for interpretation or analysis. Our report focused on the initial ECG in the ED; however, there were several instances in patients with abnormal RV by GDE where conduction, TWI, or ST-segment abnormalities were transient. We observed several instances of normalization of initial abnormal ECG patterns and improvements to RV dilatation or systolic function (determined by serial echocardiography) after administration of thrombolytic agents. Further supporting evidence for this observation is provided by Choi et al., who showed not only an association between TWI and abnlRV, but also found that reversals of T wave abnormalities were temporally associated with resolution of an abnlRV.[23] Other ECG patterns pertinent to PE, such as isolated TWI in III and V_1, were not included in this report because of lack of prognostic significance. In our database, the TWI III and V_1 pattern was present in 17.6% (73/415) and 17.4% (219/1261) of those with and without clinical deterioration, respectively. Finally, we did not examine the

TABLE 4 Univariable analysis of ECG patterns stratified by echocardiographic RV abnormalities

	Echocardiography showing RV abnormalities (47 missing)			
	No (*N* = 1100)	**Yes (*N* = 529)**	**OR**	***p*-Value**
Mean number of ECG patterns (standard deviation)	0.93 (1.13)	2.20 (1.69)		<0.0001
Number of abnormal ECG patterns				
None	485 (44.1%)	84 (15.9%)	0.24 (0.18, 0.31)	<0.001
One or more	614 (55.8%)	445 (84.1%)		
Missing	1 (0.1%)	0 (0%)		
Complete RBBB				
Absent	1043 (94.8%)	461 (87.1%)	2.70 (1.87, 3.9)	<0.001
Present	57 (5.2%)	68 (12.9%)		
Incomplete RBBB				
Absent	1046 (95.1%)	458 (86.6%)	2.95 (2.04, 4.26)	<0.001
Present	54 (4.9%)	71 (13.4%)		
Sinus tachycardia				
Absent	759 (69.0%)	241 (45.6%)	2.67 (2.15, 3.30)	<0.001
Present	341 (31.0%)	288 (54.4%)		
S1-Q3-T3 pattern				
Absent	991 (90.1%)	371 (70.1%)	3.87 (2.95, 5.08)	<0.001
Present	109 (9.9%)	158 (29.9%)		
STE V_1				
Absent	1026 (93.3%)	454 (85.8%)	2.29 (1.63, 3.22)	<0.001
Present	74 (6.7%)	75 (14.2%)		
T wave inversions V_{2-4}				
Absent	1019 (92.6%)	381 (72.0%)	4.89 (3.64, 6.57)	<0.001
Present	81(7.4%)	148 (28.0%)		
T wave inversions II, III, aVF				
Absent	1029 (93.5%)	434 (82.0%)	3.17 (2.29, 4.40)	<0.001
Present	71 (6.5%)	95 (18.0%)		
ST-segment depression V_{4-6}				
Absent	1038 (94.4%)	453 (85.6%)	2.81 (1.98, 4.00)	<0.001
Present	62 (5.6%)	76 (14.4%)		
STE aVR				
Absent	1023 (93.0%)	404 (76.4%)	4.16 (3.06, 5.55)	<0.001
Present	76 (6.9%)	125 (23.6%)		
Missing	1 (0.1%)	0 (0%)		
SVT (including atrial fibrillation with rapid ventricular response)				
Absent	1049 (95.4%)	485 (91.7%)	1.87 (1.23, 2.83)	0.004
Present	51 (4.6%)	44 (8.3%)		
LBBB associated with TWI				
Absent	1079 (98.1%)	524 (99.1%)	0.49 (0.18, 1.31)	0.214
Present	21 (1.9%)	5 (0.9%)		
LVH with TWI				
Absent	1070 (97.3%)	518 (97.9%)	0.76 (0.38, 1.52)	0.54
Present	30 (2.7%)	11 (2.1%)		

Abbreviations: ECG, electrocardiogram; LBBB, left bundle branch block; LVH, left ventricular hypertrophy; OR, odds ratio; RBBB, right bundle branch block; RV, right ventricle; STE, ST-segment elevation; SVT, supraventricular tachycardia; TWI, T-wave inversion (0.5 mV negative deflection).

TABLE 5 Full multivariable analysis model stratified by RV abnormalities (by echocardiography)

	Echocardiography with abnlRV		
Predictors	**OR**	**CI**	***p*-Value**
(Intercept)	30.78	1.09–931.83	0.046
Complete RBBB	1.23	0.74–2.04	0.422
Incomplete RBBB	2.63	1.59–4.38	<0.001
Sinus tachycardia	1.68	1.14–2.49	0.009
S1-Q3-T3 pattern	2.42	1.70–3.47	<0.001
STE V_1	1.36	0.83–2.22	0.215
T wave inversions V_{2-4}	4.00	2.64–6.12	<0.001
T wave inversions II, III, aVF	1.26	0.78–2.02	0.349
ST depression in V_{4-6}	0.79	0.47–1.31	0.364
STE aVR	2.42	1.58–3.74	<0.001
SVT	1.25	0.71–2.18	0.441
LBBB with associated TWI	0.30	0.07–0.97	0.069
LVH with associated TWI	0.33	0.11–0.86	0.029
Male	0.97	0.75–1.27	0.848
African American/Black	1.18	0.64–2.25	0.605
White	1.05	0.58–1.95	0.883
Ethnicity	1.00	1.00–1.00	0.579
Age	1.02	1.01–1.03	<0.001
Initial heart rate	1.00	0.99–1.01	0.664
Initial shock index	5.45	2.36–12.76	<0.001
Initial respiratory rate	1.04	1.02–1.08	0.003
Initial oxygen saturation on room air	0.92	0.89–0.94	<0.001
Preceding episode of syncope	3.09	2.00–4.82	<0.001
Prior history of PE or DVT	1.07	0.79–1.44	0.651
No abnormal ECG pattern	1.08	0.70–1.68	0.722
Observations	1472		
R^2 Tjur	0.284		

Abbreviations: CI, confidence interval; DVT, deep venous thrombosis; ECG, electrocardiogram; LBBB, left bundle branch block; LVH, left ventricular hypertrophy; OR, odds ratio; PE, pulmonary embolism; RBBB, right bundle branch block; STE, ST-segment elevation; SVT, supraventricular tachycardia; TWI, T-wave inversion (0.5 mV negative deflection).

prevalence of specific ECG patterns in patients without acute PE to compare with prevalence in patients with PE.

DISCUSSION

We prospectively evaluated PE patients in the ED to report on the associations of multiple specific ECG patterns for clinical deterioration and abnlRV by echocardiography using univariable and multivariable analyses. ECGs were performed early and contemporaneously with GDE assessments for abnlRV during the ED work-up. Abnormal RV (as determined by GDE) had an OR of 4.25 (3.35, 5.38) for clinical deterioration. The absence of abnormal ECG patterns on initial early ECG had ORs of 0.34 (0.26, 0.44) and 0.24 (0.18, 0.31) for clinical deterioration and abnlRV by GDE, respectively. However, on multivariable analysis, absence of any ECG patterns did not qualify as an independent predictor of clinical deterioration or abnlRV.

We found all abnormal ECG patterns had statistically significant increased ORS above unity for clinical deterioration within 5 days of PE diagnosis, except for LBBB and LVH with TWI. Of the ECG patterns, only SVT was significant as an independent predictor for clinical deterioration. For abnlRV by echocardiography, associations were significant for most ECG patterns. Sinus tachycardia, STE in aVR, incomplete RBBB, S1-Q3-T3, and TWI V_{2-4} were significant as independent predictors for abnlRV by GDE. Patients with acute pulmonary embolism may have one or more of the specific ECG abnormalities. In patients with confirmed acute pulmonary embolism, one or more of these ECG patterns should increase considerations for performance of echocardiography to look for RV abnormalities and, if present, increase concerns for early clinical deterioration. By univariable analysis, we found the absence of any of the abnormal ECG patterns was associated with lower probability of clinical deterioration and abnlRV by GDE. However, with full multivariable analysis with other ECG and non-ECG variables, the number of abnormal ECG patterns was not an independent predictor of clinical deterioration or abnlRV by echocardiography.

Most previous studies have been retrospective analyses with ECGs completed 10 to 72 hours removed from the time of PE diagnosis. Strengths of our study include its prospective design, a large database of patients representing the full spectrum of PE severity, and early timing of ECG relative to ED diagnosis of PE.

Clinical deterioration

Previous reports have shown several ECG patterns (STE, STsegment depression, S1-Q3-T3, RBBB, TWI, and SVT) were associated with clinical deterioration.[2,12] Similarly, by univariable analysis, our prospective study showed significant associations of these abnormal ECG patterns with clinical deterioration within 5 days. In addition, we noted the following ECG patterns were also associated with increased risk for clinical deterioration: STE in V_1; TWI in leads II, III, and aVF; ST depression in V_4–V_6; STE in aVR; and SVT (Table 2). However, SVT was the only ECG pattern determined to be an independent predictor from multivariable analyses in the presence of other clinical variables.

Abnormal RV by echocardiography

In our study, early echocardiography findings of abnormal RV within the ED course were significantly associated with clinical

deterioration within 5 days. In turn, abnormal ECG patterns (sinus tachycardia, STE in aVR, incomplete RBBB, S1-Q3-T3, and TWI V_{2-4}) were independent predictors of abnlRV by echocardiography even in the presence of other clinical variables. By multivariable analyses, the number of abnormal ECG patterns was not an independent predictor of clinical deterioration or abnl RV by echocardiography.

Reports on the associations between ECG and abnlRV by echocardiography in PE patients are mixed, with important differences in timing of ECG performance studies, patient cohorts, and the ECG patterns selected. A report by Vanni et al. studied ECG and echocardiography performed within 1 hour of PE diagnosis and found that patients with RBBB, precordial repolarization abnormalities, or the S1-Q3-T3 pattern complex were twice as likely to have echocardiographic signs of abnlRV than patients without any of the evaluated ECG patterns.[11] Punukollu et al. showed S1-Q3-T3 and TWI in V_{1-3} were more likely in abnlRV, especially if echocardiography was performed at admission or within 24 hours of hospitalization.[9] Daniel et al. showed that having multiple abnormal ECG patterns offered low sensitivity but high specificity for echocardiographic evidence of severe abnlRV.[4] In contrast, another report showed having a few abnormal ECG patterns had moderate sensitivity and moderate specificity for abnormal RV.[10] In one study, TWI V_{2-4} was six times more frequent in PE patients with abnlRV than those without abnlRV.[6] Another report showed TWI in V_{1-3}, tachycardia, and S wave in lead I were strong predictors of abnlRV by echocardiography, CT, or cardiac biomarkers.[24] In contrast, a retrospective study of 289 patients by Stein et al. did not find ECG to be useful in predicting the presence or absence of RV dilatation.[3] In our study, the independent predictors of abnlRV by echocardiography were sinus tachycardia, S1-Q3-T3, TWI V_{2-4}, and STE in aVR. Our study findings suggest the presence of one or more of these ECG patterns may be associated with the presence of concurrent abnlRV in suspected or confirmed acute PE patients, which in turn, warns providers the PE patient under their care has an increased risk of clinical deterioration within 5 days. The presence of an acute abnlRV may be determined by reviewing recently performed imaging (CT or echocardiography).

Reports on the prevalence of specific PE-associated ECG abnormalities vary. In a meta-analysis of 45 studies that involved 8209 PE patients, the most prevalent abnormal ECG patterns were sinus tachycardia (38%), TWI V_{1-4} (38%), and STE in aVR (36%).[12] Similarly, sinus tachycardia was the most prevalent ECG pattern in our study. Other similarities and differences between our results and previous studies are as follows. First, the prevalence of RBBB (complete or incomplete) in other studies has ranged from 4.8% to 32.0%.[2] In our study, complete RBBB and incomplete RBBB were present in 7.6% and 7.5% of 1676 patients, respectively. Second, the prevalence of precordial TWI in other studies has ranged from 14% to 59%.[3,4,6,11,16] In our study, TWI in V_2–V_4 was present in 11% of our patients. Third, the prevalence of TWI in inferior leads was as high as 48% of patients in other studies,[16] whereas only 7% of our patients had inferior TWI. Fourth, previous reports had high prevalence of STE in aVR (34%–45%) in patients with more severe PE presentation, including hypotension, need for thrombolysis or adrenergic medication use.[5,7] Although we found STE in aVR in less PE patients (12%) than previous studies, STE aVR had ORs of 2.82 for 5-day clinical deterioration and 4.16 for abnlRV in our study. We also found a prevalence of 18% for the S1-Q3-T3 pattern, whereas previous studies have reported a prevalence between 8.5% and 33.0%, and as high as 53% in PE patients with clinical shock. Like previous ECG scoring systems, sinus tachycardia and anterior precordial TWI (V_{2-4}) were strong predictors of clinical deterioration.

By univariate analysis, mean number of abnormal ECG patterns was significantly greater in those with clinical deterioration and abnl RV by echocardiography versus those without. The mean number of abnormal ECG patterns was lower in those with low-risk classification by any of the PE triaging strategies.

CONCLUSIONS

Several ECG patterns are significantly more frequent in those with clinical deterioration versus those without. SVT was the independent predictor of subsequent clinical deterioration. Sinus tachycardia, STE in aVR, incomplete RBBB, S1-Q3-T3, and TWI V_{2-4} were independent predictors of abnlRV by echocardiography. The absence of any of the abnormal ECG patterns was associated with reduced risk for both clinical deterioration and abnlRV by univariable analysis but not by multivariable analysis. By univariate analyses, higher mean number of abnormal ECG patterns was associated with presence of clinical deterioration, RV abnormality by echocardiography, and PE triaging strategy classification as "not low-risk" versus the groups without these characteristics.

FUNDING INFORMATION

This was supported by the Agency for Healthcare Research and Quality [grant number R01HS025979]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality.

AUTHOR CONTRIBUTIONS

A.J.W. designed the study. J.D.R., A.T., C.C., K.L., and A.J.W. supervised the conduct of the trial and data collection. All PESCOR investigators enrolled patients. N.O. and A.J.W. provided statistical analysis and interpretation of the data. A.J.W. managed the data, including quality control. A.J.W., J.D.R., K.L. drafted the manuscript. All authors contributed substantially to article revision for important intellectual content. A.J.W. takes responsibility for the paper as a whole.

CONFLICT OF INTEREST

All authors report no conflicts of interest.

ORCID

Anthony J. Weekes https://orcid.org/0000-0001-9309-7910

REFERENCES

1. Digby GC, Kukla P, Zhan ZQ, et al. The value of electrocardiographic abnormalities in the prognosis of pulmonary embolism: a consensus paper. *Ann Noninvasive Electrocardiol*. 2015;20(3):207-223.
2. Shopp JD, Stewart LK, Emmett TW, Kline JA. Findings from 12-lead electrocardiography that predict circulatory shock from pulmonary embolism: systematic review and meta-analysis. *Acad Emerg Med*. 2015;22(10):1127-1137.
3. Stein PD, Matta F, Sabra MJ, et al. Relation of electrocardiographic changes in pulmonary embolism to right ventricular enlargement. *Am J Cardiol*. 2013;112(12):1958-1961.
4. Daniel KR, Courtney DM, Kline JA. Assessment of cardiac stress from massive pulmonary embolism with 12-lead ECG. *Chest*. 2001;120(2):474-481.
5. Janata K, Hochtl T, Wenzel C, et al. The role of ST-segment elevation in lead aVR in the risk assessment of patients with acute pulmonary embolism. *Clin Res Cardiol*. 2012;101(5):329-337.
6. Kostrubiec M, Hrynkiewicz A, Pedowska-Wloszek J, et al. Is it possible to use standard electrocardiography for risk assessment of patients with pulmonary embolism? *Kardiol Pol*. 2009;67(7):744-750.
7. Kukla P, Dlugopolski R, Krupa E, et al. The prognostic value of ST-segment elevation in the lead aVR in patients with acute pulmonary embolism. *Kardiol Pol*. 2011;69(7):649-654.
8. Kukla P, McIntyre WF, Fijorek K, et al. Electrocardiographic abnormalities in patients with acute pulmonary embolism complicated by cardiogenic shock. *Am J Emerg Med*. 2014;32(6):507-510.
9. Punukollu G, Gowda RM, Vasavada BC, Khan IA. Role of electrocardiography in identifying right ventricular dysfunction in acute pulmonary embolism. *Am J Cardiol*. 2005;96(3):450-452.
10. Toosi MS, Merlino JD, Leeper KV. Electrocardiographic score and short-term outcomes of acute pulmonary embolism. *Am J Cardiol*. 2007;100(7):1172-1176.
11. Vanni S, Polidori G, Vergara R, et al. Prognostic value of ECG among patients with acute pulmonary embolism and normal blood pressure. *Am J Med*. 2009;122(3):257-264.
12. Qaddoura A, Digby GC, Kabali C, Kukla P, Zhan Z-Q, Baranchuk AM. The value of electrocardiography in prognosticating clinical deterioration and mortality in acute pulmonary embolism: a systematic review and meta-analysis [internet]. *Clin Cardiol*. 2017;40(10):814-824. doi:10.1002/clc.22742
13. Weekes AJ, Raper JD, Lupez K, et al. Development and validation of a prognostic tool: pulmonary embolism short-term clinical outcomes risk estimation (PE-SCORE). *PLoS One*. 2021;16(11):e0260036.
14. Weekes AJ, Thacker G, Troha D, et al. Diagnostic accuracy of right ventricular dysfunction markers in normotensive emergency department patients with acute pulmonary embolism. *Ann Emerg Med*. 2016;68(3):277-291.
15. Rudski LG, Lai WW, Afilalo J, et al. Guidelines for the echocardiographic assessment of the right heart in adults: a report from the American Society of Echocardiography endorsed by the European Association of Echocardiography, a registered branch of the European Society of Cardiology, and the Canadian Society of Echocardiography. *J Am Soc Echocardiogr*. 2010;23(7):685-713. quiz 786–8.
16. Lang RM, Badano LP, Mor-Avi V, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: an update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *Eur Heart J Cardiovasc Imaging*. 2015;16(3):233-270.
17. Addetia K, Miyoshi T, Citro R, et al. Two-dimensional echocardiographic right ventricular size and systolic function measurements stratified by sex, age, and ethnicity: results of the world Alliance of societies of echocardiography study [internet]. *J Am Soc Echocardiogr*. 2021;34:1148-1157.e1. doi:10.1016/j.echo.2021.06.013
18. Weekes AJ, Oh L, Thacker G, et al. Interobserver and Intraobserver agreement on qualitative assessments of right ventricular dysfunction with echocardiography in patients with pulmonary embolism. *J Ultrasound Med*. 2016;35(10):2113-2120.
19. Jimenez D, Aujesky D, Moores L, et al. Simplification of the pulmonary embolism severity index for prognostication in patients with acute symptomatic pulmonary embolism. *Arch Intern Med*. 2010;170(15):1383-1389.
20. Konstantinides SV, Meyer G, Becattini C, et al. 2019 ESC guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS). *Eur Heart J*. 2020;41(4):543-603.
21. Bussink BE, Holst AG, Jespersen L, Deckers JW, Jensen GB, Prescott E. Right bundle branch block: prevalence, risk factors, and outcome in the general population: results from the Copenhagen City heart study. *Eur Heart J*. 2013;34(2):138-146.
22. Kashou AH, Basit H, Malik A. ST Segment. *StatPearls*. StatPearls Publishing; 2021.
23. Choi B-Y, Park D-G. Normalization of negative T-wave on electrocardiography and right ventricular dysfunction in patients with an acute pulmonary embolism. *Korean J Intern Med*. 2012;27(1):53-59.
24. Hariharan P, Dudzinski DM, Okechukwu I, Takayesu JK, Chang Y, Kabrhel C. Association between electrocardiographic findings, right heart strain, and short-term adverse clinical events in patients with acute pulmonary embolism [internet]. *Clin Cardiol*. 2015;38(4):236-242. doi:10.1002/clc.22383

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Weekes AJ, Raper JD, Thomas AM, et al. Electrocardiographic findings associated with early clinical deterioration in acute pulmonary embolism. *Acad Emerg Med*. 2022;29:1185-1196. doi: 10.1111/acem.14554

Received: 9 May 2022 | Revised: 12 July 2022 | Accepted: 13 July 2022

DOI: 10.1111/acem.14563

ORIGINAL CONTRIBUTION

Outcomes of the National Heart, Lung, and Blood Institute K12 program in emergency care research: 7-year follow-up

Cynthia D. Morris PhD, MPH[1,2] | Jennifer N. B. Cook GCPH[3] | Amber Lin MS[3] | Jane D. Scott ScD, MSN[4] | Nathan Kuppermann MD, MPH[5] | Clifton W. Callaway MD, PhD[6] | Donald M. Yealy MD[6] | Robert A. Lowe MD, MPH[3] | Lynne D. Richardson MD[7] | Stephen Kimmel MD, MSCE[8] | James F. Holmes MD, MPH[5] | Sean Collins MD, MS[9] | Lance B. Becker MD[10] | Alan B. Storrow MD[9] | Harrison J. Newgard BS[11] | Jill Baren MD, MS, MBA[12] | Craig D. Newgard MD, MPH[3]

[1]Department of Medical Informatics and Clinical Epidemiology, Oregon Health & Science University, Portland, Oregon, USA

[2]Oregon Clinical and Translational Research Institute, Oregon Health & Science University, Portland, Oregon, USA

[3]Center for Policy and Research in Emergency Medicine, Department of Emergency Medicine, Oregon Health & Science University, Portland, Oregon, USA

[4]Division of Cardiovascular Sciences, National Heart, Lung, and Blood Institute, Bethesda, Maryland, USA

[5]Department of Emergency Medicine, UC Davis Health, Sacramento, California, USA

[6]Department of Emergency Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

[7]Department of Emergency Medicine, Department of Population Health Science & Policy, Icahn School of Medicine at Mount Sinai, New York, New York, USA

[8]Department of Epidemiology, College of Public Health and Health Professions, College of Medicine, University of Florida, Gainesville, Florida, USA

[9]Department of Emergency Medicine, Vanderbilt University Medical Center, Nashville, Tennessee, USA

[10]Department of Emergency Medicine, Zucker School of Medicine at Hofsra Northwell, Manhasset, New York, USA

[11]University of California, Berkeley, Berkeley, California, USA

[12]Department of Emergency Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Correspondence
Craig D. Newgard, MD, MPH, Department of Emergency Medicine, Center for Policy and Research in Emergency Medicine, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, mail code CR-114, Portland, Oregon 97239-3098, USA.
Email: newgardc@ohsu.edu

Funding information
This project was supported by NIH/NHLBI K12 awards at: Icahn School of Medicine at Mount Sinai (#K12 HL109005); Oregon Health & Science University (#K12HL108974); University of California, Davis Medical Center (#K12 HL108964); University of Pennsylvania (#K12 HL109009); University of Pittsburgh (#K12 HL109068); and Vanderbilt University Medical Center (#K12 HL109019). The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung, and Blood

Abstract

Background: Long-term follow-up for clinician–scientist training programs is sparse. We describe the outcomes of clinician–scientist scholars in the National Heart Lung and Blood Institute (NHLBI) K12 program in emergency care research up to 8.7 years after matriculation in the program.

Methods: This was a cohort study of faculty clinician–scientist scholars enrolled in a NHLBI K12 research training program at 6 sites across the US, with median follow-up 7.7 years (range 5.7–8.7 years) from the date of matriculation. Scholars completed electronic surveys in 2017 and 2019, with the 2019 survey collecting information for their current work setting, percent time for research, and grant funding from all sources. We used NIH RePorter and online resources to verify federal grants through March 2021. The primary outcome was a funded career development award (CDA) or research project grant (RPG) where the scholar was principal investigator. We included funding from all federal sources and national foundations.

Results: There were 43 scholars, including 16 (37%) women. Over the follow-up period, 32 (74%) received an individual CDA or RPG, with a median of 36 months (range

© 2022 Society for Academic Emergency Medicine.

Institute, the National Institutes of Health, or the U.S. Department of Health and Human Services.

9–83 months) after entering the program. Of the 43 scholars, 23 (54%) received a CDA and 22 (51%) received an RPG, 7 (16%) of which were R01s. Of the 23 scholars who received a CDA, 13 (56%) subsequently had an RPG funded. Time to CDA or RPG did not differ by sex (women vs. men log-rank test $p = 0.27$) or specialty training (emergency medicine versus other specialties, $p = 0.59$).

Conclusions: After 7 years of follow-up for this NHLBI K12 emergency care research training program, three quarters of clinician–scientist scholars had obtained CDA or RPG funding, with no notable differences by sex or clinical training.

INTRODUCTION

A 2014 Physician-Scientist Workforce Report from the National Institutes of Health (NIH) showed that the proportion of physician-scientists comprising the biomedical workforce has been declining since the 1970's.[1] The same report noted 13,700 active physician-scientists in the US, just over 1% of all American physicians.[1] There have been a number of articles detailing the virtues of physician-scientists in the biomedical workforce, recommendations for increasing the pipeline, methods to decrease attrition, and unique threats to physician-scientists during the COVID-19 pandemic.[2–5] These issues are particularly poignant for emergency care research, which has historically received a relatively small proportion of NIH research funding[6] and has lacked a robust research workforce to translate clinical questions into high-quality and actionable research protocols.[7,8]

In recognition of the decline in the physician-scientist workforce, the NIH launched two career development programs targeting clinician-scientists in the 1990's (i.e., the K08 and K23 individual mentored career development programs). The NIH also uses the K12 Clinical Scientist Institutional Career Development program as another mechanism for research training within distinct fields. Under the K12 mechanism, institutions (rather than individuals) are awarded based on the training environment, with potential scholars applying to the institutions (rather than directly to the NIH) for consideration. The K12 mechanism is often targeted by an institute for areas or specialties needing a catalyst to expand the scientific workforce. The success of these institutional awards, as evidenced by the success of individual clinician-scientists selected for the programs, are typically measured by the proportion of scholars subsequently obtaining an individual career development award (CDA) or a research project grant (RPG).[9] There are limited studies tracking outcomes for K12 programs because such analyses typically require many years of follow-up. In 2018, we reported short-term scholar outcomes from the National Heart, Lung, and Blood Institute (NHLBI) K12 career development program in emergency care research.[10] The NHLBI K12 program was the first large federal investment in emergency care research training, representing a $21.6 million investment in six sites across the US.[11] In that report, we showed that 60% of scholars had secured independent CDA or RPG funding within 3 to 5 years of starting the K12 program,[10] yet not all scholars had finished the program by the end of the follow-up period. Because successfully competing for research funding can take many years, particularly for RPG awards, longer duration follow-up would provide a better assessment of the impact of the K12 program.

In this study, we describe the outcomes of scholars from the NHLBI K12 in emergency care research up to 8.7 years after matriculation into the program, including continued participation in research, a comprehensive assessment of subsequent independent grant funding, and time to funding.

METHODS

Study design

We conducted a prospective cohort study of scholars in the NHLBI K12 program in emergency care research using a participant survey, updated scholar curriculum vitaes (CVs), publicly available databases and on-line resources for subsequent grant funding. Each participant provided consent for inclusion and the Oregon Health & Science University institutional review board approved the study.

Setting

The six sites funded in 2011 for the NHLBI K12 Program in Emergency Care Research[11] included: Icahn School of Medicine at Mount Sinai (New York, NY); Oregon Health & Science University (Portland, OR); University of California, Davis School of Medicine (Sacramento, CA); University of Pennsylvania (Philadelphia, PA); University of Pittsburgh (Pittsburgh, PA); and Vanderbilt University Medical Center (Nashville, TN). While there was a separate, distinct 5-year trans-NIH K12 program starting in 2016, we report on the initial K12 program in emergency care research[11] to provide a longer follow-up period.

Selection of participants

We obtained the names of K12 scholars affiliated with the six program sites throughout the duration of the program. The first scholar entered the program July 1, 2012 and the last scholar finished on November 8, 2017. The K12 program provided up to

3 years of support for each scholar, including salary support, comprehensive mentorship, formal research didactic training, and research expenses and education up to $25,000 annually. Programs guaranteed that each scholar had at least 75% protected time for research. There was also an annual meeting at the NIH for all sites, including K12 scholars, program directors, and program administrators from each site.

Methods of measurement

Scholars completed an initial electronic survey in 2017 and a subsequent survey in 2019 using REDCap®. The initial survey included demographics (age, gender, race, and ethnicity), education, clinical training, primary academic appointment, post-K12 time spent on research, and grant applications submitted as a principal investigator (PI) following entry to the K12 program. These results were supplemented by qualitative data from interviews and reported in the initial evaluation of K12 scholar outcomes.[10] Scholars completed a subsequent survey in July 2019, which included two questions (current work setting [academic versus clinical] and percent time for research) and a request to upload an updated CV. We abstracted all funded grants (including their role and funding sources) since the date of matriculation into the K12 program from the CVs. We used NIH RePorter (NIH, AHRQ and VA funding), the PCORI website, Grantome (https://grantome.com/), Google, and websites at the academic programs represented by scholars to verify grants through March 2021.

Outcome measures

To measure scholar productivity, we categorized funded grants with the scholar as PI as CDAs (including K01, K08, K23, VA CDA2, and similar multi-year grants from professional or research foundations) and RPGs[9] (including R01, R03, R21, R33, R34, R61, VA Merit awards, and similar multi-year grant awards from professional or research foundations). As a subset of RPGs, we also examined R01 and equivalent awards (e.g., VA Merit awards). We did not consider industry funding for scholars, given the difficulty of differentiating investigator-initiated research from industry-led research. We only considered grants with scholars listed as PI (including multiple-PI grants) to focus on a scholar's achievement of independent research grant funding and the reporting capabilities of NIH RePorter and similar online websites.

Primary data analysis

We used descriptive statistics to characterize the K12 scholars. For each scholar, we identified their first CDA or RPG in relation to the date they started the K12 program, which provided a consistent time zero for all scholars. We calculated the time interval from the date of the K12 appointment to the date of the first CDA and RPG grant and illustrated these intervals using Kaplan–Meier curves. We also compared Kaplan–Meier curves by sex and clinical specialty (emergency medicine versus other specialties) using the log-rank test. We conducted all analyses in Stata 16.1 (StataCorp, College Station, TX, USA) and considered a p value ≤0.05 to represent statistical significance.

RESULTS

The K12 emergency care research training program included 43 clinician-scientists across the 6 sites (Table 1). We obtained

TABLE 1 Description of cohort

	All participants ($N = 43$)
Age at matriculation into K12 (Median, IQR)	37 years (34–40 years)
Sex	16 F (37%), 27 M (63%)
Race/Ethnicity	
White, non-Hispanic	22 (51%)
Asian	12 (27%)
Black	5 (11%)
White, Hispanic	3 (8%)
Other	1 (3%)
Degrees on entry to K12 program:	
MD + Master's degree	30 (70%)
MD	5 (12%)
MD, PhD	4 (9%)
PhD	4 (9%)
Professional discipline:	
Emergency Medicine	18 (42%)
Pediatric Emergency Medicine	6 (14%)
Adult Critical Care[*]	5 (12%)
Cardiology	5 (12%)
Pediatric Critical Care	2 (5%)
Trauma Surgery	2 (5%)
Pediatric allergy and immunology	1 (2%)
PhD, nursing	1 (2%)
PhD, epidemiology	1 (2%)
PhD, comparative pathology	1 (2%)
PhD, clinical research, physiology and biophysics	1 (2%)
Median scholars per institution (range)	7 (6 to 8)
Median effort for research at initial assessment	75% (IQR 30–75%)
Median effort for research from current assessment	75% (IQR 25–75%, range 0–90%)

[a]One scholar trained in neurology and critical care.

[b]Many of the cohort characteristics were published in the 3–5 year outcome assessment.[10]

TABLE 2 Grant outcomes of the K12 cohort ($n = 43$)

	CDA or RPG	CDA	RPG	R01
Scholars	32/43 (74%)	23/43 (54%)	22/43 (51%)	7/43 (16%)
Time to award (beginning with date of matriculation into the K12 program)	Median 36 months (IQR 26–43, range 9–83)	Median 35 months (IQR 26–40, range 12–83)	Median 50 months (IQR 36–61, range 9–78)	Median 56 months (IQR 38–73, range 26–78)
2 years	5 (12%)	3 (7%)	2 (5%)	0
3 years	15 (35%)	11 (26%)	5 (12%)	1 (2%)
4 years	26 (60%)	18 (42%)	10 (23%)	2 (5%)
5 years	30 (70%)	21 (49%)	14 (33%)	4 (9%)
6 years	30 (70%)	21 (49%)	17 (40%)	4 (9%)
7 years	32 (74%)	23 (54%)	22 (51%)	7 (16%)

CDA, career development award; RPG, research project grant.

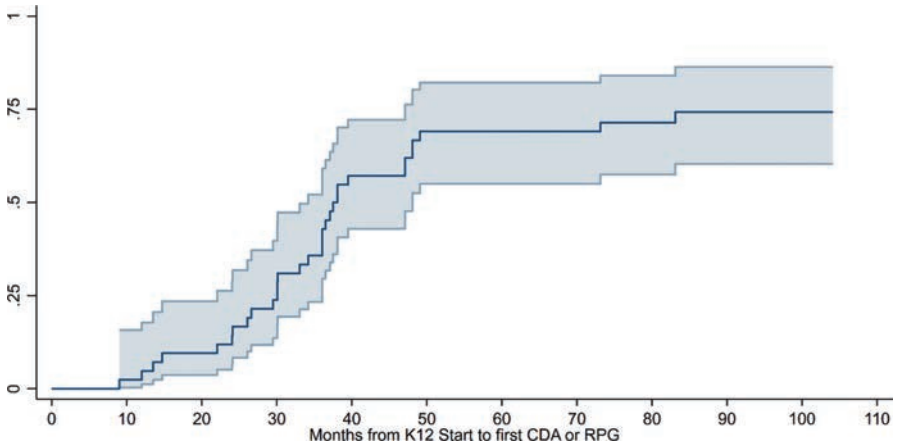


FIGURE 1 Time to receipt of first individual career development award or research project grant as principal investigator since beginning the K12 program ($n = 43$)
*95% CI shown. CDA, career development award; RPG, research project grant.

follow-up information directly from 41 of 43 (95%) scholars. For the two scholars who did not respond, we confirmed the presence versus absence of research grants through NIH RePorter and publicly-available on-line sources. The median follow-up period was 7.7 years from the date of matriculation into the K12 program (range 5.7–8.7 years). The demographics of scholars are shown in Table 1. The mean age was 37 years, 37% were female, 11% were Black, and 8% were Hispanic. The majority of scholars (25 of 43, 58%) continued to participate in research for at least 50% of their professional time. Eight (19%) scholars reported spending 25–50% time on research, and 5 (12%) reported spending up to 25% time performing research (one scholar reported no research activity and four provided no estimate).

We considered the research grant outcomes (CDA or RPG) of scholars as a measure of successful funding as PI. Of the 43 scholars, 32 (74%) received a CDA or RPG as PI during the follow-up period (Table 2, Figure 1). The time to first CDA or RPG occurred a median 36 months (range 9–83 months) after entering the K12 program. During follow-up of the 43 scholars, 23 (54%) received a CDA and 22 (51%) received an RPG, 7 (16%) of which were R01s (Table 2). Table 3 shows the grant type and funding agencies for CDAs and RPGs. Of the 23 scholars who received a CDA, 13 (56%) subsequently had an RPG funded, which occurred a median of 32 months (inter-quartile range 23 to 44 months) after the award date for the CDA. In contrast, 9 of 20 (45%) scholars who did not receive a CDA obtained RPGs (median 38 months from the start of the K12 program, IQR 30–47 months). Of the 9 scholars who obtained RPGs without a CDA, two of the RPGs were R01 awards. Although the change in rate in Figure 1 begins to plateau at approximately 50 months from entry to the K12 program, there were scholars who continued to receive grant awards after 80 months. In Figure 2, we show Kaplan–Meier curves for time-to-funding, separated by CDA versus RPG grants.

We investigated whether time to receipt of a CDA or RPG differed by scholar sex or specialty training. In Figure 3, we show that time to receipt of first CDA or RPG grant appeared longer for women (women median 37.5 months [IQR 34.1–49.1]; men median 30 months [IQR 26.1–38.1 months], $p = 0.14$), but the curves eventually align (log-rank test $p = 0.27$). When we evaluated time to grant funding by specialty training (emergency medicine versus other specialties, Figure 4), there was no difference between the groups ($p = 0.59$).

DISCUSSION

The first K12 in emergency care research resulted in a newly funded cohort of successful interdisciplinary clinician-scientists in this

TABLE 3 Funding agency and grant type for career development awards and research project grants with the scholar as principal investigator

Career development awards ($n = 23$)		
Type		
Federal - K23	15	(65%)
Federal - K08	6	(26%)
National foundation	2	(9%)
Funding agency		
NHLBI	15	(65%)
NINDS	2	(9%)
NICHD	3	(13%)
NIGMS	1	(4%)
Robert Wood Johnson Foundation	1	(4%)
American Heart Association	1	(4%)
Research project grants ($n = 22$)		
Type:		
R01	7	(32%)
R21	3	(14%)
R03	3	(14%)
R44	1	(5%)
Other	8	(36%)
Funding agency		
NHLBI	9	(41%)
NIDDK	2	(9%)
Department of Transportation	2	(9%)
Private foundation	2	(9%)
NINDS	1	(5%)
NIMHD	1	(5%)
NIA	1	(5%)
Office of Naval Research	1	(5%)
Department of Defense	1	(5%)
Veterans Administration	1	(5%)
Clinical and Translational Science Institute	1	(5%)

Abbreviations: NHLBI, National Heart, Lung, and Blood Institute; NINDS, National Institute of Neurological Disorders and Stroke; NICHD, National Institute of Child Health and Human Development; NIGMS, National Institute of General Medical Sciences; NIDDK, National Institute of Diabetes and Digestive and Kidney Diseases; NIMHD, National Institute on Minority Health and Health Disparities; NIA, National Institute on Aging.

domain, with three quarters securing CDAs or RPGs as PI during the follow-up period. However, the duration of time required to secure such awards was well beyond a 3-year training period for many scholars, particularly for RPG awards. With over 143 million annual ED visits in the US[12] and almost one in five adults[12] and children[13] requiring emergency care each year, there remains an urgent need to build a cadre of successful emergency care investigators to build the evidence base. This K12 program demonstrates that federal investment in such a training program helps to build the cadre of investigators, although securing subsequent funding takes many years.

The field of emergency care is inherently multidisciplinary, involving many medical specialties. It follows that clinician-scientists should similarly arise from a diverse group of specialties. This K12 training program included individuals across a variety of training backgrounds, where the orientation to NHLBI funding priorities within emergency care were defined by their research questions, rather than clinical specialty. Among the scholars, there was no difference in the success of grant funding by specialty training, illustrating that the ability to compete for subsequent funding is not dependent on the type of clinical training.

Increasing diversity within the research workforce improves innovation and impact[14,15] and is a high priority for the NIH.[16,17] Women are under-represented in the biomedical workforce[13] and as physician scientists.[1] We observed that overall success rates and times for achieving the next funding milestones were similar between men and women, yet the modest sample was likely under-powered to detect a difference. The Kaplan–Meier curves for women versus men suggested that women in the K12 cohort took longer to secure CDA or RPG funding than men, but the funding rates eventually merged. If we made the comparison at an earlier time point, the findings may have appeared different. These results are similar to those of a recent report on women in academic medicine, where a larger proportion of men had been promoted after 7 years, but this advancement gap narrowed by 10 years.[18] The challenges faced by women physicians and physician scientists, such as gender-based discrimination and the impact of family responsibilities, have been described in multiple reports.[1,18,19] Our results suggest that women clinician-scientists have similar grant success rates compared to their male colleagues, but may require more time, possibly due to the inequities and unique demands endured by women in healthcare and the biomedical workforce. These results provide important insights for federal funding agencies seeking to catalyze the careers and scientific discoveries of women.

Outcomes from the K12 emergency care research program compare favorably to scholar outcomes from Clinical and Translational Science Award KL2 programs across the US during the same period.[20] Within the K12 program, 70% of scholars obtained a CDA or RPG within 5 years of starting the program, while 60% of KL2 scholars had secured such funding within 5 years of completing the KL2 program.[20] All six institutions in the K12 cohort had an existing Clinical and Translational Science Award (CTSA) at the start of the K12 program, which provided focused education in research, instruction and infrastructure for team science, and mentorship. Similar to the KL2 program, K12 scholars represented many specialties. One unique difference in this K12 program compared to KL2 programs was the focus on a singular (if broad) content area – emergency care. This focus allowed development of an emergency care research community within and across sites, which may have enhanced career development opportunities, mentorship, and collaborations. In a previous qualitative evaluation of the K12 program,

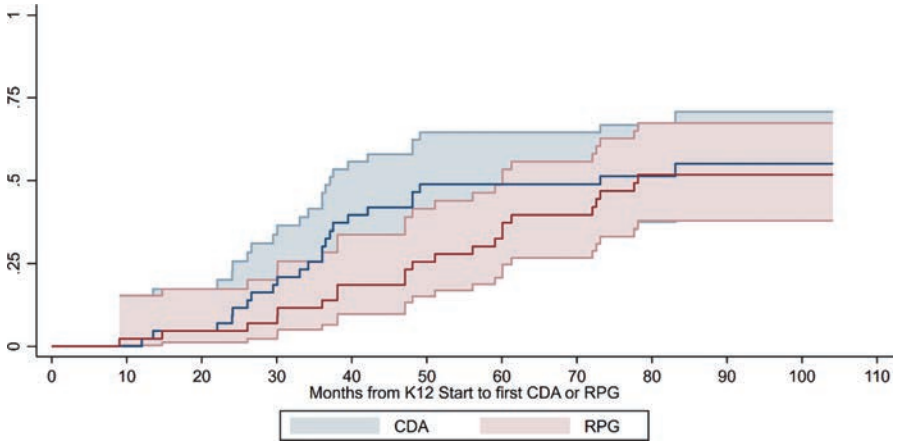


FIGURE 2 Time to receipt of first career development award and (separately) first research project grant as principal investigator since beginning the K12 program ($n = 43$)
*95% CI shown. CDA, career development award; RPG, research project grant.

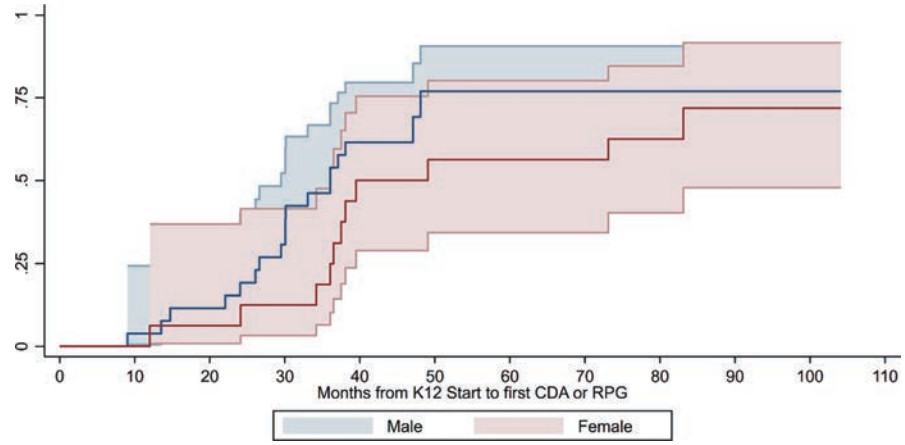


FIGURE 3 Time to receipt of first individual career development award or research project grant as principal investigator since beginning the K12 program, by scholar sex ($n = 43$)
*Log-rank test for difference in time to receipt by sex, $p = 0.27$. 95% CI shown.

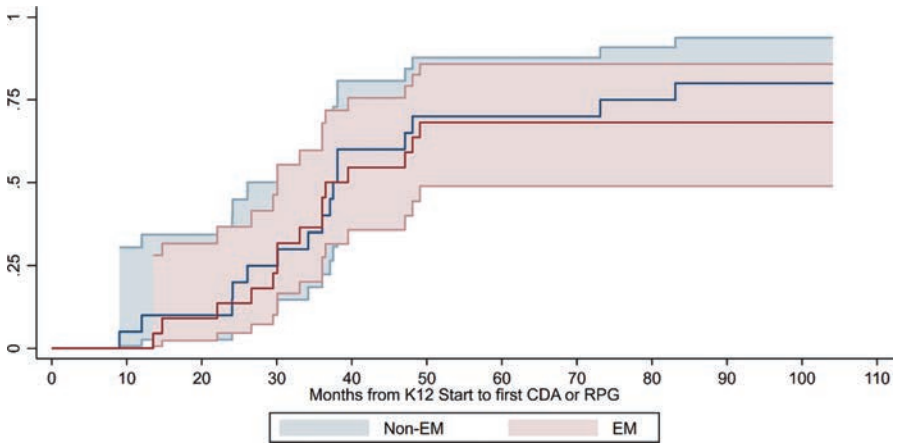


FIGURE 4 Time to receipt of first individual career development award or research project grant as principal investigator since beginning the K12 program, by specialty training ($n = 43$). EM, emergency medicine.
*Log-rank test for difference in time to receipt by training, $p = 0.59$. 95% CI shown.

participants described collaborations and networking as primary themes in the program, with some participants noting receipt of new research funding based on such programmatic collaborations.[10]

Clinician-scientists are often advised to pursue institutional research training (e.g., K12 or KL2 programs), followed by an independent CDA, prior to applying for a RPG. In our cohort, a higher proportion of scholars who obtained CDAs subsequently obtained RPGs, as compared to those who did not. While the numbers were small, this finding is consistent with an analysis of National Institute of Child Health and Human Development (NICHD) K12 programs, which found that an independent CDA, whether preceded by a K12 or not, was more likely to result in a RPG compared to a K12 alone.[21,22] A K12 provides a strategic boost to clinician-scientists embarking on a research path, which ideally includes a subsequent individual CDA after the K12. Scholars should be aware that it is uncommon to proceed directly to R01 funding after a K12 in the absence of an individual CDA; only two scholars in this cohort were successful in that pathway. Our findings also quantify and highlight the relatively long path for clinician-scientists to reach RPG funding, particularly for R01-level funding, which supports NIH results

showing the steadily increasing age of physician scientists at the time of their first RPG.[1]

The NHLBI K12 program in emergency care research may eventually have a multiplier effect on the emergency care research workforce and knowledge generation, as these clinician-scientists grow as independent investigators, mentors, and leaders. The effect of participation in this K12 program endured through 7 years after matriculation, with sustained growth of these investigators and the emergency care research community. We believe that this K12 program offered a strategic opportunity to enhance the next generation of emergency care researchers across multiple disciplines. However, impact on the field of emergency care and the patients served through the emergency care system will likely take years.

There were limitations in this study. This cohort represented a single K12 program and a modest sample size of scholars. However, the scholars came from six academic sites across the US and a variety of disciplines. Although we focused on grants received as PI, most scholars had multiple grants with other roles (i.e., co-investigator or site investigator), reflecting broader success in team science. We did not include scholars from the subsequent trans-NIH K12 program in emergency care research for consistency in measuring programmatic impact and maximizing follow-up time. Although inclusion of additional scholars would have increased the sample size, their follow-up would have been much shorter and some remain in training. Our primary focus was tracking scholar grant outcomes over an extended follow-up period.

Finally, we did not have a control group for comparison to the K12 scholars. Because this K12 program included a sizable number of scholars outside of emergency medicine, it was not possible to have a pre-post study design, as identification of a similar cohort prior to the K12 was not possible. Identification of a concurrent control group was similarly not feasible due to inherent differences that would create selection bias. We recognize that these issues limit evaluation of the impact of the NHLBI K12 program. However, description of the cohort and their grant funding over an extended period of time contributes valuable information about research workforce development in a specific scientific domain.

In summary, the NHLBI K12 program in emergency care research added to and may have enhanced a competitive workforce of interdisciplinary emergency care investigators. After 7 years of follow-up, approximately three quarters of scholars successfully obtained CDA or RPG funding, with no notable differences by scholar sex or specialty training. These results compare favorably and may exceed the success of other long-standing institutional research training programs.

MEETINGS

These results have not been presented or published in any form.

AUTHOR CONTRIBUTIONS

CM and CDN conceived of and designed the study. CDN obtained research funding specific to the K12 evaluation project. CDN, CM, DY, SC, JH, NK, LR, SK, LB, RL, CC, JB, and AS obtained funding for the six K12 sites. JNBC assisted with study coordination and data collection for the project. CM and HJN performed additional data collection. AL analyzed study data. CM and CDN supervised the study. CM and CDN drafted the manuscript. All authors helped interpret study results, refine the analysis, and contributed substantially to manuscript revision. CDN takes responsibility for the paper as a whole.

CONFLICT OF INTEREST

No author had conflicts of interest related to this study.

ORCID

James F. Holmes https://orcid.org/0000-0003-3270-2720
Craig D. Newgard https://orcid.org/0000-0003-1083-3455

REFERENCES

1. *Physician-Scientist Workforce (PSW) Report* 2014. 2014.
2. Salata RA, Geraci MW, Rockey DC, et al. U.S. Physician-scientist workforce in the 21st century: recommendations to attract and sustain the pipeline. *Acad Med*. 2018;93(4):565-573.
3. Permar SR, Ward RA, Barrett KJ, et al. Addressing the physician-scientist pipeline: strategies to integrate research into clinical training programs. *J Clin Invest*. 2020;130(3):1058-1061.
4. Rao RC, Dlouhy BJ, Capell BC, Akeju O. The endangered physician-scientist and COVID-19. *Cell Rep Med*. 2021;2(2):100190.
5. Jain MK, Cheung VG, Utz PJ, Kobilka BK, Yamada T, Lefkowitz R. Saving the endangered physician-scientist - a plan for accelerating medical breakthroughs. *N Engl J Med*. 2019;381(5):399-402.
6. Brown J. National Institutes of Health support for clinical emergency care research, 2011 to 2014. *Ann Emerg Med*. 2016;68(2):164-171.
7. Institute of Medicine, Committee on the Future of Emergency Care in the United States Health System. *Hospital-Based Emergency Care at the Breaking Point*. National Academies Press; 2006.
8. Neumar RW, Blomkalns AL, Cairns CB, et al. Emergency medicine research: 2030 strategic goals. *Acad Emerg Med*. 2021;29:241-251.
9. Definitions of Criteria and Considerations for Research Project Grant (RPG/R01/R03/R15/R21/R34) Critiques National Institutes of Health. NIH Grants & Funding, NIH Central Resource for Grants and Funding Information Web site. Accessed February 28, 2022.
10. Newgard CD, Morris CD, Smith L, et al. The first National Institutes of Health institutional training program in emergency care research: productivity and outcomes. *Ann Emerg Med*. 2018;72(6):679-690.
11. NHLBI Research Career Development Programs in Emergency Medicine Research (K12), RFA HL-11-011. National Institutes of Health. https://grants.nih.gov/grants/guide/rfa-files/rfa-hl-11-011.html. Published 2010. Accessed March 1, 2022.
12. Trency Department Services, 2009–2018. Washington, DC, 2021.
13. Health, United States, Emergency department visits within the pasts 12 months among children under age 18. National Center for Health Statistics. https://www.cdc.gov/nchs/data/hus/2019/036-508.pdf. Published 2019. Accessed.
14. Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW. Evidence for a collective intelligence factor in the performance of human groups. *Science*. 2010;330(6004):686-688.
15. AlShebli BK, Rahwan T, Woon WL. The preeminence of ethnic diversity in scientific collaboration. *Nat Commun*. 2018;9(1):5163.
16. Inviting Comments and Suggestions on the Draft NIH Chief Officer for Scientific Workforce Diversity Strategic Plan for FYs 2022–2026. National Institutes oh Health. https://grants.nih.gov/

grants/guide/notice-files/NOT-OD-22-054.html. Published 2022. Accessed March 4, 2022.
17. Bernard M. *NIH Chief Officer for Scientific Workforce Diversity Strategic Plan Draft*. Bethesda, MD; 2022.
18. Lautenberger DM, Dandar VM. *The State of Women in Academic Medicine 2018–2019 - Exploring Pathways to Equity*. American Association of Medical Colleges; 2020.
19. *United States Government Accountability Office Report to Congressional Committees: NIH Research - Action Needed to Ensure Workforce Diversity Strategic Goals Are Achieved*. August 2018 2018. GAO-18-545.
20. Sorkness CA, Scholl L, Fair AM, Umans JG. KL2 mentored career development programs at clinical and translational science award hubs: practices and outcomes. *J Clin Transl Sci*. 2020;4(1):43-52.
21. Twombly DA, Glavin SL, Guimond J, Taymans S, Spong CY, Bianchi DW. Association of National Institute of child health and human development career development awards with subsequent research project Grant funding. *JAMA Pediatr*. 2018;172(3):226-231.
22. Good M, McElroy SJ, Berger JN, Moore DJ, Wynn JL. Limited achievement of NIH research independence by pediatric K award recipients. *Pediatr Res*. 2018;84(4):479-480.

How to cite this article: Morris CD, Cook JNB, Lin A, et al. Outcomes of the National Heart, Lung, and Blood Institute K12 program in emergency care research: 7-year follow-up. *Acad Emerg Med*. 2022;29:1197-1204. doi: 10.1111/acem.14563

Received: 24 January 2022 | Revised: 31 May 2022 | Accepted: 12 July 2022

DOI: 10.1111/acem.14574

ORIGINAL CONTRIBUTION

Evaluation of the "hemoptysis" item in clinical decision rules for the diagnosis of pulmonary embolism in the emergency department

Héloïse Bannelier MD[1] | Judith Gorlicki MD[2] | Andrea Penaloza MD, PhD[3] | Delphine Douillet MD, PhD[4,5] | Pierre-Marie Roy MD, PhD[4,5] | Yonathan Freund MD, PhD[1,6] | Melanie Roussel MD[1,6]

[1]Emergency Department, Hôpital Pitié-Salpêtrière, Assistance Publique-Hôpitaux de Paris, Paris, France

[2]Emergency Department, Hôpital Avicenne, Assistance Publique-Hôpitaux de Paris, INSERM U942-MASCOT, Bobigny, France

[3]Emergency Department, Cliniques Universitaires Saint-Luc, Université Catholique de Louvain, Bruxelles, Belgium

[4]Emergency Department, Angers University Hospital, Angers, France

[5]FCRIN, INNOVTE, Saint Etienne, France

[6]Sorbonne Université, Improving Emergency Care FHU, Paris, France

Correspondence
Melanie Roussel, Service des urgences, 47-83 Bd de l'hôpital, 75013 Paris, France.
Email: melanie.roussel@chu-rouen.fr

Abstract

Background: Hemoptysis is not common in pulmonary embolism (PE) and lacks specificity for its diagnosis. However, this item is present in different validated scores that estimate the clinical probability of PE. The relevance of this item in clinical decision rules (CDRs) is not clearly established.

Objective: The aim of this study was to evaluate the impact of removing the "hemoptysis" item from the PERC, YEARS, and PEGeD CDR in patients with low clinical probability of PE.

Design: This was a post hoc analysis of two European prospective cohorts, which included 2968 patients presenting to the ED with a low clinical probability of PE (PROPER and PERCEPIC) and a 3-month follow-up. The primary endpoint was the false-negative rate of a CDR score without the hemoptysis item. Secondary endpoints included the potential reduction of chest imaging if the item hemoptysis was to be removed and risk stratification of the Geneva and Wells scores without the hemoptysis item.

Results: Of 2968 patients included (mean ± SD age 46 ± 18 years, 53% female), 87 patients (3%) had a PE diagnosed at 3 months. A total of 2908 were followed-up at 3 months and analyzed. Using the PERC rule with and without the hemoptysis item, there were 13 and 14 missed cases of PE, respectively (failure rate 0.45% [95% CI 0.25%–0.78%] and 0.48% [95% CI 0.27%–0.82%]). Using the YEARS strategy, there were 11 missed PE cases with or without the hemoptysis item (false-negative rate 0.57% [95% CI 0.30%–1.05%]). With the PERC and YEARS rule, removing the hemoptysis item would have led to a 1% reduction in chest imaging. The PEGeD strategy was not modified by the removal of the hemoptysis item.

Supervising Editor: Dr. Richard Taylor

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2022 The Authors. *Academic Emergency Medicine* published by Wiley Periodicals LLC on behalf of Society for Academic Emergency Medicine.

Conclusions: The hemoptysis item could be safely removed from the PERC, YEARS, and PEGeD CDRs. However, there was no subsequent clinically relevant reduction of chest imaging.

KEYWORDS
computed tomography pulmonary angiography, D-dimer, emergency department, emergency medicine, hemoptysis, PEGeD, PERC, probability scores, pulmonary embolism, revised Geneva, Wells, YEARS

INTRODUCTION

Pulmonary embolism (PE) is frequently suspected in emergency department (ED) patients, which often leads to the prescription of irradiative chest imaging (computed tomographic pulmonary angiogram [CTPA] in most cases).[1,2] In the past few years, rationalizing the use of chest imaging in the ED has become a priority. Indeed, an increased use of CTPA has been reported without clear benefit in terms of prognosis.[3] This increased use is reportedly associated with potential overdiagnosis of PE, increased cost, length of ED stay, and side effects from both chest imaging and undue anticoagulant treatments.[3,4]

To reduce the number of CTPA procedures performed to rule out PE, different clinical decision rules (CDRs) have been derived and validated. Among them, the Wells and Geneva scores coupled with the D-dimer assay have shown a high negative predictive value (NPV), allowing the exclusion of the diagnosis of PE without the need of chest imaging.[5,6] The D-dimer assay is sensitive but not specific causing many false positives in low-risk patients with subsequent negative CT scans. This led to the development of a clinical rule, the PERC (Pulmonary Embolism Rule-out Criteria) rule, allowing PE to be ruled out on clinical criteria only.[7] The YEARS and PEGeD CDRs were recently developed, allowing the D-dimer threshold to be raised for ordering chest imaging to 1000 ng/mL upon absence of certain criteria (a YEARS score of 0 and a low Wells clinical probability, respectively).[8,9] These three CDRs and the two scores of clinical probability all include the "hemoptysis" item, which may change the diagnostic workup if present. The literature on PE and hemoptysis is poor, and it is reported that hemoptysis is neither sensitive nor specific to PE. Historically, hemoptysis was part of a triad considered specific to PE and necessary for its diagnosis, combining dyspnea, chest pain, and hemoptysis, but studies at the time reported much higher rates of hemoptysis in PE than recent studies (29% in a 1957 study, compared to 3% in 2004 in a study by Kline et al.[10]). The main etiologies of hemoptysis are respiratory infections, neoplasia, and bronchial dilatation.[11] Whether hemoptysis is a relevant clinical item to adapt the diagnostic workup for PE in the ED is unknown, especially considering that the real PE prevalence in ED patients with suspected PE has become significantly lower than in previous validation studies. Removing the hemoptysis item from these CDRs would simplify them and could increase the number of very-low-risk patients for whom additional tests could be dispensed with. The objective of this study was to assess whether the inclusion of the hemoptysis item in CDR is relevant for patients with low implicit clinical probability and to assess the impact of removing this clinical item from the different CDR on subsequent on the failure rate of these diagnostic algorithms.

METHODS

Study design

This study was a post hoc analysis of two European cohorts, which prospectively included patients presenting to the ED with a low implicit clinical probability of PE: PROPER and PERCEPIC. The PROPER study was a cluster randomized noninferiority trial that compared the PERC-based diagnostic strategy with the conventional strategy in patients with a low clinical probability of PE.[12] The PERCEPIC study was an observational, prospective, multicenter study that aimed to prospectively assess the NPV of the PERC rule in a European cohort of patients with a low clinical probability of PE.[13] In both studies, patients were followed up at 3 months through a telephone interview, and the presence of a thromboembolic event during follow-up was confirmed by an adjudication committee. When the patient could not be reached, the patient's general practitioner was contacted. For patients lost to follow-up, death records were queried. Any death during follow-up was analyzed by the adjudication committee and a sudden unexplained death was considered a thromboembolic event if it could not be excluded.

The protocols of both studies were approved by the ethics committees. As a secondary analysis on anonymized data, patient consent was waived for this study.

Population

A total of 2968 patients were included (Figure 1). The PROPER cohort included 1916 patients (from 14 EDs in France, between 2015 and 2016) presenting to the ED with a low pretest probability of PE empirically assessed by the physician's Gestalt. The PERCEPIC cohort included 1757 patients from 12 EDs in France and Belgium, between May 2015 and April 2016, including 1052 patients with low clinical probability of PE who were included in the present study;

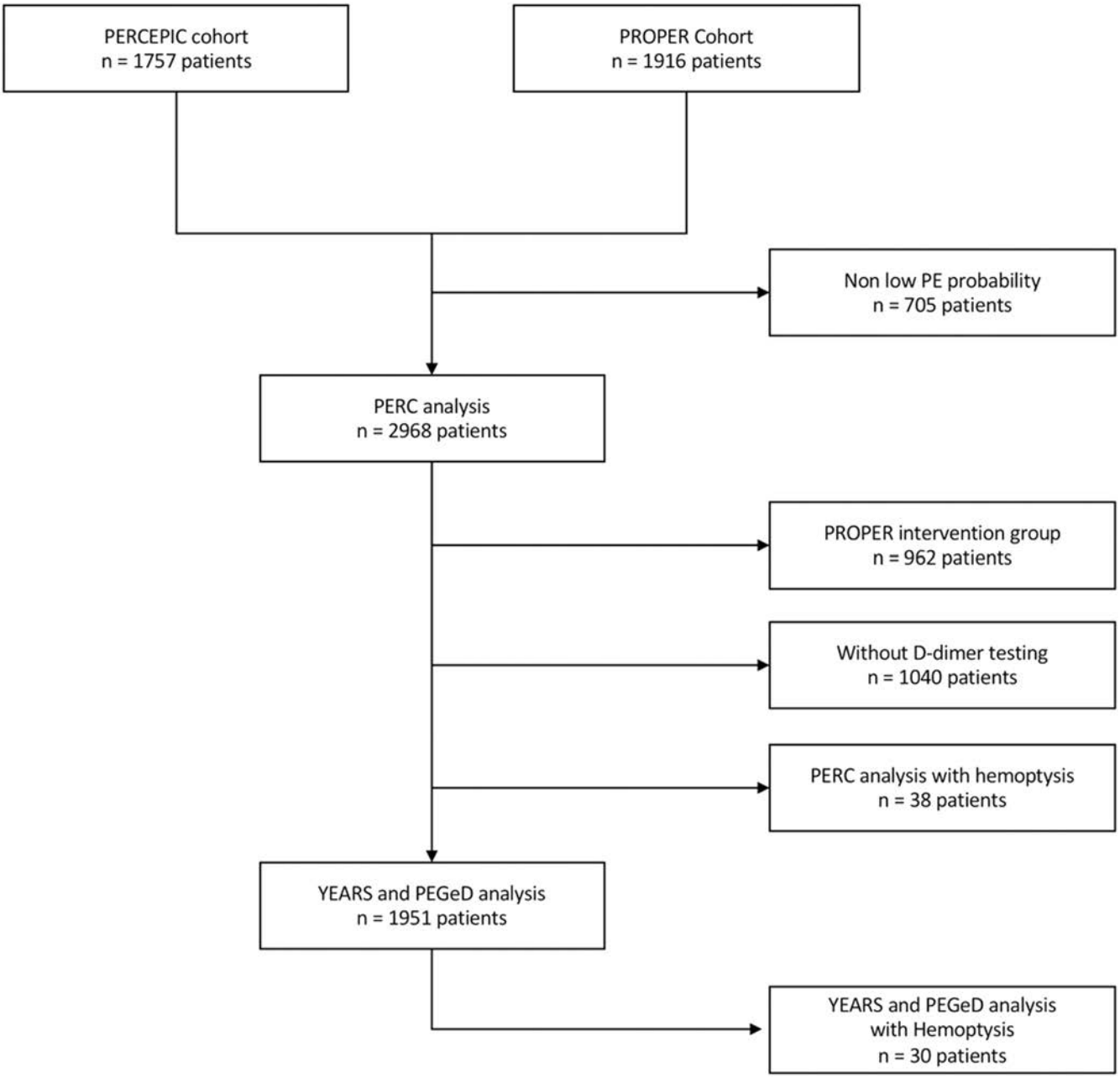


FIGURE 1 Flow chart. PE, pulmonary embolism PERC, Pulmonary Embolism Rule-out Criteria

705 patients were excluded from the PERCEPIC cohort for non–low probability of PE (558 with a moderate clinical probability, 147 with a high clinical probability of PE). In both studies, the low clinical probability was assessed by an unstructured physician's Gestalt evaluation of a risk of PE less than 15%.

Primary and secondary endpoints

The primary endpoint was the failure rate of the PERC, YEARS, and PEGeD based strategy without the hemoptysis item, when patients were followed up at 3 months. In the 60 (2%) patients lost to follow-up at 3 months in whom PE was ruled out at ED visit, we considered that they did not met the primary endpoint of PE at 3 months. The secondary endpoints included the potential reduction of CTPA if these CDR were applied without the hemoptysis item and the prevalence of PE according within the different classes of Geneva and Wells scores with and without hemoptysis item. A sensitivity analysis on PE diagnosed in the ED (and not at 3 months) was also performed.

Statistical analysis

Patient characteristics were expressed as number (percentage) for categorical variables, and as mean (standard deviation [SD]) or

median (interquartile range [IQR]) for quantitative variables, depending on their distribution. Confidence intervals (CIs) were calculated using Wilson's method with continuity correction. Recent recommendations suggested to adapt the acceptable failure rate to the prevalence of PE in the tested population. To validate the safety of a diagnostic strategy, the 2017 International Society of Thrombosis and Hemostasis recommendations of the Scientific and Standardization Committee were followed, according to which the maximum acceptable failure rate should be 1.82% + 0.0053% × prevalence.[14–16] In our population, the prevalence of PE was 3%. The upper bound of the 95% CI of our primary endpoint should then be less than 1.84% to respect these recommendations.

The analysis of the primary endpoint was performed on all patients analyzed at 3 months, i.e., 2908 patients. For the analyses performed to evaluate the YEARS and PEGeD strategy, in patients who received a conventional strategy, i.e., a D-dimer assay followed if necessary by CTPA, were included. We therefore excluded patients in the PROPER intervention group that did not undergo D-dimer testing. These analyses were thus performed on a total population of 1951 patients (Figure 1.) A sensitivity analysis that only considered PE diagnosed in the ED (and not at 3-month follow-up) was performed, which included 2968 patients.

In line with recent suggestions that the safety of a new strategy should be best evaluated among patients where the strategy was actually applied, post hoc analyses on the subgroup of patients that had a PERC, YEARS, and PEGeD score of zero, i.e., those in whom the strategy would have actually changed the workup strategy.[15,16]

RESULTS

A total of 2968 patients were included; 2908 were followed-up at 3 months and are included in the analysis. The mean (±SD) age was 46 (±18) years and 1584 (53%) were female. At 3 months, 87 PEs were diagnosed (prevalence of 3%) and seven (0.2%) patients died. Baseline characteristics are reported in Table 1. There were three (8.1%) PE cases in the 37 patients with hemoptysis and 84 (2.9%) PE cases in the 2871 patients with no hemoptysis (difference 5.2% [95% CI 0.0%–18%]].

TABLE 1 Baseline characteristics

	n		
	2968	**Population**	
Age (years), mean (±SD)		46	(±18)
Age > 50 years		1191	(40)
Gender			
Male		1384	(46)
Female		1584	(53)
HR (beats/min), mean (±SD)		83	(±17)
HR > 100 beats/min		457	(15)
SpO_2 (%), mean (IQR)		98	(97–100)
SpO_2 < 95%		217	(7)
Temperature (°C), mean (±SD)		36.7	(±1.8)
Risk factors for PE	2968		
Estrogen use		291	(10)
Signs of DVT		135	(5)
History of VTE		146	(5)
Recent immobilization		101	(3)
Hemoptysis		38	(1)
Active cancer		51	(2)
PE considered most likely diagnosis		242	(8)
D-dimer testing		2477	(84)
Positive D-dimer[a]	2477	704	(28)
D-dimer value (µg/L)	2477		
<500		1648	(67)
500–1000		471	(19)
>1000		358	(14)
Number of CTPA procedures	2968	649	(22)
PE diagnosed in the ED	2968	84	(3)
PE diagnosed at 3 months	2968	87	(3)
Death from any cause at 3 months	2968	7	(0.2)

Abbreviations: CTPA, computed tomography pulmonary angiogram; DVP, deep venous thrombosis; HR, heart rate; IQR, interquartile range; PE, pulmonary embolism; VTE, venous thromboembolism.

[a]Positive D-dimer if >500 µg/L in patients 50 years old and under and if >age x 10 µg/L in patients over 50 years old.

Primary endpoint

Among the 2908 patients included and followed up at 3 months, 1137 (39%) patients had a PERC score of 0 in whom 13 patients had a PE that was not diagnosed in the ED (failure rate 0.45% [95% CI 0.25%–0.78%]). After the hemoptysis item was removed, 1151 (40%) patients had a PERC score of 0 among whom 14 had a PE (failure rate 0.48% [95% CI 0.27%–0.82%]). One additional PE case was not diagnosed using the PERC rule without hemoptysis in a 34-year-old patient.

Among the 1925 patients included in the YEARS analysis and followed up at 3 months, 1651 (86%) patients had a YEARS score of 0 among whom 11 patients had a PE that was not diagnosed in the ED (failure rate 0.57% [95% CI 0.30%–1.05%]). After the hemoptysis item was removed, 1676 (87%) patients had a YEARS score of 0 among whom 11 had a PE (failure rate 0.57% [95% CI 0.30%–1.05%]).

Among the 1925 patients included in the PEGeD analysis and followed up at 3 months, 1237 (64%) patients had a Wells score of 0, and 11 had a PE (failure rate 0,57%% [95% CI 0.30% to 1.05%]). After the hemoptysis item was removed, 1257 (65%) patients had a Wells score of 0 and 12 patients had a PE (failure rate 0.62% [95% CI 0.34%–1.11%]).

Using the PERC rule, the removal of the hemoptysis item would have led to the reduction of six (1.0%) CTPA procedures (560 without hemoptysis item vs. 566 with hemoptysis item), and for the YEARS rule the reduction was seven (1.0%) CTPA procedures (344 without hemoptysis item vs. 351 with hemoptysis item). There was no CTPA reduction for the PEGeD rule. The results are summarized in Tables 2 and 3.

Sensitivity analysis

A sensitivity analysis was performed on the data during the ED visit. Among the 2968 patients included, 38 had hemoptysis in whom three had a PE at baseline, compared 81 in the 2830 patients without PE (7.9% vs. 2.8%. respectively; difference 5.1% [95% CI 0%–18%]).

The failure rates of the PERC rule with or without hemoptysis were 0.40% (95% CI 0.22%–0.72%) and 0.44% (95%CI 0.24%–0.77%) respectively. The failure rate of the YEARS rule with or without hemoptysis was 0.46% (95%CI 0.24%–0.87%) in both groups, and the failure rate of the PEGeD rule with or without hemoptysis was 0.62% (95% CI 0.33%–1.1%).

Revised Geneva, Wells, and simplified Wells scores

Performances of the revised Geneva, Wells, and simplified Wells scores with and without the hemoptysis item are reported in Table 4. The risk stratification remained very similar after this change.

TABLE 2 Failure rate according to the diagnostic strategy at 3 months

Diagnostic strategy	False negative (*n*)	Failure rate (% [95% CI])
PERC strategy (*n* = 2908)		
PERC	13	0.45 (0.25–0.78)
PERC without hemoptysis item	14	0.48 (0.27–0.82)
YEARS strategy (*n* = 1925)		
YEARS	9	0.57 (0.30–1.05)
YEARS without hemoptysis item	9	0.57 (0.30–1.05)
PEGeD strategy (*n* = 1925)		
PEGeD	11	0.57 (0.30–1.05)
PEGeD without hemoptysis item	12	0.62 (0.34–1.11)

Abbreviation: PERC, Pulmonary Embolism Rule-out Criteria.

Post hoc analyses

Analyses in the subgroup of patients in which the different scores were zero are reported in Table 5.

DISCUSSION

In this retrospective analysis of 2968 patients from two European cohorts, the failure rates of the PERC, YEARS, and PEGeD based

TABLE 3 Reduction of CTPA according to the diagnostic strategy

Diagnostic strategy	CTPA reduction, *n* (%)
PERC without hemoptysis item vs. PERC	6 (0.98)
YEARS without hemoptysis item vs. YEARS	7 (0.98)
PEGeD without hemoptysis item vs. PEGeD	0 (0)

Abbreviations: CTPA, computed tomography pulmonary angiogram; PERC, Pulmonary Embolism Rule-out Criteria.

TABLE 4 PE risk stratification based on Wells and revised Geneva scores

Clinical probability scores (%)	PE at 3 months	PE in the ED
Revised Geneva score		
Low risk (<2)	1.2	1.1
Intermediate risk (2–4)	5.4	5.1
High risk (>4)	19.4	18.8
Revised Geneva score without hemoptysis item		
Low risk (<2)	1.2	1.2
Intermediate risk (2–4)	5.4	5.1
High risk (>4)	20.7	20.0
Wells score		
Low risk (<2)	2.3	2.2
Intermediate risk (2–6)	7.1	6.7
High risk (>6)	17.6	17.6
Wells score without hemoptysis item		
Low risk (<2)	2.3	2.2
Intermediate risk (2–6)	7.2	6.7
High risk (>6)	15.4	15.4
Simplified Wells score		
PE unlikely (0–1)	2.5	2.4
PE probable (>1)	10.6	9.8
Simplified Wells score without hemoptysis item		
PE unlikely (0–1)	2.5	2.4
PE probable (>1)	10.5	9.7

Abbreviation: PE, pulmonary embolism.

TABLE 5 Post hoc analyses: false-negative rates among patients with negative CDR

Diagnostic strategy	Patients with negative CDR (*n*)	False negative, *n* (%)
PERC strategy		
PERC score = 0	1137	13 (1.14)
PERC score = 0 without hemoptysis item	1151	14 (1.20)
YEARS strategy		
YEARS score = 0	1651	11 (0.67)
YEARS score = 0 without hemoptysis item	1676	11 (0.66)
PEGeD strategy		
PEGeD	1605	11 (0.69)
PEGeD without hemoptysis item	1606	12 (0.75)

Abbreviation: CDR, clinical decision rule; PERC, Pulmonary Embolism Rule-out Criteria.

diagnostic strategies were below the acceptable threshold after removing the hemoptysis item. This change would have led to a slight potential reduction of chest imaging.

In this study, there was a nonsignificant association between hemoptysis and PE, with an absolute difference of 5.2% [95% CI 0.0%–18%], probably due to a lack of power as hemoptysis is reportedly associated with an increased risk of PE.[10] The prevalence of hemoptysis among patients with a low clinical probability of PE was very low; only 38 patients (1%) had hemoptysis and among them three patients (8.1%) had PE. In another recent study, similar results were found of low prevalence of hemoptysis in patients with low clinical probability of PE. Among 1414 patients with low clinical probability, 28 patients (2%) had hemoptysis and among them four patients (14.3%) had PE.[17]

The performance of the PERC score is not significantly reduced by the removal of the hemoptysis item. One PE is not diagnosed by applying the PERC strategy without hemoptysis, out of 15 patients with hemoptysis and no other PERC criteria. The prevalence of PE in patients with hemoptysis and no other PERC criteria is therefore 6.6%. Removing the hemoptysis item reduces the number of scans by only six (i.e., 1%). The performance of the YEARS and PEGeD rules are not altered by the removal of hemoptysis. A small reduction in the number of scans is achieved by applying the YEARS rule without hemoptysis. The risk stratification by the Geneva and Wells scores remains similar without the hemoptysis item.

Removing the hemoptysis item from the current CDR seems therefore safe, but this study does not provide enough argument to remove the hemoptysis item from the PERC score. The absence of modification in the YEARS and PEGeD strategies questions the interest of the hemoptysis item in these algorithms, but removing it would not lead to an important reduction of CT scans.

The main strength of this study is the large size of the cohort with a prospective recruitment in several EDs, with very few lost to follow-up. The overall prevalence of PE in this cohort was 3%, which is very different than the ones that were used for the derivation and validations of the tested CDRs and probability scores. This rate is consistent with what has been recently reported in low-risk population.[12,17,18]

However, in the Wells' score princeps article, the PE prevalence was 9.5%,[19] and this prevalence was 23% in the revised Geneva score publication.[20] Since it has been increasingly reported that the real PE prevalence in routine ED practice is below 10%, there was a need to reassess the relevance of including the chosen clinical items in real-life population. Furthermore, the interest of including hemoptysis item in the different CDR is of limited value for patients presenting with hemoptysis.

Indeed, CT scans are widely performed in these patients to localize the origin of the bleed as well as to establish the etiological diagnosis. The application of the different CDR may therefore not reduce the number of CT scans in these patients. The real benefit in terms of reducing the number of scans would be even lesser than that shown in this study. Moreover, the application of the PE diagnostic scores in patients presenting with hemoptysis systematically leads to the measurement of D-dimer levels, which delays the CT scan.[11] Therefore, it can be suggested that patients with hemoptysis should not follow usual diagnostic strategies for PE and should undergo chest CT that includes bronchial artery injection.

According to the post hoc analyses, both strategies appear to meet the requirements of the SSC 2017 guidelines even using the failure rate in the actual impacted population. All of these strategies appear to be within recommendations after applying a failure rate relative to the population actually impacted by the strategy, but a statistical analysis with determination of the 95% CI of the failure rate would be required to confirm this hypothesis.

LIMITATIONS

This study has some limitations. First, only patients with a low implicit probability of PE were included in this cohort. Although patients were included following an unstructured Gestalt estimation of <15%, the overall prevalence of PE was 3%. Hemoptysis

alone or in association with other symptoms may have changed the probability to not low. The data from this study may therefore not be applicable in patients with intermediate clinical probability of PE.

Second, there were only 38 patients (1%) with hemoptysis. This could have led to a lack of power to detect significant differences. However, the clinical impact would still be limited. Moreover, there was a small number of patients with hemoptysis in our cohorts, with subsequent wide CIs for estimating the prevalence of PE.

Third, some patients were lost to follow-up at 3 months and may have had a PE that was not diagnosed at ED visit, which could be a bias in the interpretation of the results. However, in the two included cohorts, the overall rate of missed diagnostic is approximately 0.1%. Therefore, it is likely that none of the patients lost to follow-up may have had a PE at 3 months.

Fourth, data regarding localization and severity of PEs were not collected. There may have been a difference in diagnosed PEs with or without hemoptysis. It was not possible to evaluate the association of PE severity with the presence of hemoptysis.

Fifth, the clinical characteristics of the hemoptysis were not collected, especially its estimated volume. Therefore, it was not possible to determine whether certain characteristics are more suggestive of PE than others.

CONCLUSIONS

In patients with low clinical probability of pulmonary embolism, the hemoptysis item could be safely removed from the PERC, YEARS, and PEGeD based strategies for pulmonary embolism, with a low risk of diagnostic failure. However, there was no subsequent clinically significant reduction in the rate of chest imaging.

AUTHORS CONTRIBUTIONS

Héloïse Bannelier and Yonathan Freund conceived the study. Héloïse Bannelier and Yonathan Freund developed the analysis plan. Yonathan Freund undertook the main analysis. Héloïse Bannelier and Melanie Roussel wrote the first draft of the paper. Judith Gorlicki, Andrea Penaloza, Delphine Douillet, and Pierre-Marie Roy made important critical revisions. All authors have read and approved the final version of the manuscript.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

DATA AVAILABILITY STATEMENT

Please contact author for data requests.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The protocols of both studies (PROPER and PERCEPIC) were approved by the ethics committees.

CONSENT FOR PUBLICATION

Not applicable. As a secondary analysis on anonymized data, patient consent was waived for this study.

ORCID

Yonathan Freund https://orcid.org/0000-0002-0262-3848
Melanie Roussel https://orcid.org/0000-0003-2467-8484

REFERENCES

1. Feng LB, Pines JM, Yusuf HR, Grosse SD. U.S. trends in computed tomography use and diagnoses in emergency department visits by patients with symptoms suggestive of pulmonary embolism, 2001-2009. Mark Courtney D, ed. *Acad Emerg Med*. 2013;20(10):1033-1040. doi:10.1111/acem.12221
2. Lefevre-Scelles A, Jeanmaire P, Freund Y, Joly LM, Phillipon AL, Roussel M. Investigation of pulmonary embolism in patients with chest pain in the emergency department: a retrospective multicenter study. *Eur J Emerg Med off J Eur Soc Emerg Med*. 2020;27(5):357-361. doi:10.1097/MEJ.0000000000000680
3. Wiener RS, Schwartz LM, Woloshin S. Time trends in pulmonary embolism in the United States: evidence of overdiagnosis. *Arch Intern Med*. 2011;171(9):831-837. doi:10.1001/archinternmed.2011.178
4. Wiener RS, Schwartz LM, Woloshin S. When a test is too good: how CT pulmonary angiograms find pulmonary emboli that do not need to be found. *BMJ*. 2013;347:f3368. doi:10.1136/bmj.f3368
5. Gibson NS, Sohne M, Kruip MJHA, et al. Further validation and simplification of the Wells clinical decision rule in pulmonary embolism. *Thromb Haemost*. 2008;99(1):229-234. doi:10.1160/TH07-05-0321
6. Láinez-Ramos-Bossini AJ, Moreno-Suárez S, Pérez-García MDC, Gálvez-López R, Garrido-Sanz F, Rivera-Izquierdo M. Use of clinical probability scores and D-dimer in suspected pulmonary embolism. Results from a Spanish retrospective study. *Eur Soc Emerg Med*. 2020;27(6):468-469. doi:10.1097/MEJ.0000000000000722
7. Singh B, Mommer SK, Erwin PJ, Mascarenhas SS, Parsaik AK. Pulmonary embolism rule-out criteria (PERC) in pulmonary embolism--revisited: a systematic review and meta-analysis. *Emerg Med J*. 2013;30(9):701-706. doi:10.1136/emermed-2012-201730
8. van der Hulle T, Cheung WY, Kooij S, et al. Simplified diagnostic management of suspected pulmonary embolism (the YEARS study): a prospective, multicentre, cohort study. *Lancet Lond Engl*. 2017;390(10091):289-297. doi:10.1016/S0140-6736(17)30885-1
9. Kearon C, de Wit K, Parpia S, et al. Diagnosis of pulmonary embolism with d-dimer adjusted to clinical probability. *N Engl J Med*. 2019;381(22):2125-2134. doi:10.1056/NEJMoa1909159
10. Kline JA, Mitchell AM, Kabrhel C, Richman PB, Courtney DM. Clinical criteria to prevent unnecessary diagnostic testing in emergency department patients with suspected pulmonary embolism. *J Thromb Haemost*. 2004;2(8):1247-1255. doi:10.1111/j.1538-7836.2004.00790.x
11. Abdulmalak C, Cottenet J, Beltramo G, et al. Haemoptysis in adults: a 5-year study using the French nationwide hospital administrative database. *Eur Respir J*. 2015;46(2):503-511. doi:10.1183/09031936.00218214
12. Freund Y, Cachanado M, Aubry A, et al. Effect of the pulmonary embolism rule-out criteria on subsequent thromboembolic events among low-risk emergency department patients: the PROPER randomized clinical trial. *JAMA*. 2018;319(6):559-566. doi:10.1001/jama.2017.21904
13. Penaloza A, Soulié C, Moumneh T, et al. Pulmonary embolism rule-out criteria (PERC) rule in European patients with low implicit clinical

probability (PERCEPIC): a multicentre, prospective, observational study. *Lancet Haematol*. 2017;4(12):e615-e621. doi:10.1016/S2352-3026(17)30210-7

14. Dronkers CEA, van der Hulle T, Gal GL, et al. Towards a tailored diagnostic standard for future diagnostic studies in pulmonary embolism: communication from the SSC of the ISTH. *J Thromb Haemost*. 2017;15(5):1040-1043. doi:10.1111/jth.13654
15. Behringer W, Freund Y. Clinical translation of diagnostic studies: pitfalls of the usual reported characteristics. *Eur Soc Emerg Med*. 2021;28(3):165-166. doi:10.1097/MEJ.0000000000000830
16. Freund Y, Roussel M, Kline J, Roy PM, Bloom B. The failure rate does not equal the false-negative rate: a call for tailoring diagnostic strategy validation in low prevalence populations. *J Thromb Haemost*. 2021;19(7):1832-1833. doi:10.1111/jth.15353
17. Freund Y, Chauvin A, Jimenez S, et al. Effect of a diagnostic strategy using an elevated and age-adjusted D-dimer threshold on thromboembolic events in emergency department patients with suspected pulmonary embolism: a randomized clinical trial. *Jama*. 2021;326(21):2141-2149. doi:10.1001/jama.2021.20750
18. Pernod G, Caterino J, Maignan M, et al. D-dimer use and pulmonary embolism diagnosis in emergency units: why is there such a difference in pulmonary embolism prevalence between The United States of America and countries outside USA? Fukumoto Y, ed. *PLOS one*. 2017;12(1):e0169268. doi:10.1371/journal.pone.0169268
19. Wells PS, Anderson DR, Rodger M, et al. Derivation of a simple clinical model to categorize patients probability of pulmonary embolism: increasing the models utility with the SimpliRED D-dimer. *Thromb Haemost*. 2000;83(3):416-420.
20. Le Gal G, Righini M, Roy PM, et al. Prediction of pulmonary embolism in the emergency department: the revised Geneva score. *Ann Intern Med*. 2006;144(3):165-171. doi:10.7326/0003-4819-144-3-200602070-00004

How to cite this article: Bannelier H, Gorlicki J, Penaloza A, et al. Evaluation of the "hemoptysis" item in clinical decision rules for the diagnosis of pulmonary embolism in the emergency department. *Acad Emerg Med*. 2022;29:1205-1212. doi: 10.1111/acem.14574

Received: 4 May 2022 | Revised: 16 June 2022 | Accepted: 23 June 2022

DOI: 10.1111/acem.14555

ORIGINAL CONTRIBUTION

Periosteal block versus intravenous regional anesthesia for reduction of distal radius fractures: A randomized controlled trial

Sierra Beck MD[1,2] | Alana Brunner-Parker MBChB[2] | Rosemary Stamm PhD[1] | Micheal Douglas MBChB[2] | Aileen Conboy MBChB[2]

[1]Department of Medicine, University of Otago, Dunedin, New Zealand

[2]Dunedin Hospital, Southern District Health Board, Dunedin, New Zealand

Correspondence
Sierra Beck, Department of Medicine, Otago Medical School, University of Otago, Dunedin 9054, New Zealand.
Email: sierra.beck@otago.ac.nz

Funding information
Heath Research South Research Support for Newly Appointed Clinical Academics

Abstract

Objective: We compare periosteal block and intravenous regional anesthesia (IVRA) as anesthetic techniques for reduction of distal radius fractures when performed by emergency department (ED) clinicians following brief training.

Methods: This was a single-center, nonblinded randomized controlled trial of a convenience sample of patients presenting with distal radius fractures requiring closed reduction. Primary outcome measure was patient reported fracture reduction pain score, rated on a 100-mm visual analog scale. Secondary outcomes included adjunct pain medication use, ED length of stay, remanipulation rates, participant satisfaction, clinician assessed efficacy, and clinician-assessed ease of the procedure.

Results: Eighty-one patients were randomized to receive IVRA ($n = 41$) or periosteal block ($N = 40$). Reduction pain scores were not normally distributed. Median (25th–75th percentile) pain scores in participants assigned to IVRA and periosteal block were 5 (1–27.5) and 26 (8.5–63) mm, respectively, ($p = 0.007$). Use of adjunct medications during reduction was higher for the periosteal block group compared with IVRA (57.5% vs. 22.5%, $p = 0.003$). Remanipulation rates were 17.5% for periosteal block versus 7.5% for IVRA ($p = 0.31$). There was no difference in length of stay, patient satisfaction, or clinician's assessed ease of the anesthetic technique. There was a difference in clinician's assessment of efficacy between groups, with IVRA described as "extremely effective" by 65% and periosteal block described as "extremely effective" by 25% ($p = 0.003$).

Conclusions: When performed by a diverse group of ED clinicians periosteal block provided inferior analgesia to IVRA but may provide an alternative when IVRA cannot be performed.

Supervising Editor: James Miner, MD

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2022 The Authors. *Academic Emergency Medicine* published by Wiley Periodicals LLC on behalf of Society for Academic Emergency Medicine.

INTRODUCTION

Distal radius fractures are a common injury treated in the emergency department (ED). Effective anesthesia is required for reduction of displaced fractures. Several techniques are utilized, including hematoma blocks, intravenous regional anesthesia (IVRA; also known as Bier block), regional nerve blocks, and procedural sedation. A Cochrane review of anesthetic techniques concluded that there are insufficient comparative studies.[1] The few studies that have been conducted suggest that IVRA results in lower pain scores and more adequate reduction in comparison with a hematoma block.[1–4] However, IVRA is more resource-intensive requiring monitoring for local anesthetic systemic toxicity and may require a longer ED length of stay,[1] given that the cuff must remain inflated for 30 min following anesthetic injection.

Tageldin et al.[5] described a novel technique known as a periosteal block (Figure 1) in a series of 42 patients with distal forearm fractures. This technique involves injection of local anesthetic around the periosteum of the distal radius proximal to the fracture. In this study, blocks were performed by two experienced orthopedic doctors. None of the 42 patients required remanipulation and the reduction was described as painless by 83% of patients and minimally painful by 14% of patients. Furthermore, no additional analgesia was required during the reduction. To date no randomized controlled trials have been performed comparing the periosteal block to other commonly used techniques nor has it been studied in the hands of a more diverse group of emergency medicine clinicians.

The periosteal block potentially offers an anesthetic option that requires fewer ED resources and more efficient treatment. It is also a relatively straightforward technique to learn and requires no specialized equipment. This study aims to assess the effectiveness of periosteal block in comparison to IVRA, which was standard of care in our ED at the time of this study.

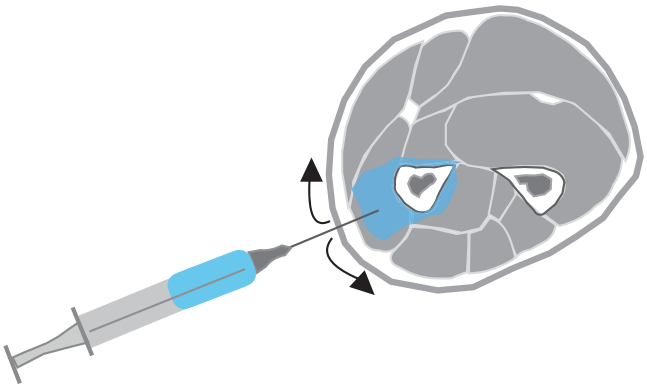


FIGURE 1 Periosteal block technique: 15 ml of 1% lidocaine is inserted 6 cm proximal to the wrist joint. A small wheal is placed under the skin with a 25-gauge needle, and then a 20-gauge needle is used to infiltrate around the radius in close proximity to the bone. By pulling traction on the skin a single injection site can be used to access the radial, dorsal, and volar surfaces of the radius. When an ulnar styloid fracture is present the process is repeated on the ulna using 3 ml of anesthetic.

Our hypothesis was that periosteal block performed by emergency physicians after brief training provides equivalent anesthesia compared to IVRA during distal radius fracture reduction. The primary aim of the study was to compare reduction procedural pain scores using visual analog scale (VAS)[6] in patients randomized to either IVRA or periosteal block. The secondary aims were to compare adjunct pain medication use during the reduction, remanipulation rates, clinician assessed ease of the anesthetic technique, clinician-assessed efficacy of the anesthetic technique, patient satisfaction, ED length of stay, and adequacy of reduction as assessed by blinded orthopedic review.

METHODS

Study design

We performed a single-center nonblinded, randomized, controlled trial comparing the efficacy of periosteal block to IVRA in a convenience sample of patients presenting with displaced distal radius fractures requiring closed reduction. The trial was registered with the Australia New Zealand Clinical Trials registry (ACTRN12618000226202) and approved by the New Zealand Health and Disability Ethics Committee (Ref 18/NTB/11). The trial protocol was reviewed by the Ngāi Tahu Research Consultation Committee. The study was designed and reported in keeping with CONSORT guidelines[7] for reporting of randomized controlled trials. Adverse events were monitored by the primary investigator (SB). All participants provided written informed consent.

Study setting and population

Dunedin Hospital ED is a metropolitan, regional referral center with an annual census of approximately 45,000 patient visits. Trained ED registrars and consultants who were the treating clinicians performed patients' screening, enrollment, study procedures, and data collection. Subjects were enrolled when a study investigator was available.

We included adult patients ≥16 years of age who presented with a displaced distal radius fracture requiring closed reduction. A convenience sample of participants was recruited when a trained clinician was present in the ED during the period between May 2018 and November 2020. Patients were excluded if they were unable to provide informed consent; had an open fracture, evidence of compartment syndrome, a known allergy to local anesthetic, a history of sickle cell disease or Raynaud's, a severe distracting injury, severe peripheral vascular disease, a systolic blood pressure >200 mm Hg, or severe hepatic failure; or needed bilateral manipulation.

Study protocol

All consultants and trainees in the ED were offered training on how to perform the periosteal block and enroll patients in the study.

Participating physicians received a 60-min training session on the periosteal block and IVRA protocols, which concluded with a practical session using a gelatin forearm model and an assessment of competency.

Participants were randomized to receive either IVRA or periosteal block. Group allocations were performed in blocks of 20 through random computer-generated numbers with 1:1 allocation to each group. Assignments were placed in sealed envelopes and opened by the clinician after consent was obtained to participate in the study.

Guides with instructions for how to perform each procedure (Appendix S1) were placed in these sealed envelopes. IVRA was performed using a double-cuff pneumatic tourniquet and 3 mg/kg 0.5% prilocaine injected intravenously through a peripheral cannula in the hand of the injured extremity. Periosteal blocks were performed using the protocol described by Tageldin et al.[5] using 15 ml or 1% lidocaine inserted 6 cm proximal to the wrist joint. A small wheal was placed under the skin with a 25-gauge needle at the lateral aspect of the radius. This was exchanged for a 20-gauge needle that was used to infiltrate around the lateral aspect of the radius in close proximity to the bone (Figure 1). By pulling traction on the skin, the same injection site was used to direct the needle along the dorsal and volar aspects of the radius without removing the needle. If an ulnar styloid fracture was present the process was repeated, and a second injection was performed around the ulna using 3 ml of lidocaine. Both arms of the study included a 10-min wait time between administration of anesthetic and reduction of the fracture to allow time for the block to take effect. During the reduction, clinicians were able to provide additional pain medications to patients as clinically indicated if pain was not controlled by the block alone.

Outcome measures

The primary outcome was the difference in reduction procedural pain scores in the two arms of the study measured on a 100-mm VAS immediately following the reduction. Participants were asked to rate their baseline pain prior to the start of either anesthetic technique. Then anesthetic technique and the fracture reduction were performed, and patients were asked to rate the pain of the reduction immediately afterward. Participants were awake during the procedures and could not be blinded to the anesthetic technique they received. The clinicians who performed the procedure were not blinded. Clinicians recorded whether remanipulation was required for each participant, details of the procedure, including medications administered prior to randomization, and any adjunct medications required during the reduction procedure. For each participant, clinicians ranked the ease of performing the anesthetic technique as extremely easy, very easy, somewhat easy, difficult, or extremely difficult. For each participant, clinicians ranked the efficacy of the anesthetic technique as extremely effective, very effective, somewhat effective, or not at all effective. Length of stay was calculated from timestamps in the hospital electronic medical record system. A blinded review of X-rays was performed by an orthopedic surgeon to determine whether the reduction was adequate given the fracture pattern present or whether they would have repeated an attempt at closed reduction in ED, in which case the reduction was classified as inadequate.

Data analysis

A clinically significant difference in pain scores is context-dependent,[8] and this has not previously been defined for pain during distal radius fracture reduction. Prior ED-based studies of patients with acute traumatic pain have defined scores between 9 and 13 mm as the minimally clinically significant difference.[9–11] Based on this we estimated 10 mm to be the minimum clinically significant difference in pain score to show equivalence of periosteal blocks for pain control during manipulation.[12] Based on the prior study by Tageldin et al.,[5] the expected standard deviation in VAS in this population was 15 mm. Using 80% power and type 1 error rate of 0.05, we estimated that we would need a total sample size of 78 participants.

The a priori primary outcome was comparison of mean fracture reduction pain scores. However, summary statistics for the primary outcome variable did not show a normal distribution and thus we compared and report medians and (25th–75th percentile). An intention-to-treat design was used and differences between the randomized groups were assessed using the Wilcoxon rank sum test due to nonnormality of the data. The main outcome was also analyzed in a per-protocol analysis. Summary statistics for secondary outcome variables are presented as mean ± SD or *n* (%). Secondary outcomes were compared using chi-square tests or Fisher's exact tests where appropriate for categorical data and Student's *t*-test for numeric data. All data were analyzed in R (version 3.6.1) using R studio (version 1.2.5019). A two-sided p value of <0.05 was considered statistically significant. A post hoc analysis was performed in per-protocol fashion to compare mean change in pain scores from baseline given baseline pain scores were slightly higher in the periosteal arm compared to IVRA.

RESULTS

Out of 125 people screened, we enrolled 81 participants and randomized them to either IVRA or periosteal block (Figure 2). One participant randomized to the IVRA group was excluded because data forms were not completed; therefore, 40 participants from each group were included for intention-to-treat analyses. Of the participants randomized into the IVRA group three were switched to receive periosteal block: two because the discomfort from the cuff was too severe to proceed with injection of prilocaine and in one participant because the cuff did not fit their arm. One participant randomized to the periosteal block group received IVRA because the pain was too severe after the block to start the reduction. Therefore, 38 participants in the IVRA and 42 participants in the periosteal block arm were included for per-protocol analyses.

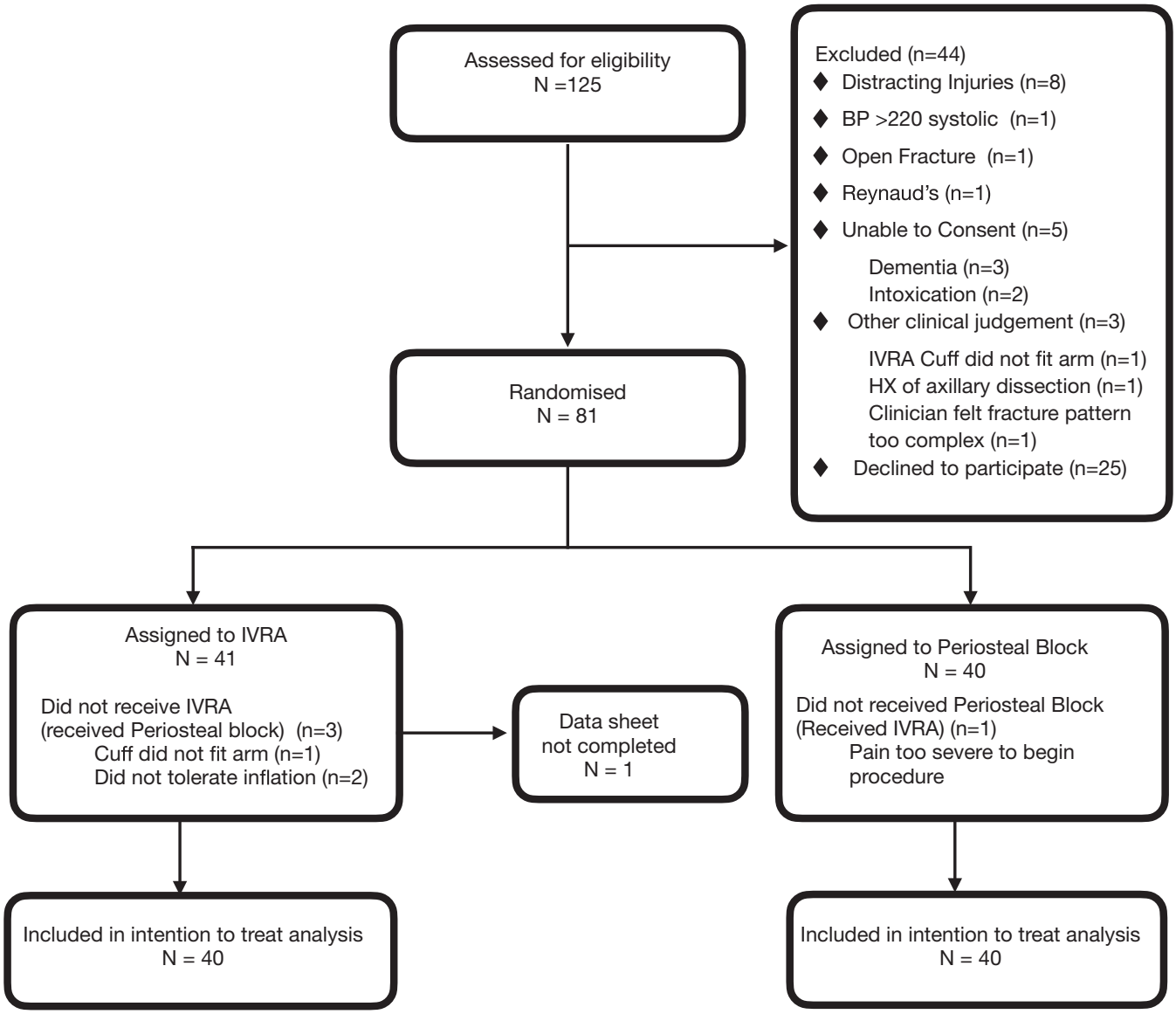


FIGURE 2 Participant flow diagram. HX, history; IVRA, intravenous regional anesthesia.

Baseline characteristics of the participants according to their randomized groups are presented in Table 1. Pain scores prior to reduction were normally distributed, with a mean ± SD of 49.4 ± 26.5 mm and ranged from 0 to 100 mm. Both groups had a high rate of premedication use. The most commonly administered premeditations were paracetamol and fentanyl in both groups. There were 27 different clinicians involved in the trial, the median number of participants enrolled by each clinician in the trial was two and ranged from one to 13. The median (range) number of prior anesthetic procedures performed by the clinicians was 15 (0–300) in the IVRA group and 3.5 (0–20) in the periosteal block group.

For the primary outcome procedural VAS, Figure 3 depicts the prereduction (Figure 3A) and postreduction (Figure 3B) pain scores in both groups. Pain score data are evenly distributed across the measure at prereduction. There is notable skewing of the data for the pain scores in the IVRA group following the reduction, with most of the data centered around low pain scores, indicating a nonnormal distribution. Comparison of median pain scores showed that there was a statistically significant difference in the median pain scores reported between participants assigned to receive IVRA compared with participants assigned to receive the periosteal block (median [25th–75th percentile] 5 [1–27.5] mm and 26 [8.5–63] mm, respectively, $p = 0.007$). Per-protocol analysis revealed similar findings between participants who received IVRA compared with participants who received the periosteal block (median [25th–75th percentile] 3 [1–23.8] mm and 27.5 [10–63] mm, respectively, $p = 0.001$).

Table 2 presents the findings of the secondary outcomes of the study. There were no statistically significant differences between groups for the rate of remanipulation ($p = 0.31$), length of stay ($p = 0.96$), participant satisfaction ($p = 0.11$), or the clinician's assessment of the ease of the anesthetic technique ($p = 0.07$). The clinician's assessment of the efficacy of the anesthetic technique differed between groups ($p = 0.003$). Using IVRA, the efficacy of the anesthetic was described as "extremely effective" 65% of time,

TABLE 1 Demographic and baseline characteristics by randomization group

	IVRA ($n = 40$)	Periosteal block ($n = 40$)
Age (years)	55.8 ± 19.6	59.2 ± 18.6
Gender		
Female	30 (75)	33 (82.5)
Male	10 (25.0)	7 (17.5)
Ethnicity		
European	34 (85.0)	37 (92.7)
Māori	5 (12.5)	1 (2.5)
Other	1 (2.5)	2 (5.0)
Use of premedication	36 (90.0)	31 (77.5)
Paracetamol	18 (45.0)	14 (35.0)
Ibuprofen	8 (20.0)	7 (17.5)
Codeine	9 (22.5)	5 (12.5)
Morphine	9 (22.5)	7 (17.5)
Fentanyl	12 (30.0)	11 (27.5)
Midazolam	1 (2.5)	0 (0.0)
Patient-administered inhaled nitrous oxide/oxygen	5 (12.5)	5 (12.5)
Pain before the procedure	45.3 ± 27.1	53.6 ± 25.7

Note: Data are reported as mean ± SD or *n* (%).
Abbreviation: IVRA, intravenous regional anesthesia.

while the periosteal block was described as "extremely effective" only 25% of the time. Use of adjunct medications during the reduction was significantly higher for the periosteal block group compared with IVRA group (57.5% vs. 22.5%, respectively, $p = 0.003$). On blinded orthopedic review four participants (10%) in the periosteal block group had inadequate reduction compared to zero in the IVRA group. Given that a zero creates numeric inaccuracy in statistical analysis, Fisher exact test was not performed.

Post hoc analysis of change in pain score from baseline showed a mean reduction of 25.3 mm for IVRA compared to 15.3 mm in the periosteal block arm ($p = 0.23$). One participant in the periosteal block arm experienced vasovagal syncope following injection and recovered without intervention. There were no other adverse events.

DISCUSSION

This randomized controlled trial demonstrated lower median pain scores for IVRA compared to periosteal block when performed by a diverse group of ED clinicians following brief training. Together with the secondary outcomes this study suggests that periosteal block did not provide equivalent procedural pain control given the higher median pain scores, the lower clinician-assessed efficacy, and the increased use of adjunct medications in the periosteal block group compared with IVRA group. When performed by ED clinicians with brief training, the periosteal block resulted in less favorable pain control compared to IVRA.

Tageldin et al.[5] found that 83% of patients receiving a periosteal block reported no pain during fracture reduction and 14% reported minimal pain, defined as pain scored between 1 and 3 on a 10-point VAS, in a case series of 42 patients. Wan Ali et al.[13] also reported on using the periosteal block, in a case series of 19 patients where four patients reported no pain, 12 reported minimal pain, and three reported a pain score of 4/10, which was slightly above the criteria for minimal pain. Their study was also performed by a small group of orthopedic specialists. These findings are in contrast with our own, where the median pain score was 26 mm and just under half of patients in the periosteal group reported a pain score above 30 mm, compared with less than a quarter in the IVRA group. Tageldin et al. also reported that no patients required remanipulation within their study, whereas seven of 40 participants in the periosteal arm of our study required remanipulation.

There are several possibilities for the apparent discrepancies between our study and that of Tageldin et al. The use of a different VAS may lead to differing answers by patients in reporting pain scores. It is possible that discrepant findings are the result of inadequate training and experience performing the periosteal block by clinicians in our study. Pain control achieved through periosteal block may improve with experience. Although clinicians at our center were more experienced with IVRA, the median number of prior procedures was still only 15. Being able to identify and cannulate a peripheral vein and apply a pneumatic cuff tourniquet are the main technical skills required for success of IVRA. As these are common skills among ED physicians, the success of IVRA is likely a less operator- and experience-dependent technique. In contrast the periosteal block requires performing a ring block adjacent to the periosteum of the radius, which was novel to clinicians. Anatomic differences such as the depth of subcutaneous tissues at the wrist and other patient variation may have impacted the success of the block in those with less experience. We did not have enough participants in our study to stratify pain scores based on clinician experience with the technique.

Effective analgesia is needed to achieve adequate reduction during manipulation. In this study, four of 40 (10%) of participants in the periosteal block group were defined as having inadequate reduction on orthopedic review, while no participants had inadequate reduction in the IVRA group. There was no statistically significant difference in remanipulation rates between the two groups, albeit a higher proportion of participants in the periosteal block arm underwent remanipulation. The rate of inadequate reduction and remanipulation rates were secondary outcomes of the study and sample size calculations were based on the primary outcome of VAS pain scores. It is possible that a larger trial may have shown significant differences between groups in remanipulation rates given the relationship between effective analgesia and adequate reduction. As part of the study protocol, participants received adjunct medications to manage pain during the reduction. We found that a considerably higher proportion of participants required adjunct medications in the periosteal block group during the manipulation.

Emergency clinicians tailor selection of anesthetic techniques for reduction of distal radius fractures considering several factors

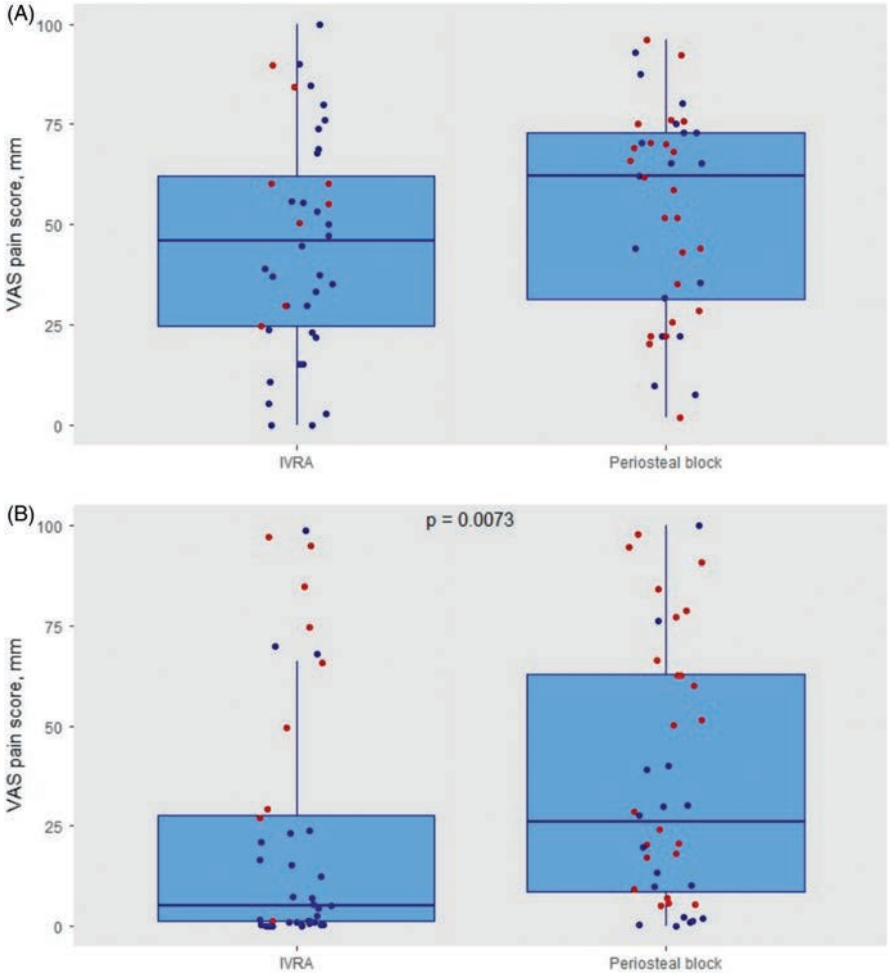


FIGURE 3 Box-and-whisker plots showing the distribution of VAS pain scores at baseline (A) and reduction (B) in the randomized groups. Individual data points are shown on the plots in blue where adjunct medications were not used and in red where adjunct medications were used. *p*-values were calculated for postreduction pain scores between groups using the Wilcoxon rank sum test. IVRA, intravenous regional anesthesia; VAS, visual analog scale.

including patient preferences and characteristics, department staffing and resources, and their own skill and experience. We were surprised that there was no difference in length of stay, which we felt would be a potential benefit of periosteal block. It is likely that factors unrelated to the reduction (i.e., time in waiting room, waits for discharge planning, waiting with nursing staff with casting skills) were more important determinants of length of stay.

In this study the periosteal block with use of adjunct medications was performed safely and facilitated successful reduction in the majority of participants. Given that there are several contraindications to IVRA, the periosteal block may provide an alternative in these situations. Further research comparing periosteal block to hematoma block would help to better define the role for this technique. Similar to our results, prior studies of hematoma block have shown it to have higher pain scores and increased rates of remanipulation when compared to IVRA,[2] but the relative efficacy of these local infiltration-based techniques is unknown.

LIMITATIONS

This was a single-center study in which participants were enrolled as a convenience sample when a trained clinician was available. Overall clinician familiarity and experience were higher on average with IVRA compared to the periosteal block. In our study a diverse group of clinicians underwent brief training to perform the periosteal block, and it is possible that the periosteal block may have performed better following increased experience with the technique. Baseline pain scores were slightly higher in the periosteal block group (53.6 mm) versus (45.3 mm) in the IVRA group; this must be considered when interpreting the study results.

Adjunct pain medication was not administered via standardized protocol and different doses and types of medications will have had variable impact on reported pain scores. Pain medications used were at the discretion of the treating clinician who could not be blinded to the anesthetic technique they had performed; this may have further biased pain medication administration. Although this may limit direct comparison of the two techniques, it models real-world ED practice where we tailor medication administration to the needs of the individual patient. Pain scores were not measured at multiple time points during the procedure, which may have further helped to show the impact of adjunct medications. Lastly, our study was not powered to detect a difference in rates of remanipulation. As reduction success is related to adequacy of analgesia, a study and analysis plan specifically designed to assess this may have detected a difference in the adequacy of reduction.

TABLE 2 Secondary outcomes

	IVRA $n = 40$	Periosteal block $n = 40$	p value
Remanipulation	3 (7.5)	7 (17.5)	0.311
Adjunct medications	9 (22.5)	23 (57.5)	0.003
Morphine	1 (11.1)	4 (17.4)	—
Fentanyl	6 (66.7)	6 (26.1)	—
Ketamine	1 (11.1)	0 (0.0)	—
Midazolam	1 (11.1)	0 (0.0)	—
Patient-administered inhaled nitrous oxide/oxygen	3 (33.3)	14 (60.9)	—
Length of stay (min)	296 ± 146	295 ± 118	0.961
Inadequate reduction	0 (0.0)	4 (10.0)	—
Patient satisfaction			0.114
Not at all satisfied	1 (2.5)	0 (0.0)	
Partially satisfied	2 (5.0)	3 (7.5)	
Satisfied	7 (17.5)	8 (20.0)	
More than satisfied	5 (12.5)	13 (32.5)	
Extremely satisfied	25 (62.5)	16 (40.0)	
Clinician assessed ease of anesthetic technique			0.066
Extremely easy	12 (30.0)	8 (20.0)	
Very easy	24 (60.0)	20 (50.0)	
Somewhat easy	3 (7.5)	10 (25.0)	
Difficult	0 (0.0)	2 (5.0)	
Extremely difficult	1 (2.5)	0 (0.0)	
Clinician assessed efficacy of the anesthetic technique			0.003
Extremely effective	26 (65.0)	10 (25.0)	
Very effective	6 (15.0)	15 (37.5)	
Somewhat effective	7 (17.5)	12 (30.0)	
Not at all effective	1 (2.5)	3 (7.5)	

Note: Data are reported as *n* (%) or mean ± SD. *p*-values were calculated using Student's *t*-test for numeric variables and chi-square tests or Fisher's exact tests where appropriate for categoric variables.

Abbreviation: IVRA, intravenous regional anesthesia.

CONCLUSIONS

When performed by a diverse group of ED clinicians, periosteal block provided inferior analgesia to intravenous regional anesthesia but may provide an alternative when intravenous regional anesthesia cannot be performed.

AUTHOR CONTRIBUTIONS

Sierra Beck conceived the study. Sierra Beck, Aileen Conboy, and Alana Brunner-Parker designed the trial. Sierra Beck and Alana Brunner-Parker supervised the conduct of the trial and data collection. Sierra Beck and Alana Brunner-Parker managed the data, including quality control. Rosemary Stamm analyzed the data. Sierra Beck and Rosemary Stamm drafted the manuscript, and all authors contributed substantially to its revision. Sierra Beck takes responsibility for the paper as a whole.

ACKNOWLEDGMENTS

The authors thank the patients who participated in the trial and the following clinicians for their contributions enrolling patients in this study: Oliver Aughton, Emel Barbirou, Amr Bin Sadiq, Caroline Collins, Rene Cesson, Conor Cosgrave, Sue Creighton, Kevin Jan, Bruce Lambie, Wan Lim, Ben Nuttall, Adam Pomerleau, Matt Reay, Tim Ritchie, Rebecca Roberts, Helena Robinson, Sean Sanders, Ralph Scorgie, Joshua Smith, Rich Stephenson, Tim Tennyson, Yulia Tychinskaya, Rosa Tobin Steckings, Jacques Van der Mere, Shong Min Voon, and Jeremy Wong. The authors also thank Anne Daniels and Bernie Don who assisted with the study administration. Open access publishing facilitated by University of Otago, as part of the Wiley - University of Otago agreement via the Council of Australian University Librarians. Open access publishing facilitated by University of Otago, as part of the Wiley - University of Otago agreement via the Council of Australian University Librarians.

FUNDING INFORMATION

This research was funded by Heath Research South Research Support for Newly Appointed Clinical Academics.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

ORCID

Sierra Beck https://orcid.org/0000-0002-1614-5058

REFERENCES

1. Handoll HH, Madhok R, Dodds C. Anaesthesia for treating distal radial fracture in adults. *Cochrane Database Syst Rev*. 2002;3:Cd003320.
2. Kendall JM, Allen P, Younge P, Meek SM, McCabe SE. Haematoma block or Bier's block for Colles' fracture reduction in the accident and emergency department--which is best? *J Accid Emerg Med*. 1997;14(6):352-356.
3. Abbaszadegan H, Jonsson U. Regional anesthesia preferable for Colles' fracture. Controlled comparison with local anesthesia. *Acta Orthop Scand*. 1990;61(4):348-349.
4. Cobb AG, Houghton GR. Local anesthetic infiltration versus Bier's block for Colles' fractures. *Br Med J (Clin Res Ed)*. 1985;291(6510):1683-1684.
5. Tageldin ME, Alrashid M, Khoriati AA, Gadikoppula S, Atkinson HD. Periosteal nerve blocks for distal radius and ulna fracture manipulation--the technique and early results. *J Orthop Surg Res*. 2015;10:134.
6. Bijur PE, Silver W, Gallagher EJ. Reliability of the visual analog scale for measurement of acute pain. *Acad Emerg Med*. 2001;8(12):1153-1157.
7. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ*. 2010;340:c332.
8. Olsen MF, Bjerre E, Hansen MD, et al. Pain relief that matters to patients: systematic review of empirical studies assessing the minimum clinically important difference in acute pain. *BMC Med*. 2017;15(1):35.
9. Kelly AM. Does the clinically significant difference in visual analog scale pain scores vary with gender, age, or cause of pain? *Acad Emerg Med*. 1998;5(11):1086-1090.
10. Todd KH, Funk KG, Funk JP, Bonacci R. Clinical significance of reported changes in pain severity. *Ann Emerg Med*. 1996;27(4):485-489.
11. Gallagher EJ, Liebman M, Bijur PE. Prospective validation of clinically important changes in pain severity measured on a visual analog scale. *Ann Emerg Med*. 2001;38(6):633-638.
12. Myles PS, Myles DB, Galagher W, et al. Measuring acute postoperative pain using the visual analog scale: the minimal clinically important difference and patient acceptable symptom state. *Br J Anaesth*. 2017;118(3):424-429.
13. Wan Ali WAA, Soh EZF, Abdullah S, Narin Singh PSG, Ahmad AA, Sapuan J. Pain perception during the phases of manual reduction of distal end radius fracture with a periosteal block. *Cureus*. 2021;13(1):e12691.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Beck S, Brunner-Parker A, Stamm R, Douglas M, Conboy A. Periosteal block versus intravenous regional anesthesia for reduction of distal radius fractures: A randomized controlled trial. *Acad Emerg Med*. 2022;29:1213-1220. doi: 10.1111/acem.14555

Received: 14 February 2022 | Revised: 27 July 2022 | Accepted: 29 July 2022

DOI: 10.1111/acem.14576

ORIGINAL CONTRIBUTION

Changes in scientific characteristics of abstracts accepted to the Society for Academic Emergency Medicine Annual Meeting, 1990–2020

Carlos A. Camargo Jr. MD, DrPH | **Krislyn M. Boggs MPH** | **Rebecca E. Cash PhD, MPH** | **Vishal P. Doshi BS** | **Henry H. Isaacson BA** | **Kohei Hasegawa MD, MPH** | **Ali S. Raja MD, MBA, MPH**

Department of Emergency Medicine, Massachusetts General Hospital, Boston, Massachusetts, USA

Correspondence
Carlos A. Camargo Jr., Department of Emergency Medicine, Massachusetts General Hospital, 125 Nashua Street, Suite 920, Boston, MA 02114, USA.
Email: ccamargo@partners.org

Abstract

Objectives: Since its founding in 1989, the Society for Academic Emergency Medicine (SAEM) has accepted thousands of abstracts for presentation at its annual meeting. We reviewed abstracts to characterize temporal changes in study design, abstract topics, quality scores, and proportion of abstracts published as manuscripts.

Methods: In this serial cross-sectional study, we compiled accepted SAEM abstracts at 5-year intervals (1990, 1995, 2000, 2005, 2010, 2015, 2020) and then randomly selected 100 abstracts from each year for review by two investigators. We documented each abstract's study design, sample size, and whether it was a single-center or multicenter study. We assigned each abstract to the most appropriate topic category. Applying SAEM's abstract scoring system from 2020, we calculated the mean overall quality score per year. Finally, we searched PubMed to determine if abstracts from 1990–2015 meetings were published as manuscripts.

Results: The number of accepted abstracts increased from 180 in 1990 to 879 in 2020 (+388%). The most common study design changed from laboratory study in 1990 (22%) to cohort study in 2020 (44%; $p < 0.001$). The median study sample size increased over time, from 105 (interquartile range [IQR] 25–389) in 1990 to 544 (IQR 102–2067) in 2020 ($p < 0.001$). Multicenter studies have become more common (19% in 1990 vs. 40% in 2020; $p = 0.001$). The most common topic categories also changed from cardiology/pulmonary/airway (40%) and orthopedic/trauma/burn (17%) in 1990 to health services research/health policy/operations (25%) and cardiology/pulmonary/airway (22%) in 2020. There was a 20% increase in overall quality scores ($p < 0.001$). Between 37% and 49% of the abstracts reviewed from each year were later published as manuscripts, with no significant change over time ($p = 0.33$).

Conclusions: Over the past 30 years, there have been significant changes to the study designs, topics, and quality scores of SAEM meeting abstracts. However, conversion of abstracts to published manuscripts remains a challenge.

PresentationsPresented virtually at the Society for Academic Emergency Medicine Annual Meeting, May 2021.

Supervising Editor: Jody Vogel MD, MSc

 © 2022 Society for Academic Emergency Medicine.

INTRODUCTION

Scientific abstracts are frequently presented at the major meetings of professional and academic societies. Importantly, these presentations facilitate the dissemination of new research, often prior to the publication of a scientific manuscript based on that same data.[1] In emergency medicine (EM), one of the major meetings at which abstracts are presented is the Society for Academic Emergency Medicine (SAEM) Annual Meeting.[2]

Briefly, SAEM was formed in 1989 via the merger of the University Association for Emergency Medicine (UAEM) and the Society of Teachers of Emergency Medicine (STEM).[3] This merger was undertaken in recognition of the overlap between research and education and, accordingly, SAEM has always focused on furthering both research and education. Research abstracts are selected for presentation at every SAEM Annual Meeting.[2] As part of the selection process, SAEM initially used a simple 4-item scale for abstract scoring, which was replaced by a more detailed approach in 1998; another iteration, which was made publicly available, was launched in 2015.[4] The evolving scoring system has provided a more objective approach for judging abstract quality and also educated EM researchers on essential aspects of high-quality research.

Although SAEM has included scientific abstract presentations as part of its annual meeting since its founding, it is unclear if the characteristics of these abstracts have changed over time. We aimed to characterize temporal changes in the study design, abstract topics, quality scores, and the proportion of meeting abstracts subsequently published as manuscripts.

METHODS

Study design

In this retrospective study, we reviewed a random sample of abstracts from among all of those accepted to the SAEM Annual Meeting between 1990 and 2020. The Mass General Brigham Human Research Committee reviewed this study and classified it as exempt. This report adheres to the STROBE guidelines.[5]

Data collection

We compiled accepted SAEM abstracts at 5-year intervals, from 1990 to 2020 (i.e., for 1990, 1995, 2000, 2005, 2010, 2015, and 2020). We then randomly selected 100 abstracts from each chosen year for review using = RAND(), the random number feature of Microsoft Excel. Briefly, we assigned a random number between 0 and 1 to every accepted abstract and then sorted these random numbers from lowest to highest; the reviewed abstracts were those with the lowest 100 numbers in each year. We chose 100 abstracts per year because this would provide sufficiently narrow 95% confidence intervals (CIs) for key characteristics to allow meaningful comparisons across years.

Each abstract was reviewed by at least two trained research assistants (VPD, HHI). When there were any discrepancies (<10%), one of the authors (KMB) reviewed the two items to resolve the discrepancy. To keep KMB blinded to the year when reviewing discrepant scores, abstract text was copied over from the source PDFs to a separate document without identifying information; individual abstracts could be identified by searching for the abstract number or title. Additional methods and results describing agreement between trained research assistants and one investigator are available as supplemental materials accompanying the online article (Table S1).

Measurements

For each abstract, we documented the following information: study type (basic, clinical, population [Methods S1]), study design (laboratory, case report, case series, cross-sectional, case–control, cohort, randomized controlled trial, meta-analysis, qualitative, other), study sample size, and whether the study was a single-center or multicenter study.

We also created a list of 18 abstract topic categories from those provided by SAEM over the 30-year period and assigned each abstract to the most appropriate category: abdominal/gastrointestinal/genitourinary, cardiology/pulmonary/airway, education/ethics/faculty, geriatrics/palliative, infectious diseases, international EM, health services research (HSR)/health policy/operations, neurology, obstetrics/gynecology, orthopedic/trauma/burn, pediatrics, emergency medical services (EMS)/disaster/wilderness, psychiatry/social EM, public health, research methods, technology/simulation, toxicology/pharmacology, and other diseases (Methods S2).

Using the latest (2020) version of SAEM's scoring system,[6] we also scored all abstracts on a scale of 0–2 based on seven criteria: clarity of objectives, choice of approach, validity, statistics, scope, importance of topic, and publication readiness. The maximum score (highest quality) is 14 points. Finally, we documented whether there was a corresponding publication in PubMed for abstracts accepted 1990–2015; thus, even the 2015 abstracts had 5 years to achieve manuscript publication. We documented that a publication was directly from the abstract if the content of the manuscript was substantially the same as that of the abstract, even if the title and authors of the manuscript varied slightly from the original SAEM Annual Meeting abstract (Methods S3). Additional methods describing measurements are available as supplemental materials accompanying the online article.

Data analysis

Data analysis included descriptive statistics, presented as proportions or as medians with interquartile ranges (IQRs). We conducted unadjusted analyses using chi-square tests, Fisher's exact test, and

Kruskal-Wallis tests, as appropriate, to assess the association between abstract characteristics and abstract year. Analyses were done using Stata 15.1 (StataCorp). A two-tailed $p < 0.05$ was considered statistically significant. We did not adjust for multiple testing so statistically significant results should be interpreted with caution.

RESULTS

Overall, the number of abstracts accepted to the SAEM Annual Meeting increased from 180 in 1990 to 879 in 2020 (+388%). Of the 7 years reviewed (e.g., 1990, 1995, and so forth), the year with the highest number of abstracts accepted was 2015 (1037; Figure 1).

Study type and design

The proportion of studies that were clinical remained consistent throughout the study period (64% in 1990 vs. 63% in 2020); by contrast, the proportion of basic science studies decreased (20% vs. 2%), and the proportion of population health studies increased (16% vs. 35%; Table 1). Over the study period, the most common study design changed from laboratory study in 1990 (22%) to cohort study in 2020 (44%, $p < 0.001$). Notably, only 1% of abstracts in 2015 and 2020 were laboratory studies. Twenty percent of abstracts in 1990 and 1995 were cohort studies—this proportion rose to 34% in 2000 and remained at least 34% for all subsequent years. Case–control studies decreased over time, from 14% in 1990 to 4% in 2020 (Figure 2, Figures S1 and S2).

Study sample sizes

The median sample size increased over time, from 105 (IQR 25–389) in 1990 to 544 (IQR 102–2067) in 2020 ($p < 0.001$). There were some intervals where the median sample size increased dramatically. Specifically, between 1990 and 1995, the median sample size increased from 105 (IQR 25–389) to 199 (IQR 52–592). Between 2015 and 2020, it increased from 295 (IQR 100–1370) to 544 (IQR 102–2067). Multicenter studies became more common over time (19% of studies in 1990 vs. 40% in 2020, $p = 0.001$; Table 1, Figure S2).

Abstract topic categories

In 1990, the most common topic categories were cardiology/pulmonary/airway (40%), orthopedic/trauma/burn (17%), EMS/disaster/wilderness (16%), and toxicology/pharmacology (14%). In 2020, the most common topic categories were HSR/health policy/operations (25%), cardiology/pulmonary/airway (22%), technology/simulation (18%), and education/ethics/faculty (15%). Looking across the 30-year study period, cardiology/pulmonary/airway remained the most common topic category from 1990 until 2005. From 2010 through 2020, HSR/health policy/operations was the most common topic category. Cardiology/pulmonary/airway has remained the second most common topic category during these more recent years. Other topics that significantly *increased* between 1990 and 2020 included technology/simulation (2% vs. 18%, respectively), psychiatry/social EM (0% vs. 11%), education/ethics/faculty (6% vs. 15%), and infectious diseases (5% vs. 12%, all $p < 0.05$). Other topics that significantly *decreased* between 1990 and 2020 included EMS/disaster/wilderness (16% vs. 3%) and toxicology/pharmacology (14% vs. 7%, both $p = 0.01$; Table 1).

Abstract quality scores

Overall quality scores increased from a median of 10 (IQR 8–11) in 1990 to 12 (IQR 10.5–13) in 2020. Scores increased among the following categories between 1990 and 2020: clarity of objectives, choice of approach, statistics, scope, and publication readiness (all $p < 0.01$). Scores for validity ($p = 0.08$) and importance of topic ($p = 0.42$) did not significantly change between 1990 and 2020 (Table 2).

Between 37% and 49% of the abstracts reviewed from each year were later published as manuscripts, with no significant change over time ($p = 0.33$; Figure S2). Of these, between 21% (2000) and

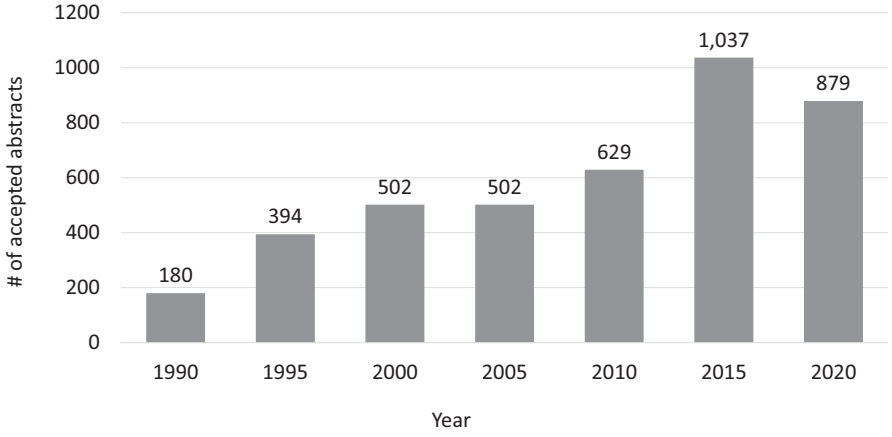


FIGURE 1 Number of abstracts accepted to the SAEM Annual Meeting, 1990–2020.

TABLE 1 Characteristics of abstracts accepted to the SAEM Annual Meeting, 1990–2020

	1990	1995	2000	2005	2010	2015	2020	*p*-value
Total number of accepted abstracts	180	394	502	502	629	1037	879	
Study type								<0.001*
Basic	20	15	10	6	3	0	2	
Clinical	64	58	67	84	68	67	63	
Population	16	27	23	10	29	33	35	
Multicenter study	19	15	27	21	29	34	40	0.001*
Study sample size, median (IQR)	105 (25–389)	199 (52–592)	182 (67–480)	209 (70–658)	271 (91–1797)	302 (100–1500)	544 (102–2067)	<0.001*
Topic (>1 topic possible)								
Abdominal/gastrointestinal/genitourinary	4	3	5	2	4	1	9	0.15
Cardiology/pulmonary/airway	40	31	35	41	15	18	22	<0.001*
Education/ethics/faculty	6	11	9	6	14	20	15	0.02*
Geriatrics/palliative	1	0	1	0	4	8	4	0.002*
Infectious diseases	5	4	3	4	15	12	12	0.002*
International EM	1	0	3	2	4	2	1	0.49
HSR/health policy/operations	10	25	15	27	26	32	25	0.003*
Neurology	6	5	7	8	7	3	4	0.73
Obstetrics/gynecology	0	0	3	0	4	0	2	0.02*
Orthopedic/trauma/burn	17	18	15	15	10	15	10	0.56
Pediatrics	6	7	14	5	9	6	10	0.25
EMS/disaster/wilderness	16	4	7	9	10	5	3	0.01*
Psychiatry/social EM	0	4	9	5	6	7	11	0.04*
Public health	2	6	2	5	5	4	2	0.57
Research methods	4	3	2	1	2	0	3	0.53
Technology/simulation	2	10	6	11	17	15	18	0.002*
Toxicology/pharmacology	14	6	4	2	5	4	7	0.01*
Other diseases	2	3	3	5	7	2	4	0.56
Published manuscript in PubMed	46	42	38	49	49	37	—	0.33
Published within 1 year of abstract	23	24	21	25	23	22	-	
Published 1.1–2 years of abstract	17	11	10	9	8	12		
Published 2.1–3 years of abstract	2	2	4	8	11	3		

Abbreviations: HSR, health services research; IQR, interquartile range.

*$p < 0.05$.

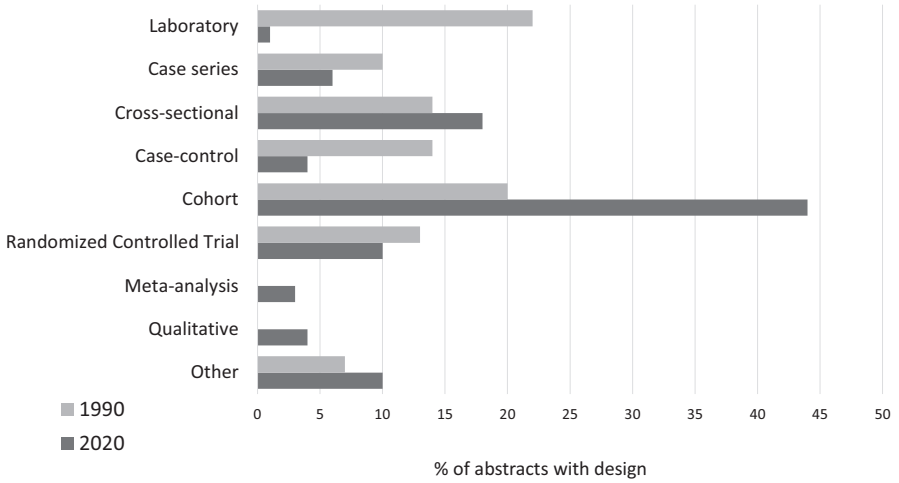


FIGURE 2 Study design of abstracts accepted to the SAEM Annual Meeting in 1990 ($n = 100$) versus 2020 ($n = 100$).

TABLE 2 Changes (%) in scores of abstracts accepted to the SAEM Annual Meeting, 1990–2020

	1990	1995	2000	2005	2010	2015	2020	*p*-value*
Score								
Overall, median (IQR)	10 (8–11)	9 (8–11)	11 (9–12)	10.5 (9.5–12)	11 (9–12)	11 (10–12)	12 (10.5–13)	<0.001*
Clarity of objectives								<0.001*
0	8	6	7	8	10	1	0	
1	37	31	14	21	20	14	19	
2	55	63	79	71	70	85	81	
Choice of approach								<0.001*
0	3	4	2	1	1	2	0	
1	34	28	23	25	20	9	10	
2	63	68	75	74	79	89	90	
Validity								0.08
0	4	10	2	2	5	4	2	
1	39	29	31	33	26	25	26	
2	57	61	67	65	69	71	72	
Statistics								0.001*
0	16	24	10	9	17	7	6	
1	36	31	28	34	30	39	18	
2	47	40	58	57	50	50	72	
Not applicable	1	5	4	0	3	4	4	
Scope								<0.001*
0	28	29	15	21	17	20	16	
1	33	35	36	40	34	27	28	
2	18	18	36	32	43	50	55	
Not applicable	21	18	13	7	6	3	1	
Importance of topic								0.42
0	19	15	15	17	15	9	9	
1	52	58	54	54	61	56	52	
2	29	27	31	29	24	35	39	
Publication readiness								<0.001*
0	1	3	4	2	1	4	1	
1	29	25	24	28	15	11	6	
2	70	72	72	70	84	85	93	

Abbreviation: IQR, interquartile range.

25% (2005) were published within 1 year, 8% (2010) and 17% (1990) published between 1.1 and 2 years of the abstract, and 2% (1990) and 11% (2010) published between 2.1 and 3 years of the abstract (Table 1).

DISCUSSION

From 1990 to 2020, there was a clear increase in the number of scientific abstracts accepted for presentation at the SAEM Annual Meeting. Moreover, there were important changes in the most common study designs, increases in both sample size and the proportion of studies that were multicenter, and changes in most common topic categories. There was also an increase in overall abstract quality scores, although not in every scoring category. There was also no significant change in the proportion of abstracts subsequently published as manuscripts. To our knowledge, this is the first study to assess the change in characteristics of abstracts accepted to the SAEM Annual Meeting over the three decades since the Society's founding.

Our results are similar to those found in a 1999 study by Singer and colleagues.[7] Briefly, they reviewed UAEM abstracts from 1974 and 1983, then SAEM abstracts from 1989 and 1997. They noted changes in abstract characteristics and quality even in the early years after SAEM's founding. For example, they found an increase in the proportion of studies with power calculations and tests of significance, which would likely mean an increase in "statistics" score and overall quality score, as we found. They similarly found an increase in study sample size during the early years of SAEM (25 [IQR 24–182] in 1989 vs. 129 [IQR 46–416] in 1997).

We also found similar results to Porturas and Taylor,[8] who investigated abstracts of EM articles from 1980 to 2019 within six highly cited journals. Although our topic groupings were not identical, Porturas and Taylor found a decrease in research involving cardiac arrest over the study period, while we found a decrease in abstracts focused on cardiology/pulmonary/airway. Similar to our results, Porturas and Taylor also found a large decrease in the proportion of abstracts classified as basic science. On the other hand, Porturas and Taylor found only a modest increase (0.6%) in the proportion of abstracts classified as mental health and substance abuse research, while we found an 11% increase in abstracts classified as psychiatry/social EM over the study period. The inclusion of social EM in our topic category may account for part of this. Further, there was a relatively high number of psychiatry/social EM abstracts in 2020, which was after the study period used in the Porturas and Taylor publication.

One area that requires improvement is the proportion of meeting abstracts that are later published as full manuscripts. Although there probably is no one "right" answer, we estimate that that at least 80% of abstracts should proceed to manuscript publication. We found that approximately one-third to one-half were later published as manuscripts over the study period. Other studies similarly found that less than half of SAEM abstracts were published as manuscripts. For example, Callaham and colleagues[9] found that 41% of SAEM abstracts in 1991 were published as manuscripts, similar to the 46% we found to be published from the 1990 meeting. Li and colleagues reviewed all SAEM abstracts from 1997 and from 1999–2001. They found that between 35% and 40% were published as manuscripts per year.[10] Among abstracts accepted in 2000, they found that 38% were published manuscripts (compared to the 40% we found published in our sample from 2000).[10] They found that 19% to 24% were published within 1 year (compared to our 1990–2015 review, which found 21%–24% published within 1 year), 12%–17% 1 year later (compared to 8%–17%), and 10%–11% 2 years later (compared to 2%–11%).[10] Likewise, Gottlieb and colleagues[11] reviewed abstracts accepted to multiple EM conferences, including SAEM, from 2011 to 2015. They found that 49% of SAEM abstracts from during this time were later published as manuscripts. This was the same proportion of abstracts that we identified as published manuscripts in 2010 (49%) but higher than the proportion in 2015 (37%).

In related work, Clemency and colleagues[12] examined EMS abstracts and publication following abstract presentation at multiple meetings, including SAEM, and found that 53% EMS abstracts from SAEM were later published. Adhikari and colleagues[13] examined publication rates of ultrasound, EMS, and toxicology abstracts submitted to both SAEM and ACEP from 1998 to 2008. Publication rates were low (23%, 23%, and 21%, respectively), suggesting little difference by topic.

The publication rate for SAEM abstracts since 1990 is middling compared to those of abstracts accepted to non-EM meetings. For example, publication rates of SAEM abstracts were *lower* than those accepted to: the Dutch Surgical Society meetings (69%),[14] American Society of Therapeutic Radiology and Oncology (56%),[15] Orthopedic Society of North America (53%),[16] Skeletal Society of Radiology (51%),[17] and the American Heart Association (50%).[18] Publication rates of SAEM abstracts were *similar* to those of abstracts accepted to the American College of Cardiology (43%), the European Society of Cardiology (38%),[18] and the North American Menopause Society (38%).[19] On the other hand, publication rates of SAEM abstracts were *higher* than those abstracts accepted to the American Society of Gastrointestinal Endoscopy (25%)[20] and two wound conferences (12%).[21] Even the highest proportions suggest ample room for improvement.

LIMITATIONS

Given the subjective nature of some of the data collection (e.g., the SAEM scoring system), it is possible that a different set of reviewers would have documented different results for these items. However, our approach was aligned with SAEM's actual scoring system, where each abstract is assigned to multiple reviewers. Furthermore, any discrepancies, between both the subjective and the objective items, were further reviewed for reconciliation. While it is possible that information bias was introduced (e.g., data abstracters were aware of the year of each abstract), the investigator who resolved scoring discrepancies in the SAEM scoring was blinded to the abstract

year. Another possible limitation relates to the need to collapse topic categories given the substantial changes in potential abstract topics from 1990 to 2020, but these decisions were made with input of multiple experienced researchers, and these arguably more "general" categories still convey a clear change in common topic categories over time. Finally, we note that many of the characteristics overlap (e.g., multicenter studies generally will have larger sample sizes) but we did not perform multivariable analyses to address this relatedness. Our goal was to describe what happened over the past 30 years and the observed trends are true regardless of which factor stands out in a stratified or multivariable model. We speculate that increases in collaborative research and in the availability of large administrative datasets underlie some of the observed results. We also speculate that widespread adoption of the SAEM abstract scoring criteria has led to significant improvements in the quality of the accepted research.

CONCLUSIONS

Research presented at the Society for Academic Emergency Medicine Annual Meeting has changed significantly over the past 30 years. Overall, studies have become larger and are now more likely to be multicenter, and the most common topics have changed over time. Further, the scientific quality of abstracts has improved, reflected by an increase in scores over time. One area for improvement is the subsequent publication of accepted abstracts as full manuscripts. We look forward to witnessing (and participating) in the continued evolution of research presented at the Society for Academic Emergency Medicine Annual Meeting, as this research will remain a key component of emergency medicine research dissemination worldwide.

AUTHOR CONTRIBUTIONS

Carlos A. Camargo and Krislyn M. Boggs conceived and designed the study. Krislyn M. Boggs, Vishal P. Doshi, Henry H. Isaacson, and Ali S. Raja acquired the data. Carlos A. Camargo and Rebecca E. Cash provided statistical advice. Rebecca E. Cash performed the analysis. Carlos A. Camargo, Krislyn M. Boggs, Rebecca E. Cash, Kohei Hasegawa, and Ali S. Raja interpreted the data. Carlos A. Camargo and Krislyn M. Boggs drafted the article. All authors critically revised the manuscript for important intellectual content and approved the final version. Carlos A. Camargo takes responsibility for the paper as a whole.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

ORCID

Carlos A. Camargo Jr. https://orcid.org/0000-0002-5071-7654
Krislyn M. Boggs https://orcid.org/0000-0002-8493-2731
Rebecca E. Cash https://orcid.org/0000-0002-0355-1014

REFERENCES

1. Kelly JA. Scientific meeting abstracts: significance, access, and trends. *Bull Med Libr Assoc*. 1998;86:68-76.
2. Society for Academic Emergency Medicine. Meetings and events. https://www.saem.org/meetings-and-events/meetings-and-events. Accessed: February 11, 2022.
3. Society for Academic Emergency Medicine. About SAEM. https://www.saem.org/about-saem. Accessed: February 11, 2022.
4. Kuczmarski TM, Raja AS, Pallin DJ. How do medical societies select science for conference presentation? How should they? *West J Emerg Med*. 2015;16:543-550.
5. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol*. 2008;61:344-349.
6. Society for Academic Emergency Medicine. Submitting your abstract to the SAEM Annual Meeting. https://www.saem.org/docs/default-source/annual-meeting/2019-annual-meeting/education/abstract-scoring-8-22-18.pdf?sfvrsn=ad3b09fd_2. Accessed: February 11, 2022.
7. Singer AJ, Homan CS, Brody M, Thode HCJ, Hollander JE. Evolution of abstracts presented at the annual scientific meetings of academic emergency medicine. *Am J Emerg Med*. 1999;17:638-641.
8. Porturas T, Taylor RA. Forty years of emergency medicine research: uncovering research themes and trends through topic modeling. *Am J Emerg Med*. 2021;45:213-220.
9. Callaham M, Weber E, Wears R. Citation characteristics of research published in emergency medicine versus other scientific journals. *Ann Emerg Med*. 2001;38:513-517.
10. Li SF, Umemoto T, Crosley P, Cassidy C. SAEM abstracts to articles: 1997 and 1999-2001. *Acad Emerg Med*. 2004;11:985-987.
11. Gottlieb M, Ryan K, Alcorn T, et al. From presentation to paper: assessment of successful abstract publications in emergency medicine over a five-year period. *Acad Emerg Med*. 2021;28:679-682.
12. Clemency BM, Thompson JJ, Lindstrom HA, Gurien S, Jaison BA, Grates-Sciarrino AA. Frequency of manuscript publication following presentation of EMS abstracts at national meetings. *Prehosp Disaster Med*. 2014;29:294-298.
13. Adhikari S, Blaivas M, Frrokaj I, Shostrom V. Meeting abstracts to published manuscripts: how does emergency ultrasound compare? *J Ultrasound Med*. 2011;30:1275-1279.
14. de Meijer VE, Knops SP, van Dongen JA, Eyck BM, Vles WJ. The fate of research abstracts submitted to a national surgical conference: a cross-sectional study to assess scientific impact. *Am J Surg*. 2016;211:166-171.
15. Papagikos MA, Rossi PJ, Lee WR. Publication rate of abstracts from the annual ASTRO meeting: comparison with other organizations. *J Am Coll Radiol*. 2005;2:72-75.
16. Jackson KR, Daluiski A, Kay RM. Publication of abstracts submitted to the annual meeting of the pediatric Orthopaedic Society of North America. *J Pediatr Orthop*. 2000;20:2-6.
17. Tritz D, Bautista L, Scott J, Vassar M. Conversion of skeletal Society of Radiology annual meeting abstracts to publications in 2010-2015. *Peer J*. 2018;6:e5817.
18. Fosbøl EL, Fosbøl PL, Harrington RA, Eapen ZJ, Peterson ED. Conversion of cardiovascular conference abstracts to publications. *Circulation*. 2012;126:2819-2825.
19. Schnatz PF, Romegialli A, Abrantes J, Marakovits K, Cunningham D, O'Sullivan DM. The North American Menopause Society: from abstract to publication. *Menopause*. 2008;15:996-1001.
20. Eloubeidi MA, Wade SB, Provenzale D. Factors associated with acceptance and full publication of GI endoscopic research originally published in abstract form. *Gastrointest Endosc*. 2001;53:275-282.

21. Dumville JC, Petherick ES, Cullum N. When will I see you again? The fate of research findings from international wound care conferences. *Int Wound J*. 2008;5:26-33.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Camargo CA, Boggs KM, Cash RE, et al. Changes in scientific characteristics of abstracts accepted to the Society for Academic Emergency Medicine Annual Meeting, 1990–2020. *Acad Emerg Med*. 2022;29:1221-1228. doi: 10.1111/acem.14576

Received: 6 December 2021 | Revised: 17 March 2022 | Accepted: 25 March 2022

DOI: 10.1111/acem.14492

SYSTEMATIC REVIEWS

Screening tools to identify patients with unmet palliative care needs in the emergency department: A systematic review

Scott W. Kirkland MSc[1] | Esther H. Yang MSc[1] | Miriam Garrido Clua MSc[2] | Maureen Kruhlak BSc[3] | Sandra Campbell MLS[4] | Cristina Villa-Roel MD, PhD[1] | Brian H. Rowe MD, MSc[1,5]

[1]Department of Emergency Medicine, Faculty of Medicine & Dentistry, University of Alberta, Edmonton, Alberta, Canada

[2]Department of Orthopaedic Surgery, St. Joseph's Healthcare Hamilton, Hamilton, Ontario, Canada

[3]School of Veterinary Medicine, St. George's University, West Indies, Grenada

[4]J.W. Scott Health Sciences Library, University of Alberta, Edmonton, Alberta, Canada

[5]School of Public Health, University of Alberta, Edmonton, Alberta, Canada

Correspondence
Brian H. Rowe, Department of Emergency Medicine, University of Alberta, 1G1.43 WMC, 8440-112 Street NW, Edmonton, Alberta T6G 2B7, Canada.
Email: brian.rowe@ualberta.ca

Funding information
Funding for this review was provided by the Ministry of Health, Government of Alberta. Dr. Rowe's research is supported by a Scientific Director's grant (SOP 168483) from the Canadian Institutes of Health Research (CIHR) through the Government of Canada (Ottawa, ON). These funding organizations had no involvement in any aspect of the conduct, analysis, and manuscript preparation of this study; CIHR takes no responsibility for the conduct or results of this review.

Supervising Editor: Dr. Tammie Quest

Abstract

Objectives: This systematic review identified and assessed psychometric properties of the available screening tools to identify patients with unmet palliative care (PC) needs in the emergency department (ED).

Methods: A comprehensive search of electronic databases and the gray literature was conducted. Two independent reviewers completed study screening and inclusion, data extraction, and quality assessment. A descriptive summary of the results was reported using median of medians and interquartile ranges (IQRs).

Results: A total of 35 studies were included, involving the assessment of 14 unique screening tools. The most commonly used screening tool was the surprise question (SQ; $n = 12$ studies), followed by the Palliative Care and Rapid Emergency Screening (P-CaRES) tool ($n = 8$), and the screening for palliative and end-of-life care needs in the emergency department (SPEED) instrument ($n = 4$). Twelve of the included studies reported on the psychometric properties of the screening tools, of which eight of these studies assessed the performance of the SQ to predict patient mortality. Overall, the median sensitivity (63%, IQR 38%–78%) and specificity (75%, IQR 57%–84%) of the SQ to predict mortality at 1 or 12 months was moderate. While the median positive predictive value of the SQ was low (32%, IQR 16%–40%), the median negative predictive value was high (91%, IQR 88%–95%). Across the studies, the proportion of patients identified as having unmet PC based on the criteria of the screening tools ranged from 5% to 83%.

Conclusions: This review identified 14 unique screening tools used to identify adult patients with unmet PC needs in the ED. One screening tool, the SQ, was found to have moderate sensitivity and specificity to accurately predict future patient mortality. Additional research is needed to better understand the clinical value of this and the other available tools prior to their widespread implementation.

KEYWORDS
emergency department, end of life, palliative care, screening tools

 © 2022 Society for Academic Emergency Medicine.

INTRODUCTION

With the increase in emergency department (ED) visits among patients with advanced or end-of-life (EOL) conditions,[1,2] early integration of palliative care (PC) services into the ED has been recommended by initiatives like the Choosing Wisely Campaign.[3] For many patients, the ED offers a unique opportunity to initiate these connections; however, identifying patients with unmet PC needs and finding the time to initiate discussions in a chaotic, fast-paced environment such as the ED can be challenging. As such, several screening instruments/tools have been developed for the purpose of assisting emergency physicians and other health care providers to identify patients with unmet PC needs, including the surprise question (SQ)[4] and the Palliative Care and Rapid Emergency Screening (P-CaRES) tool.[5,6]

Despite the availability of multiple screening tools to identify patients with unmet PC needs in the ED, the performance of these tools has not been well established making the recommendation for one particular tool over the others difficult.[7] With the publication of various new studies, this systematic review conducted an updated assessment of the available screening tools to identify ED patients with unmet PC needs in the ED.

METHODS

Protocol

A study protocol was developed and registered on PROSPERO (registration #CRD42018086178) to define the objectives, selection criteria, data collection, and analysis. This review followed the recommendations of the PRISMA guidelines for reporting of systematic reviews.[8]

Inclusion/exclusion

Studies assessing the ability of a screening tool to identify adult patients with unmet PC needs in the ED were eligible for inclusion in this review. To be eligible studies had to apply the screening tool on adult patients presenting to the ED. Studies did not need to utilize a control group or reference standard nor to report on the psychometric properties of the screening tools to be included in the review. While prospective and retrospective studies were eligible for inclusion, reviews, editorials, or case studies were excluded.

Outcome

The primary outcome of interest pertained to the performance of the screening tools, including psychometric properties such as sensitivity, specificity, positive predictive values (PPV), negative predictive values (NPV), and/or likelihood ratios (LRs). Secondary outcomes of interest included metrics on the reliability and validity of the tools, proportion of patients with unmet PC needs, referrals to PC consultations or hospice services, disposition of index visit (e.g., discharge or hospitalizations), mortality (either within the ED or following ED discharge), and subsequent health care utilization (e.g., hospitalization, ED visits).

Search methods for identification of studies

A search was conducted by an experienced health research librarian (S.C.) on the following databases: OVID Medline, Ovid EMBASE, OVID Health and Psychosocial Instruments, EBSCO-CINAHL, SCOPUS, ProQuest Dissertations and Theses Global, Cochrane Library, and PROSPERO using controlled vocabulary and key words representing the concepts "emergency department" and "palliative care" and "screening." Searches were adjusted appropriately for different databases. Searches were last updated in August 2021. No publication or language limits were applied. Results were exported to RefWorks citation management system. Detailed search strategies are available in Appendix 1.

A search of the gray literature consisted of a search of Google Scholar, clinical trial registries (e.g., clinicaltrials.gov, Cochrane Central Register of controlled trials, and controlled-trials.com), backward and forward SCOPUS searches of included studies, bibliographies from known reviews, and abstracts from emergency medicine conferences including Canadian Association of Emergency Physicians (2015–2021), Society for Academic Emergency Medicine (2015–2021), and the American College of Emergency Physician Abstracts (2015–2021).

Selection of studies

The title and abstract of studies identified in the literature search were screened for eligibility by two independent reviewers (M.G.C., M.K., E.H.Y., or S.W.K.). The full text of studies identified as potentially eligible were retrieved and reviewed for inclusion based on predefined inclusion criteria by two independent reviewers (M.G.C., M.K., E.H.Y., or S.W.K.). Disagreements were resolved via consensus.

Data extraction

Two independent reviewers (M.G.C., M.K., E.H.Y., or S.W.K.) extracted predefined outcomes onto standardized electronic forms. Disagreements were resolved via discussion. If a study reported outcomes such as mortality over multiple time points, only the last time point was considered in the meta-analysis.

Quality assessment

For studies reporting the psychometric properties of the screening tools, the methodologic quality was assessed using the Quality

Assessment of Diagnostic Accuracy Studies (QUADAS)[9] and Standards for Reporting Diagnostic accuracy studies (STARD)[10] tools. The quality of observational cohort studies that did not assess the psychometric properties was assessed using the Newcastle-Ottawa Scale (NOS).[11] Quality assessment was conducted by independent author reviewers and disagreements were resolved via consensus. Two authors (E.H.Y., S.W.K.) reviewed 34 of the 35 studies, and the one study[12] for which S.W.K. was the lead author was instead reviewed by E.H.Y. and J.M.

Data synthesis

A descriptive summary of the results was reported. Summary tables were created to report on the study characteristics, screening tools, psychometric properties, and other outcomes. Medians and interquartile ranges (IQRs) were calculated, and a median of medians approach was taken for reporting the "average" effect. No sensitivity or subgroup analyses were conducted. While some studies reported outcomes based on ED visits or presentations,[12,13] for the purposes of this review, the unit of analysis was based on individual patient data. If available, performance of screening tools was assessed based on the positive (+) or negative (−) LRs, with +/− LRs ~0.5 to 2 rated as unhelpful (poor performance), mid-range ($5 < +LR < 10$; $0.5 < -LR > 0.1$) rated as moderate performance, and the extremes ($+LR > 10$; $-LR < 0.1$) rated as excellent performance.[14]

RESULTS

Search results

Once duplicates were removed, a total of 5491 title and abstracts were screened following database and gray literature searches (see Figure 1). A total of 5237 studies were found to be irrelevant and excluded from the study, resulting in the full-text review of 254 studies. The full texts of 219 studies were excluded for various reasons (see Figure 1), resulting in the inclusion of 35 studies in the review.

Study characteristics

The majority of studies were conducted in the United States (see Table 1). Twelve of the included studies were conference abstracts.[15-26] The majority of studies reported screening unmet PC needs in ED patients with specific EOL conditions or patients requiring admission (see Table 1).[12,13,15-21,24-42] Across the 35 included studies, a total of 14 unique screening tools were assessed (see Table 1 and Appendix 2). The majority of studies did not utilize another screening tool as a reference standard, except for three studies that used screening tools that had been previously used to identify unmet PC needs in patients in other clinical settings.[37,43,44] Other studies utilized medical or death records as a reference standard for the assessment of the accuracy of the screening tool to predict patient mortality.[4,19,27-29,41,42,44,45] One study[42] and one conference abstract[18] were found to utilize a similar database of patients. Both studies reported different objectives and outcomes, except for one overlapping outcome (proportion of patients with unmet PC needs). To reduce the risk of duplicate data, only the proportion of patients with unmet PC needs as reported in the published study[42] was included in the analysis.

Screening tool characteristics

SQ

The most commonly assessed screening tool was the SQ (see Table 1). While most of the included studies applied the original SQ (Would you be surprised if the patient died in the next 12 months?),[4,16,19,28,44,45] some studies utilized a modified SQ (Would you be surprised if the patient died in the next 30 days?).[15,17,18,27,29,42] A negative response (e.g., the physician would not be surprised if the patient died during a given time frame) was suggestive that the patient had unmet PC needs. Assessment of unmet PC needs was reported to have been completed by the attending emergency physicians in all but two of the included studies (see Table 1).[16,44] A small number of studies applied the SQ to patients with a specific clinical condition including heart failure,[45] advanced cancer,[28] chronic obstructive pulmonary disease (COPD),[29] and sepsis,[17] while the majority of studies applied the SQ to patients without a specific clinical condition.[4,15,16,18,19,27,44] Some studies further specified that they strictly applied the SQ to patients who were to be admitted to hospital,[15,17,18,27,29,42] admitted to the intensive care unit (ICU),[19] or to any patients older than 65[4] or 70 years.[44]

P-CaRES tool

Eight studies[12,13,20,21,30,31,32,33] utilized the P-CaRES tool, which consists of a list of potential EOL conditions including advanced dementia or central nervous system disease, cancer, COPD, chronic renal failure, end-stage liver disease, septic shock in patients over 65 years old, and other conditions with a high chance of accelerated death.[5] If the patient presented with at least one EOL condition, the assessors completed the second domain, selecting risk factors for unmet PC needs from a checklist that included previous ED visits or hospitalizations in the past 6 months, uncontrolled symptoms, functional decline, uncertainty about goals of care and/or caregiver distress, and a negative response to the SQ.[5] Based on the criteria of the P-CaRES tool, patients were identified as having unmet PC needs if they presented with at least one EOL condition and two or more risk factors for unmet PC needs. Some studies modified either the checklist of EOL conditions listed on the screening tool[12] or the list of risk factors for unmet PC needs (see Appendix 2).[32] Three studies strictly applied the P-CaRES tool on older or geriatric patients.[13,30,33] For the majority of studies, emergency physicians completed the

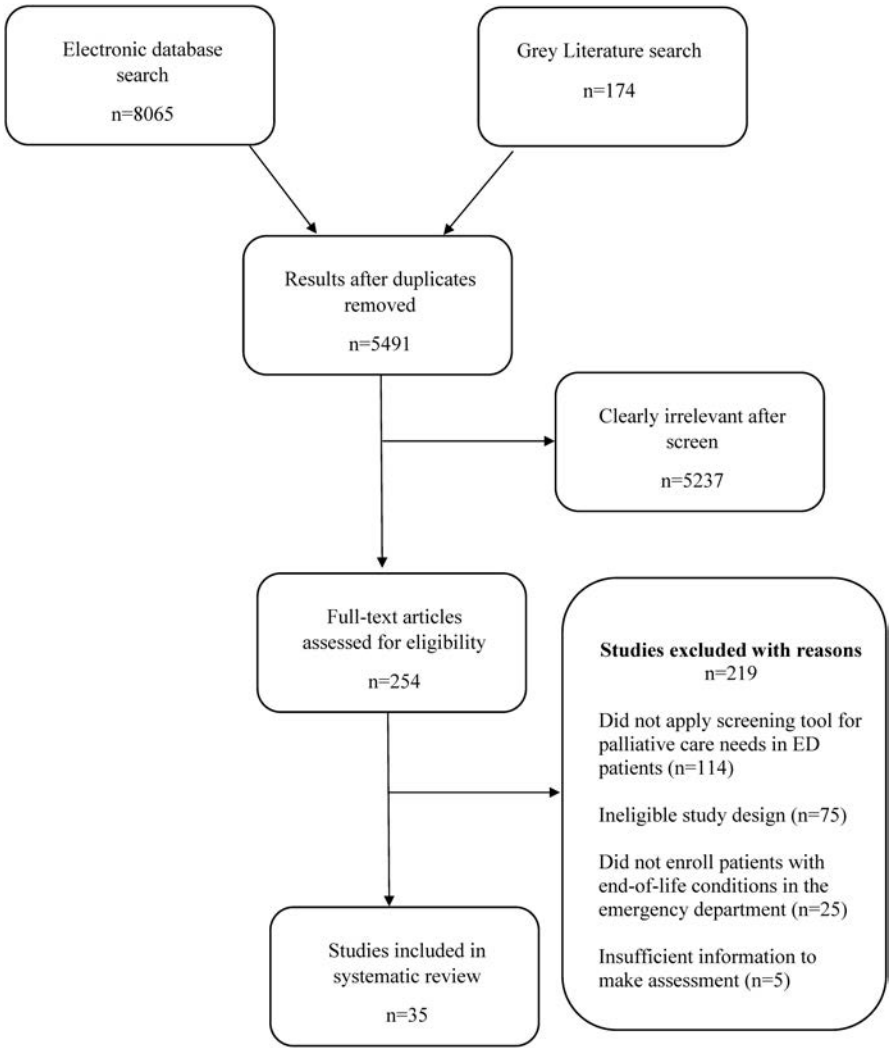


FIGURE 1 Literature search flow diagram for a systematic review on studies assessing screening tools to identify patients with unmet PC needs in the ED. PC, palliative care

P-CaRES tool,[12,21,30,31,32] with one study comparing the accuracy of screening with the P-CaRES tool between emergency physicians and PC chaplain fellows (see Table 1).[13]

Screening for palliative needs in the emergency department (SPEED) instrument

Two studies used the original SPEED instrument, which consists of 13 items assessing social, therapeutic, physical, psychological, and spiritual needs.[22,43] Unlike many of the other screening tools, the original SPEED tool does not have prespecified criteria for a patient to be considered to have "unmet PC needs," but, rather, aims to assist physicians identify specific needs of the patient. Two studies utilized a modified five-item version of the SPEED instrument, which includes five questions regarding difficulties in pain management, home care, medication management, psychological support, and goals of care.[23,46] The items of the five-item SPEED tool are scored on a Likert scale, however, the threshold score for each domain to identify a patient with unmet PC needs was not reported.[23,46] Three studies applied the SPEED instrument on patients with active cancer[23,43,46] while one study only screened patients with end-stage heart failure.[22] Research assistants[43] and ED nurses[23,46] were reported to have completed the SPEED instrument except in one study which did not specify who completed the screening (see Table 1).[22]

Other screening tools

The 11 remaining studies each assessed a different screening tool[24,25,26,34-41] and additional details are provided in Appendix 2. The majority of the screening tools consisted of a checklist of questions, such as the simplified screening tool (SST) included physicians completing two domains, the first of which consisting of an 11-point scale assessing ambulation, activity, disease progression, self-care, oral intake, and level of consciousness, as well as a checklist in which physicians identified at least one indicator of unmet PC needs.[37] Another study assessed whether various risk factors for mortality (age, quick sequential organ failure assessment score, cancer, performance score, and do-not-resuscitate status), recognized as the A-qCPR risk score, could identify patients with unmet PC needs.[41] Two studies utilized a battery of available screening tools for assessing symptom management, frailty, quality of life, as well as assessments of physical, psychological, and financial burden of patients and their

TABLE 1 Characteristics of studies assessing screening tools to identify patients with unmet PC needs in the ED

Study ID	Screened participants	No. of patients	Assessor completing the screening tool	PC screening tool
Surprise Question (SQ)				
Bond 2014[15] United States	ED patients admitted to hospital	4478	ED physicians	Modified SQ - 30 days
Yoshida 2014[16] United States	ED patients with serious illness	NA	Not specified	SQ
Strout 2016[17] United States	ED patients with sepsis admitted to hospital	330	ED physicians	Modified SQ - 30 days
Strout 2017[18] United States	ED patients admitted to hospital	10,334	ED physicians	Modified SQ - 30 days
Glick 2018[19] United States	Nontrauma patients being admitted to an ICU	313	ED physicians	SQ
Ouchi 2018[4] United States	ED patients aged 65 years and older	207	ED physicians	SQ
Ouchi 2019[42] United States	ED patients aged 65 years and older admitted to hospital	10,737	ED physicians	Modified SQ - 30 days
Aaronson 2019[45] United States	Patients with CHF	193	ED physicians	SQ
Haydar 2019[27] United States	Hospitalized ED patients	6122	ED physicians	Modified SQ - 30 days
Verhoef 2020[28] Netherlands	Adult patients with end-stage cancer	245	ED physicians	SQ
Tripp 2021[29] United States	Patients admitted to the hospital for an acute exacerbation of COPD	381	ED physicians	Modified SQ - 30 days
Gaffney 2022[44] Ireland	ED patients aged 70 years and older	191	Non-consultant hospital doctors working in the ED (not specified)	SQ
Palliative Care and Rapid Emergency Screening (P-CaRES) tool				
Ouchi 2017[30] United States	ED patients aged 65 years and older with at least one EOL condition including advanced dementia or CNS disease, end-stage renal disease, advanced COPD, advanced CHF, end-stage liver disease, septic shock, and provider discretion (high chance of accelerated death)	207	ED physicians	P-CaRES
Duffy 2019[20] Canada	ED patients (order than 18 years) with a triage score of 2–5 who were triaged to the subacute or acute areas of the ED with at least one EOL condition including advanced dementia or CNS disease, end-stage renal disease, advanced COPD, advanced CHF, end-stage liver disease, septic shock, and provider discretion (high chance of accelerated death)	417	ED nurses and physicians	P-CaRES
Kostenberger 2019[31] Austria	Adult ED patients with at least one EOL condition including advanced dementia or CNS disease, end-stage renal disease, advanced COPD, advanced CHF, end-stage liver disease, septic shock, and provider discretion (high chance of accelerated death)	195	ED physicians	P-CaRES
Yang 2020[32] Taiwan	Acute, critically ill adult patients over 20 years old who were to be admitted to the emergency ICU following their ED visit with at least one EOL condition including advanced dementia or CNS disease, end-stage renal disease, advanced COPD, advanced cardiovascular disease, end-stage liver disease, severely frail, septic shock, acute respiratory disease syndrome, multiple organ failure, or impending death	796	ED physicians	Modified P-CaRES
Howard 2021[13] United States	Geriatric patients over the age of 80 with at least one EOL condition including advanced dementia or CNS disease, end-stage renal disease, advanced COPD, advanced CHF, end-stage liver disease, septic shock, and provider discretion (high chance of accelerated death)	58	ED physicians and PC chaplains-fellows	P-CaRES
Kirkland 2021[12] Canada	Adults ED patients with at least one EOL condition including advanced dementia, progressive CNS disease, end-stage renal disease, advanced COPD, advanced CHF, and end-stage liver disease	663	ED physicians	Modified P-CaRES

(Continues)

TABLE 1 (Continued)

Study ID	Screened participants	No. of patients	Assessor completing the screening tool	PC screening tool
Te Paske 2021[33] United States	ED patients aged 55 years who were to be admitted to hospital with at least one EOL condition including advanced dementia or CNS disease, end-stage renal disease, advanced COPD, advanced CHF, end-stage liver disease, septic shock, and provider discretion (high chance of accelerated death)	131	Coinvestigators of physicians from various specialties	P-CaRES
Tolia 2020[21] United States	ED patients with at least one EOL condition including advanced dementia or CNS disease, end-stage renal disease, advanced COPD, advanced CHF, end-stage liver disease, septic shock, and provider discretion (high chance of accelerated death)	641	ED physicians	P-CaRES
Screening for palliative and EOL care needs in the emergency department (SPEED) instrument				
Richards 2011[43] United States	ED patients over 21 years old with active cancer, English speaking, and capable of providing informed consent	53	Research assistants	The SPEED instrument
Lo 2016[22] United States	Adult patients with CHF	81	Not specified	The SPEED instrument
Marshall 2017[23] United States	Adult patients with active cancer	293	ED nurse	The 5-item SPEED instrument
Reuter 2019[46] United States	Adult ED patients with active cancer	817	ED nurse	The 5-item SPEED instrument
Other screening tools				
Mahony 2008[34] United States	Patients with progressive end-stage illness aged 65 years and older, with multiple organ failure but rejected for ICU admission, was hospice eligible, and required access to community services	291	Nurse practitioner	PC trigger tool
Grudzen 2011[35] United States	Patients 65 years or over with functional deficit and a history of either cancer requiring chemotherapy, radiation, and surgery other than local excision; CHF; end-stage liver disease; chronic renal failure on hemodialysis; stroke; oxygen-dependent pulmonary disease; or dementia of any time	50	Trained interviewer (not specified)	Battery of tests including: NEST-13, ESAS, MQOL, and the Caregiver Strain Index
Glajchen 2011[36] United States	Elderly patients older than 65 years old with life-limiting illness (including CNS, COPD, CHF, AIDS)	1587	ED physicians	Two-stage BriefPal screening protocol.
Minor 2013[24] United States	ED patients waiting for admission to hospital	1554	ED staff unspecified	IPAL-EM screening tool
Earle 2014[25] United States	Patients with severe dementia, CNS disease with progressive decline, metastatic cancer, AIDS with progressive decline, CHF with recurrent admissions or requiring home oxygen, severe advanced disease, living in a nursing home with unclear care goals, or severe decline in functional status	155	Not specified	Unnamed palliative screening tool
Schulman 2015[26] United States	ED patients suitable for ICU admission	52	Not specified	7-item PC screening
Cotogni 2017[37] Italy	Seriously ill patients with either chronic organ failure, progressive neurologic disease, or advanced cancer who are waiting to be admitted to hospital	257	Physician specialist in anesthesia and intensive care	Simplified Screening Tool (SST)
Koh 2019[38] Singapore	Adult patients with metastatic cancer who require home hospice care or admission to an inpatient palliative unit or hospice	340	ED physicians and members of the PC team	Modified screening checklist (ED-PALS)

TABLE 1 (Continued)

Study ID	Screened participants	No. of patients	Assessor completing the screening tool	PC screening tool
Bakan 2020[39] Turkey	Patients who had been diagnosed with diseases of the heart, kidney, and respiratory system along with cancer, stroke, and other conditions and diseases that shorten the lifespan	208	Not specified	The Criteria for Receiving PC
Tineo 2020[40] Peru	Adult patients aged 18 years or older with advanced disease	60	ED physicians	Unnamed tool
Wang 2021[41] Taiwan	Any patients aged 15 years or older who were admitted to hospital through the ED	4627	Not specified	A-qCPR risk score

Abbreviations: AIDS, acquired immune deficiency syndrome; CHF, congestive heart failure; CNS, central nervous system; COPD, chronic obstructive pulmonary disease; EOL, end of life; ESAS, Edmonton symptom assessment scale; HCS, home health care hospice and community services; ICU, intensive care unit; MQOL, McGill quality of life questionnaire; NEST-13, needs near the end-of-life care screening tool; PC, palliative care; QOL, quality of life; SQ, surprise question.

caregivers.[35,36] Various personnel completed the screening including emergency physicians,[36,40] nurse practitioners,[34] a specialist in anesthesia and intensive care,[37] unspecified ED staff,[24] or a trained interviewer (see Table 1).[35]

Risk-of-bias assessment

Thirteen studies were assessed using the STARD and QUADAS tools.[4,13,15,19,27-29,37,41-45] Based on the STARD criteria, the majority of studies clearly reported assessing the methods for participant sampling, the reference standard, the patient characteristics, and the tools clinical applicability (see Appendix 3). The majority of studies, however, did not report whether assessors were blinded to the outcomes of the index test or reference standard, nor provided a cross-tabulation of the results, test reproducibility, or accuracy variability. The majority of studies sufficiently reported on most of the criteria for the QUADAS tool including the patient population, the screening tools and the reference standard, and uninterpretable results (see Appendix 4).

The quality of 22 studies were assessed using the NOS (see Appendix 5).[12,16,17,18,20-26,30-36,38-40,46] The majority of these studies included a representative cohort of participants and sufficient follow-up. The majority of studies consisted of a single cohort, and studies with multiple groups did not report on controlling or adjusting for any potential differences among the groups. Only half of the studies reported how the outcomes were assessed.

Primary outcome

Performance

Twelve of the 35 included studies reported the psychometric properties of the PC screening tools. From these 12 studies, various psychometric properties of the SQ were reported in nine of the studies,[4,15,19,27-29,42,44,45] with the performance of the P-CaRES tool,[13] the SST,[37] and A-qCPR[41] being assessed in three different studies. None of the included studies reported the psychometric properties of the SPEED tool or the other screening tools identified in this review.

Of the nine studies assessing the SQ, eight assessed its ability to predict patient mortality at 30 days[27,29,42] or 12 months[4,19,28,44,45] respectively using either official death records or the patients' medical charts as the reference standard (see Table 2, Figure 2). The overall median sensitivity of the SQ to predict mortality was 63% (IQR 38%–78%), while the median specificity was 75% (IQR 57%–84%). The overall median PPV of the SQ was 32% (IQR 16%–40%) while the median NPV was 91% (IQR 88%–95%). The sensitivity of the SQ to predict patient mortality at 30 days was lower (range 12%–43%)[27,29,42] than at 12 months (range 57%–89%;[4,19,28,44,45] see Figure 2). Similarly, the specificity of the SQ was higher among studies assessing patient mortality at 30 days (range 82%–95%)[27,29,42] than at 12 months (range 40%–79%).[4,19,28,44,45] One study reported that the SQ had a +LR of 2.68 (95% confidence interval [CI] 0.68–10.64) and a −LR of 0.92 (95% CI 0.76–1.11) for predicting 30-day mortality,[29] while another study reported a +LR of 2.5 (95% CI: 1.7, 3.7) and a −LR of 0.56 (95% CI: 0.4, 0.8) for predicting 12 month mortality.[44] One study reported that the modified-SQ had poor predictive properties for mortality at 6 and 12 months (see Table 2).[42] One study reported that the SQ had poor predictive properties for admission from the ED, prolonged length of stay (LOS), frailty, and readmission to hospital (see Table 3),[44] while another study reported that the SQ had low sensitivity (35%), high specificity (89%), and high NPV (97%) when used to predict the use of PC resources.[15]

Psychometric properties of three other screening tools were reported. First, the SST reported high sensitivity (98%, 95% CI 92%–99%), specificity (93%, 95% CI 88%–96%), and accuracy (95%) to identify ED patients with unmet PC needs when compared with five criterion from the Italian Society of Anesthesia, Analgesia, Resuscitation, and Intensive care (SIAARTI).[37] Another study compared the ability of chaplains trained in PC assessment to use the P-CaRES tool to accurately identify patients with unmet PC needs compared to emergency physicians, which found a trend for more false positives by chaplains (see Table 2).[13] Finally, the A-qCPR tool had excellent sensitivity (99, 95% CI 98%–99%) and NPV

TABLE 2 Performance of screening tools to identify patients with unmet PC needs in the ED

Study ID	Target condition	Screening tool	Reference standard	Outcome	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)	Accuracy, % (95% CI)	AUC, (95% CI)	Validity
SQ											
Bond 2014[15]	Admitted patients unspecified	SQ	Electronic health records	Patients receiving PC consultation	35	89	NA	97	NA	NA	NA
Glick 2018[19]	ICU admission unspecified nontrauma	SQ	Medical record	Inpatient 12-month mortality	68	71	32	92	70	NA	NA
Ouchi 2018[4]	≥65 years unspecified	SQ	Death records from Massachusetts Department of Health Vital Records	12-month mortality	77	56	32	90	NA	0.72	NA
Ouchi 2019[42]	≥65 years admitted patients unspecified	SQ	Death records from the National death index	30-days mortality	43	82	20	93	78	0.73 (0.72–0.74)	NA
				6-month mortality	30	83	37	84	75	NA	NA
				12-month mortality	30	77	36	85	73	NA	NA
Aaronson 2019[45]	Heart failure	SQ	Death records from the Massachusetts Department of Health	12-month mortality	79	57	43	87	NA	NA	NA
Haydar 2019[27]	Admitted patients unspecified	SQ	Medical record	Inpatient 30-days mortality	32 (25–40)	85 (84–86)	6 (4–7)	98 (98–98)	NA	NA	NA
Verhoef 2020[28]	End-stage cancer	SQ	Survival data from the Dutch population register	12-month mortality	89 (84–93)	40 (27–55)	85 (82–88)	50 (37–63)	NA	0.56 (0.53–0.63)	NA
Tripp 2021[29]	COPD	SQ	Mortality data from electronic medical records, HealthInfoNet, and online obituary searches	30-days mortality	12 (2–38)	95 (93–97)	11 (1–33)	96 (94–98)	NA	NA	NA

TABLE 2 (Continued)

Study ID	Target condition	Screening tool	Reference standard	Outcome	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)	Accuracy, % (95% CI)	AUC, (95% CI)	Validity
Gaffney 2022[44]	≥70 years unspecified	SQ	Electronic health records	12-month mortality	57 (39–74)	79 (70–83)	36 (23–50)	89 (82–94)	NA	0.67 (0.58–0.76)	Moderate-strong correlation between SQ and CBS (r =−0.35) and RISC (r =−0.68). Poor correlation between SQ and CSF and EQ-5D (<0.1)
				Frailty[α]	41 (32–50)	88 (78–94)	84 (72–92)	49 (40–58)	NA	0.64 (0.58–0.70)	
				Admission	33 (26–42)	84 (70–93)	88 (76–95)	27 (20–36)	NA	0.59 (0.52–0.65)	
				Prolonged LOS	43 (31–55)	79 (67–88)	68 (53–81)	56 (46–66)	NA	0.60 (0.52–0.68)	
				Hospital readmission	48 (29–68)	73 (63–81)	30 (17–46)	85 (76–92)	NA	0.60 (0.50–0.71)	
P-CaRES											
Howard 2021[13]	≥80 years unspecified	P-CaRES	Emergency physicians' screening	Agreement between the chaplain and emergency physician's screening	NA	NA	Chaplain: 34% Physician: 44%	Chaplain: 76% Physician: 88%	NA	NA	NA
SPEED											
Richards 2011[43]	Cancer	SPEED	NEST	Performance of SPEED compared to domains of NEST	NA	NA	NA	NA	NA	NA	Cronbach coefficient for survey scales (0.716–0.991), indicating strong correlation.
Other screening tools											
Cotogni 2017[37]	Admitted patients End-stage organ failure, progressive neurologic disease, or cancer	Simplified screening tool (SST)	The Italian Society of Anesthesia, Analgesia, Resuscitation, and Intensive Care (SIAARTI) tool	Agreement with SIAARTI to identify patients in need of PC	98 (92–99)	93 (88–96)	NA	NA	95 (91–97)	NA	NA

(Continues)

TABLE 2 (Continued)

Study ID	Target condition	Screening tool	Reference standard	Outcome	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)	Accuracy, % (95% CI)	AUC, (95% CI)	Validity
Wang 2021[41]	Admitted patients unspecified	A-qCPR	Electronic health records	12-month mortality	Overall, 99 (98–99) Low risk group (≤4 points), 5 (4–7) Intermediate-risk group (4 < score ≤ 9 points), 55 (51–58) High-risk group (>9 points), 40 (37–43)	Low-risk group (≤4 points), 52 (50–54) Intermediate-risk group (4 < score ≤ 9 points), 43 (41–44) High-risk group (>9 points), 6 (5–6)	Low-risk group (≤4 points), 1.7 (1–2.4) Intermediate-risk group (4 < score ≤ 9 points), 21 (20–23) High-risk group (>9 points), 65 (60–69)	Overall, 99 (98–100) Low-risk group (≤4 points), 98 (97–98) Intermediate-risk group (4 < score ≤ 9 points), 76 (74–78) High-risk group (>9 points), 36 (32–40)	NA	Overall Trained cohort, 0.84 (0.83–0.86) Cross-validation, 0.84 (0.83–0.85)	NA

Abbreviations: AUC, area under the curve; CBS, caregiver burden score; CFS, clinical frailty score; COPD, chronic obstructive pulmonary disease; ICU, intensive care unit; NA, not applicable; NEST, needs at the end-of-life screening tool; NPV, negative predictive value; PC, palliative care; P-CaRES, palliative care and rapid emergency screening; PPV, positive predictive value; RISC, risk instrument for screening in the community; SIAARTI, Screening tool from the Italian Society of Anesthesia, Analgesia, Resuscitation, and Intensive care; SPEED, screening for palliative and end-of-life care needs in the emergency department; SST, simplified screening tool; SQ, surprise question; α = frailty as measured via a comprehensive geriatric assessment.

(99%, 95% CI 98%–100%) to predict 1-year patient mortality.[41] The study recommended that a threshold A-qCPR risk score of nine (high-risk group) as indicative of patients with unmet PC needs, as the PPV of the high-risk group was higher (65%, 95% CI 60%–69%) compared to patients with lower risk scores (see Table 2).[41]

Secondary outcomes

Reliability/validity

One study reported that the 13-item SPEED instrument had strong concurrent validity with domains of the Needs at the End-of-life Screening tool (NEST), with a Cronbach coefficient alpha of the SPEED instrument ranging from 0.716 to 0.991.[43] Following a corrected item correlation, the Cronbach coefficient of the SPEED instrument ranged from 0.326 to 0.970.[43] Another study reported that the SQ had moderate to strong correlation with the Caregiver Burden Score and Risk Instrument for Screening in the Community, but a poor correlation between the Clinical Frailty scale and the EQ-5D.[44] While Tineo et al.[40] did not assess the psychometric properties of their screening tool, their study reported that an unpublished pilot study found that their screening tool had acceptable internal consistency (Cronbach coefficient = 0.699).

Patients with unmet PC needs

The proportion of patients identified as having unmet PC needs varied among studies evaluating different screening tools and even among the studies assessing the same instruments (see Figure 3 and Table 3). Based on the criteria of the SQ, the median proportion of patients found to have unmet PC needs was 31% (IQR 15%–49%),[4,15,17,19,27-29,42,44,45] while the median proportion of patients identified as having unmet PC needs using the P-CaRES and SPEED instrument was 32% (IQR 30%–55%)[12,13,20,21,30-33] and 62% (IQR 52%–72%),[21,22,23,43] respectively. The range of patients identified with the other palliative screening tools ranged considerably from 12% to 80% (median 36%, IQR 25%–71%; see Figure 3 and Table 3).[24,26,34,36,37,40,41]

PC consultations

Eighteen studies reported on PC consultations among patients with unmet PC needs; however, outcome reporting varied considerably (see Table 3). Two studies reported on the proportion of patients with unmet PC needs who had received an ED-based PC consultation, which ranged from 1.4%[12] to 12%.[27] The proportion of patients with unmet PC needs who were hospitalized and received a PC consultation was reported to be 8%,[24] 13%,[15] 47%,[29] and 55%[19] across four studies. Patients with unmet PC needs who received a PC consultation postdischarge was reported to be 5%,[28] 6%,[31] 11%,[37,38] and 29%[12] in five studies. Five studies[23,25,26,36,46] applied an intervention

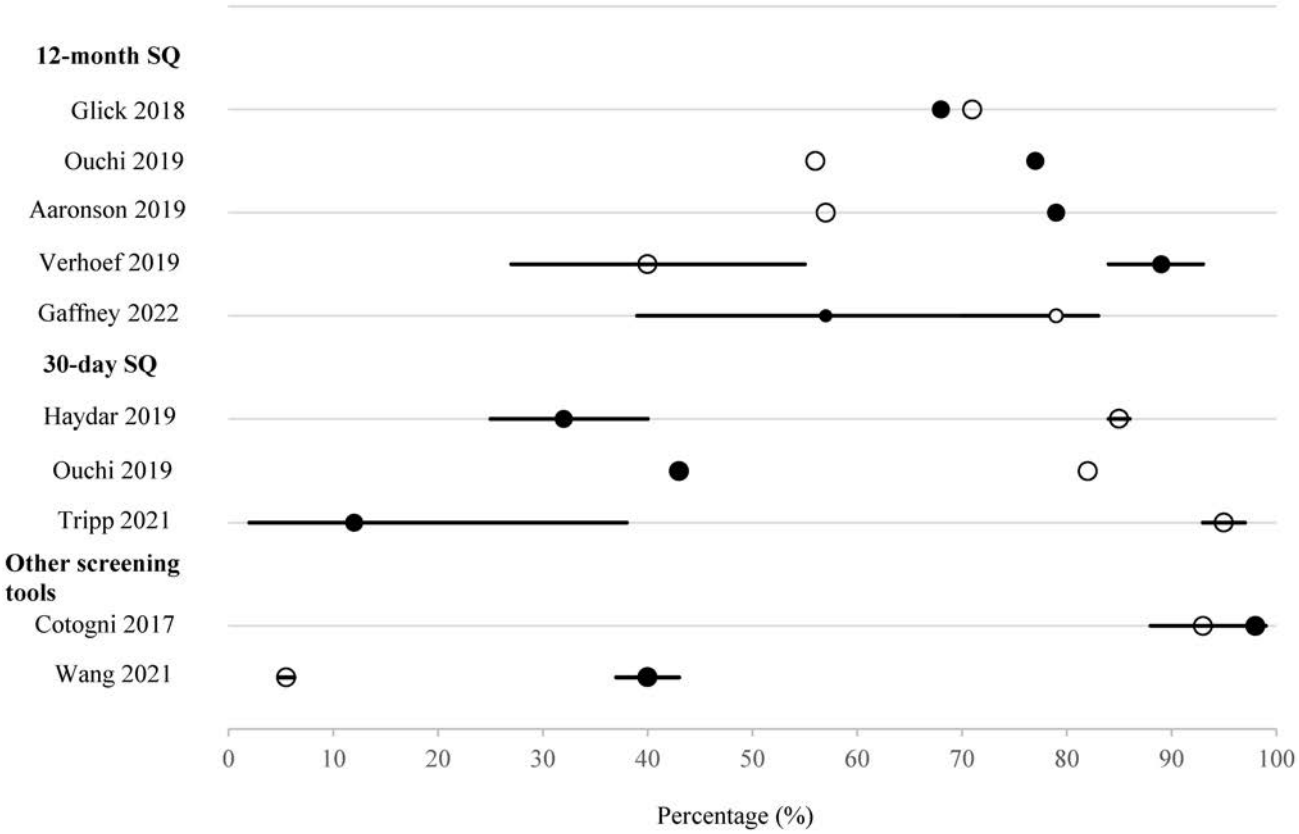


FIGURE 2 Performance of the screening tools to predict patient mortality among patients in the ED. The horizontal lines indicate the 95% CI for sensitivity (%) and specificity (%) and the filled circle symbols in the middle of the horizontal lines indicate the sensitivity and empty circle symbols represent the specificity

with the objective of promoting PC consultations and reported that a range between 32% and 47% of patients with unmet PC needs received a PC consultation in the ED or hospital.[23,25,26,46] One of the intervention studies reported that 90% of patients with unmet PC needs received a referral to PC post-discharge; however, only 43% of the patients attended the PC consultation.[36]

Disposition

The majority of patients identified as having unmet PC needs were admitted to hospital, with a median proportion of 64% (IQR 55%–77%; see Table 3).[12,20,22,23,27,28,30,31,38,39,43-46] The majority of patients identified as having unmet PC needs based on the P-CaRES (range: 69% to 79%)[12,20,30,31] and SPEED instrument (55%–59%)[22,23,43,46] were admitted to hospital. Alternatively, the median proportion of patients with unmet PC needs discharged from the ED was 23% (IQR 14%–28%).[12,16,19,20,23,25,29,38,39,45,46]

Mortality

The timeline of when mortality was reported ranged from within hospital to over 30 days, 6 months, 8 months, or a 12-month period (see Table 3). The median proportion of patients who died in hospital was 12% (IQR 5%–25%),[12,16,17,18,19,25,27,32,39] while the median proportion of patients with unmet PC needs who died following discharge from the ED (ranging from 1 to 12 months) was 51% (IQR 25%–69%; see Table 3).[4,13,15,28,33,34,36,41,45]

Subsequent health care utilization

Five studies assessed the proportion of patients with unmet PC needs who returned to the ED following the index visit; the median proportion was 32% (IQR 8%–53%; see Table 3).[12,23,28,34,46] Three studies reported 30%–77% of patients with unmet PC needs were subsequently hospitalized following discharge from their index visit.[28,29,44]

DISCUSSION

With an increase in ED visits among patients with EOL conditions, there is a need to identify those with unmet PC needs. While a previous systematic review published in 2016 included seven studies assessing different screening tools or interventions,[7] the current systematic review conducted a comprehensive and updated search of the literature and identified 35 studies utilizing 14 unique

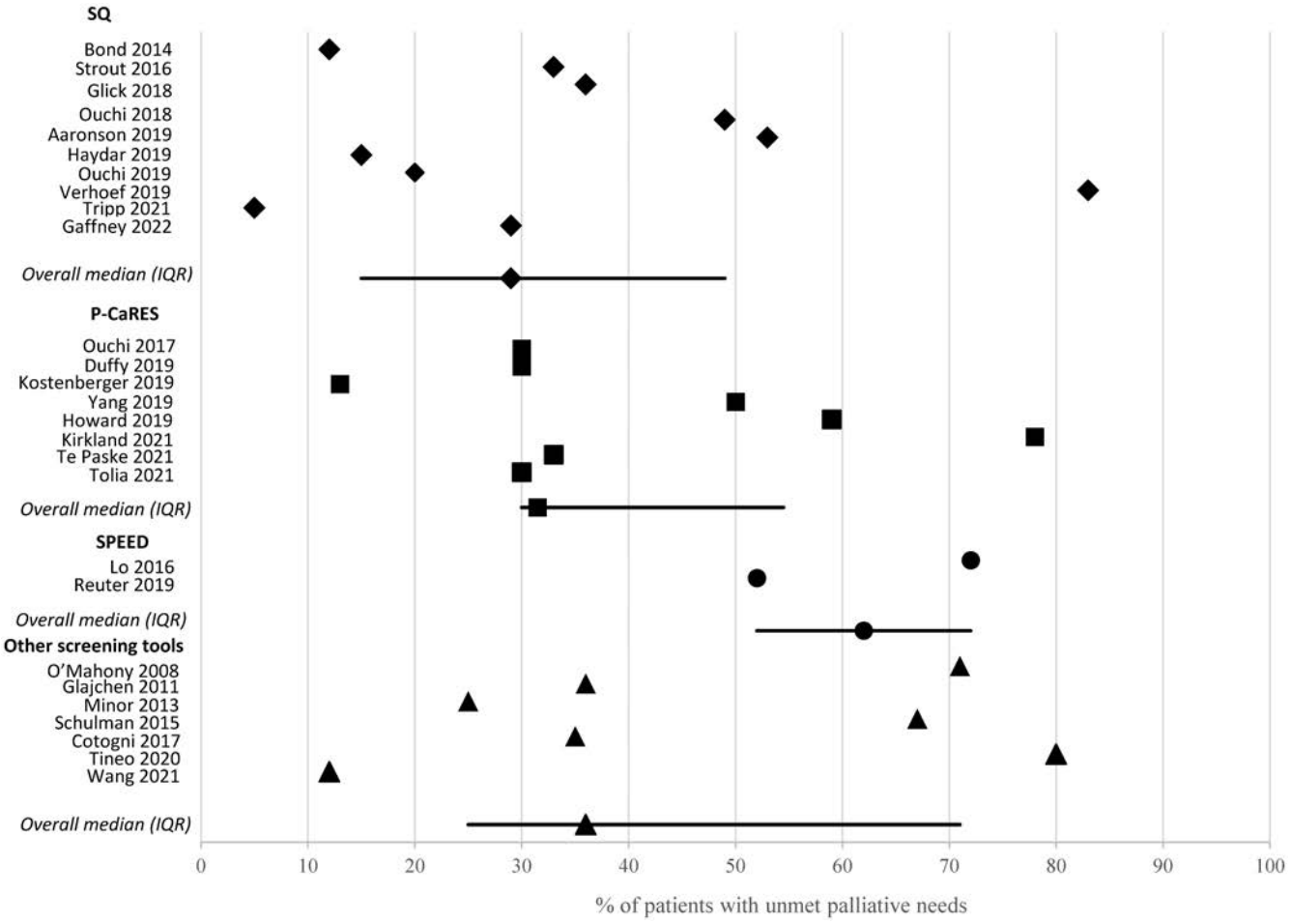


FIGURE 3 Proportion of patients in the ED identified as having unmet PC needs using various screening tools. The diamond (SQ), square (P-CaRES), circle (SPEED), and triangles (other screening tools) represent the average proportion of patients with unmet PC needs for the given studies. PC, palliative care

screening tools to identify unmet PC needs among adult patients presenting to the ED. Despite the increase in published studies, the performance of the majority of screening tools remain infrequently reported, limiting the ability to make any clear recommendations for their widespread implementation.

Of the screening tools assessed across the included studies, the psychometric properties of the SQ were the most commonly reported. The SQ exhibited moderate ability to predict mortality; however, performed better at predicting long-term mortality (within 12 months) than short-term mortality (within 30 days), while offering limited assistance to clinicians. For example, assuming a pretest probability of unmet PC needs in the populations at baseline of approximately 30%, given these results, the LRs are insufficient to confidently *rule in* a case when positive (30% increased to 45%) and insufficient to confidently *rule out* a case when negative (30% decreased to 18%) with certainty. Overall, strictly using the SQ to identify patients with unmet PC needs will result in referral of a large number of patients to PC services who may not benefit from them. Considering the costs and limited availability of PC services in the ED/hospital, screening tools with excellent performance are recommended to ensure the appropriate use of limited resources and improved quality of care.

The results of our review are similar to those of another systematic review that assessed the performance of the SQ across various clinical settings (e.g., ED, inpatient ward, ICU, general practice clinics, cancer centers).[47] That review reported that the SQ performed poorly to moderately as a predictive tool for death (sensitivity 67%, 95% CI 55.7%–76.7%; specificity 80.2%, 95% CI 73.3%–85.6%; +LR 3.4, 95% CI 2.8–4.1; −LR 0.41, 95% CI 0.32–0.54) and that the SQ performed worse among patients with non-cancer conditions.[47] Unfortunately, the current review was unable to provide further clarity due to a limited reporting across the included ED studies. For example, the majority of studies assessing the SQ enrolled patients with various medical conditions; however, these studies did not specify how the SQ performed among any subgroup of patients with advanced conditions. In addition, the majority of included studies utilized emergency physicians as the assessors of the SQ, and so it is unclear whether the performance of the SQ varies among user groups (e.g., emergency physicians, nurses, or other specialists) or whether prior training in PC impacts performance. Overall, given the limited availability of PC services in the ED/hospital settings, PC screening tools applied to patients with EOL conditions need to be both highly sensitive and

TABLE 3 Outcomes of patients identified as having unmet PC needs in ED

Study ID	Patients with unmet PC needs, *n* (%)	Patients with unmet PC needs discharged, *n* (%)	Patients with unmet PC needs hospitalized, *n* (%)	Mortality for patients with unmet PC needs, *n* (%)	Patients with unmet PC needs receiving PC consults, *n* (%)	Subsequent healthcare utilization, *n* (%)
SQ						
Bond 2014[15]	536/4478 (12)	NA	NA	NA	Inpatient PC consults: 69/536 (13)	NA
Yoshida 2014[16]	48 patients	20/48 (42)	NA	In-hospital mortality: 8/48 (17)	48 consults requested from the ED	NA
Strout 2016[17]	108/330 (33)	NA	NA	In-hospital mortality: 31/330 (9)	NA	NA
Strout 2017[18]	2023/10334 (20)	NA	NA	In-hospital mortality: 110/2023 (5.4)	NA	NA
Glick 2018[19]	112/313 (36)	17/112 (15)	NA	In-hospital mortality: 36/112 (32)	Inpatient PC consult: 62/112 (55)	NA
Ouchi 2018[4]	102/207 (49)	NA	NA	12-month mortality: 34/102 (33)	NA	NA
Ouchi 2019[42]	2104/10737 (20)	NA	NA	NA	NA	NA
Aaronson 2019[45]	103/193 (53)	6/103 (6)	20/103 (19)	12-month mortality: 44/103 (43)	NA	NA
Haydar 2019[27]	918/6122 (15)	NA	45/918 (5)	In-hospital mortality: 107/918 (12%)	ED-based PC consult: 114/918 (12)	NA
Verhoef 2020[28]	203/245 (83)	NA	157/203 (77)	12-month mortality: 141/203 (69)	Postdischarge PC consult: 11/203 (5)	Subsequent ED utilization: 93/203 (46) Subsequent hospitalization: 157/203 (77)
Tripp 2021[29]	19/381 (5)	11/19 (58)	NA	30-day mortality: 2/19 (1)	Inpatient PC consult: 9/19 (47)	Subsequent hospitalization: 8/19 (42)
Gaffney 2022[44]	56/191 (29)	NA	49/56 (88)	12-month mortality: 20/56 (36)	NA	Subsequent hospitalization: 13/43 (30)
P-CaRES						
Ouchi 2017[30]	62/207 (30)	NA	48/62 (77)	NA	NA	NA
Duffy 2019[20]	56/188 (30)	15/56 (27)	41/56 (73)	NA	Site of PC consult not specified: 9/56 (16)	NA
Kostenberger 2019[31]	145/1096 (13)	NA	108/145 (74)	NA	Outpatient PC consult: 8/145 (6)	NA
Yang 2020[32]	396/796 (50)	NA	NA	In-hospital mortality: 161/396 (41)	NA	NA
Howard 2021[13]	Chaplain: 41/58 (71) Physician: 34/58 (59)	NA	NA	12-month mortality: 10/58 (17)	Site of PC consult not specified: 18/58 (31)	NA

(Continues)

TABLE 3 (Continued)

Study ID	Patients with unmet PC needs, *n* (%)	Patients with unmet PC needs discharged, *n* (%)	Patients with unmet PC needs hospitalized, *n* (%)	Mortality for patients with unmet PC needs, *n* (%)	Patients with unmet PC needs receiving PC consults, *n* (%)	Subsequent healthcare utilization, *n* (%)
Kirkland 2021[12]	518/663 (78)	142/508 (28)	352/508 (69)	In-hospital mortality: 89/489 (18)	ED-based PC consult: 6/518 (1.4) Postdischarge PC consult: 36/126 (29)	Subsequent ED use: 129/430 (30)
Te Paske 2021[33]	43/131 (33)	NA	NA	6-month mortality: 22/43 (51)	NA	NA
Tolia 2020[21]	194/641 (30)	NA	NA	NA	NA	NA
SPEED						
Richards 2011[43]	53 patients	NA	29/53 (55)	NA	NA	NA
Lo 2016[22]	58/81 (72)	NA	29/53 (55)	NA	NA	NA
Marshall 2017[23]	293 patients	66/293 (23)	172/293 (59)	NA	ED-based consults following implementation of a PC consult intervention: 103/293 (35)	Subsequent ED utilization: 8/103 (8)
Reuter 2019[46]	422/817 (52)	92/422 (22)	247/422 (59)	NA	ED-based consults following implementation of a PC consult intervention: 167/422 (40)	Subsequent ED utilization: 13/167 (8)
Other screening tools						
Mahony 2008[34]	670/950 (71)	NA	NA	50-day mortality: 90/291 (30.9) 180-day mortality: 197/291 (67.8)	NA	Subsequent ED utilization: 172/291 (59)
Glajchen 2011[36]	51/140 (36)	NA	NA	8-month mortality: 41/51 (80)	Postdischarge PC referral following an PC consult intervention: 46/51 (90) Attended the PC consult: 20/46 (43)	NA
Grudzen 2011[35]	Financial needs: 36/50 (72) Access to care needs: 29/50 (58) Caregiver needs: 36/50 (72) Physical health needs: 47/50 (94) Mental health needs: 31/50 (62)	NA	NA	NA	NA	NA
Minor 2013[24]	387/1544 (25)	NA	NA	NA	Inpatient PC consult: 30/387 (8)	NA
Earle 2014[25]	155 patients	16/105 (10)	NA	In-hospital mortality: 7/105 (5)	ED-based consult following initiation of a PC consult intervention 50/155 (32)	NA

TABLE 3 (Continued)

Study ID	Patients with unmet PC needs, *n* (%)	Patients with unmet PC needs discharged, *n* (%)	Patients with unmet PC needs hospitalized, *n* (%)	Mortality for patients with unmet PC needs, *n* (%)	Patients with unmet PC needs receiving PC consults, *n* (%)	Subsequent healthcare utilization, *n* (%)
Schulman 2015[26]	35/52 (67)	NA	NA	NA	ED and inpatient-based PC consults following initiation of a PC consult intervention: 18/38 (47)	NA
Cotogni 2017[37]	91/257 (35)	NA	NA	NA	NA	NA
Koh 2019[38]	340 patients	49/340 (14)	279/340 (82)	NA	Postdischarge PC consult: 39/340 (11)	NA
Bakan 2020[39]	NA	88/208 (42)	115/208 (55)	In-hospital mortality: 5/208 (2)	NA	NA
Tineo 2020[40]	50/60 (80)	NA	NA	NA	PC requested by ED physician: 12/60 (20)	NA
Wang 2021[41]	569/4627 (12)	NA	NA	12-month mortality: 364/569 (64)	NA	NA

Abbreviations: NA, not applicable; PC, palliative care; P-CaRES, Palliative Care and Rapid Emergency Screening; SPEED, screening for palliative and end-of-life care needs in the emergency department; SQ, surprise question.

specific (90% or higher). Further assessment into factors (e.g., specific EOL conditions, assessors, setting/context where this screening tool is applied) that could influence the performance of the SQ to predict mortality appear warranted.

The other most commonly assessed screening tools were the SPEED instrument[22,23,43,46] and the P-CaRES tool.[12,13,20,21,30-33] While one study found that the SPEED instrument had high reliability compared to similar survey questions from previous validated tools,[43] the accuracy of SPEED is unclear. The original 11-item SPEED instrument was subsequently shortened to include 5 items to allow the tool to be more readily integrated into the patient's bedside assessment[46]; however, the psychometric properties of the modified 5-item SPEED instrument have not been reported. The P-CaRES tool was found to have excellent content validity[5] and to have excellent accuracy in correctly recommending PC consults when used by emergency physicians in various written case vignettes.[6] In the current review, the accuracy of the P-CaRES tool when applied to adult ED patients was infrequently reported, with only one study comparing the ability of trained chaplains to use the P-CaRES tool to identify ED patients with unmet PC needs compared to emergency physicians.[13] While the evidence suggests that the SPEED[43] and P-CaRES[5,6] tools have excellent content reliability/validity, additional studies assessing their accuracy in identifying unmet PC needs in ED patients are needed. For the other screening tools identified in this review, there is insufficient evidence to recommend any one over the SQ, P-CaRES, or SPEED tools for identifying unmet PC needs in the ED. One study found that the SST tool had high sensitivity and specificity to identify patients with unmet PC needs[37]; however, additional studies are needed. Importantly, given the poor to moderate performance of the SQ, additional studies assessing and reporting the performance characteristics of other available screening tools such as P-CaRES and SPEED are necessary.

The proportion of patients identified as having unmet PC needs varied considerably across the included studies, even among studies using the same screening tools, which could be the result of many factors. First, the patient population that was screened for unmet PC needs varied, with some studies screening only patients 65 years or older,[4,44] while other studies specifically screened patients who were admitted to hospital or the ICU[18,19,27] or any patients with specific clinical conditions[29,45] or specifically patients with established EOL conditions.[28,30,35] As previously mentioned, the performance of the SQ to predict mortality appears to decrease among patients with non–cancer-related conditions,[47] and at this point, it is unclear whether other patient demographic or clinical characteristics impacted the performance of the other PC screening tools. Second, it was unclear how the screening tools were implemented within the ED, including when and how the screening tool was completed, whether health care staff were trained in the use of the screening tool, and time needed to complete the screening process. It is possible that differences in how the screening tools were implemented could have impacted the performance of the tools across the studies. Finally, the variation in the proportion of

patients identified with unmet PC needs by the screening tools may be related to the inconsistent criteria for unmet PC needs across the tools. For example, some screening tools such as the P-CaRES and SPEED tools incorporated combinations of factors including frequent ED/hospitalizations, lack of family support or GOC planning, and uncontrolled symptoms in identifying whether or not patients were considered to have unmet PC needs. Other tools, such as the SQ, strictly used the expected likelihood of death in a prespecified period of time as their sole indicator of unmet PC needs, without considering any other physical, psychological, or socio-related factors. Given the complexity of the concept of unmet PC needs among patients with EOL conditions, it may be difficult to establish simple and quick screening tools such as the SQ to effectively and accurately identify patients in need of PC services.

While the proportion of patients with unmet PC needs identified across the studies varied, the summary estimates do provide some useful information; for example, the median proportion of patients with unmet PC needs ranged between 32% and 35% and could be used in future study planning. In addition, the proportion of patients with unmet PC needs who received a PC consult either in the ED or in hospital varied among the studies yet tended to be low. It is possible that some studies may have been conducted in EDs/hospitals that provide additional PC services compared to other studies, which could explain some of the variation in PC consultation referrals among the studies. It is unclear, however, whether screening for unmet PC needs prompted emergency physicians to more likely request PC consultations across some of the included studies, despite no explicit intervention to promote PC interventions. Various factors have been found to impact PC consultation rates including patient race and age,[48] and research has also suggested that inpatients with noncancer EOL conditions were more likely to undergo delays in receiving referrals to in-hospital PC consultations.[49] Finally, the available evidence suggests that approximately two of every three patients with unmet PC needs are hospitalized and that patients with unmet PC needs are likely to make subsequent ED visits. Further research assessing the demographics and clinical characteristics/ED management of patients identified as having unmet PC needs could help physicians and providers understand their health care needs.[12,50]

LIMITATIONS

There are some limitations of the methods employed in this review to consider. While the risk of publication bias exists, multiple steps were taken to attenuate these risks, including an extensive search of the literature and no limitations based on language, publication status, or year of publication. Moreover, steps were taken to reduce the risk of selection bias including the utilization of two independent reviewers to identify eligible studies to minimize the risk of missing any available studies. In addition, there were some inconsistencies identified by the STARD and QUADAS quality assessment tools with respect to the reporting of performance including blinding, test reproducibility, and assessors training. The methodologic quality of the majority of studies that did not assess psychometric properties was low primarily due to a lack of a comparative group and inconsistent outcome and follow-up reporting. Moreover, there was a lack of a clear "reference standard" for the identification of patients with PC needs across the studies. The majority of studies assessing the accuracy of the SQ to predict mortality utilized death certificates or medical records as their reference standard to confirm patient death. Otherwise, a handful of studies utilized screening tools used other clinical settings to identify patients with unmet PC needs[37,43,44]; however, it is unclear how effective these "reference standards" are in identifying ED patients with unmet PC needs. Finally, few comparative studies were performed, making selection of a preferred tool(s) for practice difficult.

This review was unable to assess the ease in which any of these tools could be implemented into ED practice due to a limited reporting of fidelity/implementation strategies across the included studies. While information related to training and time needed to administer the screening tools was not reported across the studies, it is likely that the low complexity of the SQ, which required a single yes or no response from attending physicians, makes it an appealing screening tool to utilize in a busy environment such as the ED and is perhaps one reason why it was the most commonly investigated. Given the fast-paced environment of the ED, screening tools that have high accuracy, but are complex or cumbersome to apply, will likely not gain widespread acceptance.

CONCLUSIONS

This review identified 14 unique screening tools used to identify ED patients with unmet PC needs; however, the clinical utility of the majority of these tools to identify patients who would benefit from timely PC services remains unclear due to limited reporting on their performance. The SQ was found to have moderate predictive ability for long-term mortality; nonetheless, given the state of the evidence, there is insufficient evidence to recommend one particular screening tool for widespread use in the ED setting to accurately identify patients with unmet PC needs. While additional research assessing the performance and utility of these tools is recommended, before doing so health care providers and administrators should assess the context in which screening tools have been tested and identify those they believe might be of clinical value in their unique settings.

ACKNOWLEDGMENTS

The authors thank Dr. Kei Ouchi and Dr. Tania Strout for responding to our requests for additional information on their studies. We also thank Jillian Meyer for her assistance with the quality assessment.

CONFLICT OF INTEREST

All authors except Sandy Campbell were co-authors on one study that was included in this systematic review; however, no authors have other conflicts of interest to declare.

AUTHOR CONTRIBUTIONS

Scott Kirkland contributed to the conception and design of the review, writing the protocol, study screening, quality assessment, acquisition of data, data analysis and interpretation, and manuscript preparation. Esther Yang contributed to study screening, quality assessment, acquisition of data, data analysis and interpretation, and manuscript preparation. Miriam Garrido Clua and Maureen Kruhlak contributed to study screening, quality assessment, acquisition of data, and manuscript preparation. Sandra Campbell contributed to the development of the electronic search and manuscript preparation. She has no conflicts of interest to declare. Cristina Villa Roel contributed to the design of the review, writing the protocol, study screening, and manuscript preparation. Brian Rowe is the guarantor of the review. He secured funding and contributed to the conception and design of the review, writing the protocol, data analysis of interpretation, and manuscript preparation.

ORCID

Scott W. Kirkland https://orcid.org/0000-0002-6043-0999
Brian H. Rowe https://orcid.org/0000-0003-3446-2097

REFERENCES

1. Chan GK. End-of-life models and emergency department care. *Acad Emerg Med*. 2004;11:79-86.
2. Bailey CJ, Murphy R, Porock D. Dying cases in emergency places: caring for the dying in emergency departments. *Soc Sci Med*. 2011;73:1371-1377.
3. Five things physicians and patients should question. 2013. Accessed June 17, 2019. https://www.choosingwisely.org/societies/american-academy-of-hospice-and-palliative-medicine/
4. Ouchi K, Jambaulikar G, George NR, et al. The "surprise question" asked of emergency physicians may predict 12-month mortality among older emergency department patients. *J Palliat Med*. 2018;21:236-240.
5. George N, Barrett N, McPeake L, Goett R, Anderson K, Baird J. Content validation of a novel screening tool to identify emergency department patients with significant palliative care needs. *Acad Emerg Med*. 2015;22:823-837.
6. Bowman J, George N, Barrett N, Anderson K, Dove-Maguire K, Baird J. Acceptability and reliability of a novel palliative care screening tool among emergency department providers. *Acad Emerg Med*. 2016;23:694-702.
7. George N, Phillips E, Zaurova M, Song C, Lamba S, Grudzen C. Palliative care screening and assessment in the emergency department: a systematic review. *J Pain Symptom Manag*. 2016;51:108-119.
8. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151:264-269.
9. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. 2003;3:25.
10. Bossuyt PM, Reitsma JB, Bruns DE, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *BMJ*. 2003;326:41-44.
11. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. Accessed June 2019. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp
12. Kirkland SW, Garrido-Clua M, Kruhlak M, et al. Comparison of characteristics and management of emergency department presentations between patients with met and unmet palliative care needs. *PLoS One*. 2021;16:e0257501.
13. Howard EH, Schwartz R, Feldstein B, et al. Harnessing the chaplain's capacity to identify unmet palliative needs of vulnerable older adults in the emergency department. *J Palliat Care*. 2021;8258597211003359. https://doi.org/10.1177/08258597211003359
14. McGee S. Simplifying likelihood ratios. *J Gen Intern Med*. 2002;17:646-649.
15. Bond AG, Strout TD, Haydar SA, Han PKJ. Emergency physician responses to the surprise question and identification of unmet palliative care needs. *Acad Emerg Med*. 2014;21:S193.
16. Yoshida C, Ing M, Uemura D, Fischberg D. Collaboration between a palliative care department and the emergency department improves early access to palliative care. *J Palliat Med*. 2014;17:A-21.
17. Strout TDS, Haydar SA, Eager E, Han PKJ. Identifying unmet palliative care needs in the ED: use of the 'surprise question' in patients with sepsis. *Acad Emerg Med*. 2016;23:S196.
18. Strout TD, Haydar SA, Vogt AO. Identifying unmet palliative care needs: use and utility of the "surprise question" in emergency and inpatient settings. *Acad Emerg Med*. 2017;24:S149.
19. Glick J, Gallo MB, Chelluri J, Courtright K. Utility of the "surprise question" in critically ill emergency department patients. *Ann Emerg Med*. 2018;72:S68.
20. Duffy J, Crump M, O'Connor E. Identifying unmet palliative care needs in the emergency department. *CJEM*. 2019;21:S75.
21. Tolia VM, Yourman L, Brennan JJ, Cronin AO, Castillo EM. Initial evaluation of a palliative care screening tool in the emergency department. *Ann Emerg Med*. 2020;76:S70.
22. Lo AX, Rundle AG, Pearce MR, Kaushik D, Quest TE. Emergency department-based assessment of palliative needs in adults with heart failure. *Acad Emerg Med*. 2016;23:S131.
23. Marshall A, Reuter Q, Powell ES, et al. Emergency department-based palliative interventions. *Ann Emerg Med*. 2017;70:S113-S114.
24. Minor SK. Impact of emergency department identification of palliative care patients on hospital average length of stay. *J Palliat Med*. 2013;16:A15.
25. Earle BI. Analysis of an emergency room palliative care trigger tool: diagnosis and barriers to its use. *J Palliat Med*. 2014;17:A14-A15.
26. Schulman KA, Zalenski R, Johnson C. Palliative care screening in the emergency department. *Acad Emerg Med*. 2015;22:S283-S284.
27. Haydar SA, Strout TD, Bond AG, Han PK. Prognostic value of a modified surprise question designed for use in the emergency department setting. *Clin Exp Emerg Med*. 2019;6:70-76.
28. Verhoef MJ, de Nijs EJM, Fiocco M, Heringhaus C, Horeweg N, van der Linden YM. Surprise question and performance status indicate urgency of palliative care needs in patients with advanced cancer at the emergency department: an observational cohort study. *J Palliat Med*. 2020;23:801-808.
29. Tripp D, Janis J, Jarrett B, et al. How well does the surprise question predict 1-year mortality for patients admitted with COPD ? *J Gen Intern Med*. 2021;36:2656-2662.
30. Ouchi K, Block SD, Schonberg MA, et al. Feasibility testing of an emergency department screening tool to identify older adults appropriate for palliative care consultation. *J Palliat Med*. 2017;20:69-73.
31. Kostenberger M, Neuwersch S, Weixler D, Pipam W, Zink M, Likar R. Prevalence of palliative care patients in emergency departments. *Wien Klin Wochenschr*. 2019;131:404-409.
32. Yang C, Yang TT, Tsou YJ, et al. Initiating palliative care consultation for acute critically ill patients in the emergency department intensive care unit. *J Chin Med Assoc*. 2020;83:500-506.
33. Paske JRT, DeWitt S, Hicks R, Semmens S, Vaughan L. Palliative care and rapid emergency screening tool and the palliative performance scale to predict survival of older adults admitted to the

hospital from the emergency department. *Am J Hosp Palliat Care*. 2021;38:800-806.
34. Mahony SO, Blank A, Simpson J, et al. Preliminary report of a palliative care and case management project in an emergency department for chronically ill elderly patients. *J Urban Health*. 2008;85:443-451.
35. Grudzen CR, Stone SC, Morrison RS. The palliative care model for emergency department patients with advanced illness. *J Palliat Med*. 2011;14:945-950.
36. Glajchen M, Lawson R, Homel P, Desandre P, Todd KH. A rapid two-stage screening protocol for palliative care in the emergency department: a quality improvement initiative. *J Pain Symptom Manag*. 2011;42:657-662.
37. Cotogni P, De Luca A, Evangelista A, et al. A simplified screening tool to identify seriously ill patients in the emergency department for referral to a palliative care team. *Minerva Anestesiol*. 2017;83:474-484.
38. Koh MYH, Lee JF, Montalban S, Foo CL, Hum AYM. ED-PALS: a comprehensive palliative care service for oncology patients in the emergency department. *Am J Hosp Palliat Care*. 2019;36:571-576.
39. Bakan G, Ozen M, Azak A, Erdur B. Determination of the characteristics and outcomes of the palliative care patients admitted to the emergency department. *Int Emerg Nurs*. 2020;53:100934.
40. Tineo JPA, Alva RV, Pastrana RH, Manrique PZH, Espinoza TO. Need for palliative care in patients admitted to emergency departments of three tertiary hospitals: evidence from a latin-american city. *Palliat Med Pract*. 2020;14:236-240.
41. Wang RF, Lai CC, Fu PY, et al. A-qCPR risk score screening model for predicting 1-year mortality associated with hospice and palliative care in the emergency department. *Palliat Med*. 2021;35:408-416.
42. Ouchi K, Strout T, Haydar S, et al. Association of emergency clinicians' assessment of mortality risk With actual 1-month mortality among older adults admitted to the hospital. *JAMA Netw Open*. 2019;2:e1911139.
43. Richards CT, Gisondi MA, Chang CH, et al. Palliative care symptom assessment for patients with cancer in the emergency department: validation of the screen for palliative and end-of-life care needs in the emergency department instrument. *J Palliat Med*. 2011;14:757-764.
44. Gaffney L, Jonsson A, Judge C, Costello M, O'Donnell J, O'Caoimh R. Using the "surprise question" to predict frailty and healthcare outcomes among older adults attending the emergency department. *Int J Environ Res Public Health*. 2022;19:1709.
45. Aaronson EL, George N, Ouchi K, et al. The surprise question can be used to identify heart failure patients in the emergency department who would benefit from palliative care. *J Pain Symptom Manag*. 2019;57:944-951.
46. Reuter Q, Marshall A, Zaidi H, et al. Emergency department-based palliative interventions: a novel approach to palliative care in the emergency department. *J Palliat Med*. 2019;22:649-655.
47. Downar J, Goldman R, Pinto R, Englesakis M, Adhikari NK. The "surprise question" for predicting death in seriously ill patients: a systematic review and meta-analysis. *CMAJ*. 2017;189:E484-E493.
48. Price M, Howell EP, Dalton T, et al. Inpatient palliative care utilization for patients with brain metastases. *Neurooncol Pract*. 2021;8:441-450.
49. Maddison AR, Malik S, Smaggus A. Inpatient palliative care consultations from a Canadian clinical teaching unit: who is referred and when? *J Palliat Care*. 2018;33:204-208.
50. Verhoef MJ, de Nijs E, Horeweg N, et al. Palliative care needs of advanced cancer patients in the emergency department at the end of life: an observational cohort study. *Support Care Cancer*. 2020;28:1097-1107.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Kirkland SW, Yang EH, Garrido Clua M, et al. Screening tools to identify patients with unmet palliative care needs in the emergency department: A systematic review. *Acad Emerg Med*. 2022;29:1229-1246. doi: 10.1111/acem.14492

Received: 18 November 2021 | Revised: 4 April 2022 | Accepted: 25 April 2022

DOI: 10.1111/acem.14513

RESEARCH METHODS AND STATISTICS

Pragmatic clinical trial design in emergency medicine: Study considerations and design types

Cameron J. Gettel MD, MHS[1,2] | Maame Yaa A. B. Yiadom MD, MPH, MSCI[3] | Steven L. Bernstein MD[4] | Corita R. Grudzen MD, MSHS[5] | Bidisha Nath MBBS, MPH[1] | Fan Li PhD[6] | Ula Hwang MD, MPH[1,7] | Erik P. Hess MD, MSc[8] | Edward R. Melnick MD, MHS[1,6]

[1]Department of Emergency Medicine, Yale School of Medicine, New Haven, Connecticut, USA

[2]Center for Outcomes Research and Evaluation, Yale School of Medicine, New Haven, Connecticut, USA

[3]Department of Emergency Medicine, Stanford University School of Medicine, Stanford, California, USA

[4]Geisel School of Medicine at Dartmouth, Hanover, New Hampshire, USA

[5]Ronald O. Perelman Department of Emergency Medicine and Department of Population Health, New York University Grossman School of Medicine, New York, New York, USA

[6]Department of Biostatistics, Yale School of Public Health, New Haven, Connecticut, USA

[7]Geriatrics Research, Education and Clinical Center, James J. Peters VA Medical Center, Bronx, New York, USA

[8]Department of Emergency Medicine, Vanderbilt University School of Medicine, Nashville, Tennessee, USA

Correspondence
Cameron J. Gettel, MD, MHS, Department of Emergency Medicine, Yale School of Medicine, New Haven, CT 06520, USA
Email: cameron.gettel@yale.edu

Funding information
This work is supported within the National Institutes of Health (NIH) Health Care Systems Research Collaboratory by the NIH Common Fund through cooperative agreement U24AT009676 from the Office of Strategic Coordination within the Office of the NIH Director and cooperative agreement (UH3DA047003) from the National Institute on Drug Abuse of the NIH. Dr. Gettel is also supported by the Claude D. Pepper Older Americans Independence Center at Yale School of Medicine (P30AG021342), the National Institute on Aging (NIA; R03AG073988), and the NIA Imbedded Pragmatic Alzheimer's and AD-Related Dementias Clinical Trials (IMPACT) Collaboratory (U54AG063546). Dr. Yiadom is supported by the Institute for Human-centered Artificial Intelligence and the Emergency

Abstract

Pragmatic clinical trials (PCTs) focus on correlation between treatment and outcomes in real-world clinical practice, yet a guide highlighting key study considerations and design types for emergency medicine investigators pursuing this important study type is not available. Investigators conducting emergency department (ED)-based PCTs face multiple decisions within the planning phase to ensure robust and meaningful study findings. The PRagmatic Explanatory Continuum Indicator Summary 2 (PRECIS-2) tool allows trialists to consider both pragmatic and explanatory components across nine domains, shaping the trial design to the purpose intended by the investigators. Aside from the PRECIS-2 tool domains, ED-based investigators conducting PCTs should also consider randomization techniques, human subjects concerns, and integration of trial components within the electronic health record. The authors additionally highlight the advantages, disadvantages, and rationale for the use of four common randomized study design types to be considered in PCTs: parallel, crossover, factorial, and stepped-wedge. With increasing emphasis on the conduct of PCTs, emergency medicine investigators will benefit from a rigorous approach to clinical trial design.

Presented at the Society for Academic Emergency Medicine Annual Meeting, New Orleans, LA, May 2022.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2022 The Authors. *Academic Emergency Medicine* published by Wiley Periodicals LLC on behalf of Society for Academic Emergency Medicine.

Department Benchmarking Alliance. Dr. Grudzen is supported by the NIH (UG3AT009844 and UH3AT009844) and the Patient-Centered Outcomes Research Institute (PLC-1609-36,306). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or other sponsors.

Supervising Editor: Dr. Craig Newgard.

KEYWORDS
clinical trial, emergency, pragmatic, study design

INTRODUCTION

Randomized clinical trials are considered the standard design to assess an intervention's effect due to their ability to minimize confounding and potential selection biases that could influence the outcome of interest.[1] However, prior analyses of clinical trials have shown that the majority are too small in sample size, thereby lacking statistical power, or are prone to other critical issues in design and execution.[2] With explanatory clinical research often conducted in highly controlled patient populations with strict inclusion criteria, results may not be replicated under real-world conditions.[3] To develop valid and reliable treatment effect evidence, the number of clinical trials in the emergency department (ED) setting is increasing, particularly with "pragmatic" clinical trials (PCTs) focusing on the intervention effect on outcomes in routine clinical practice. The term "pragmatic" was first used by Schwartz and Lellouch[4] to describe approaches to trial design comparing two treatments under the conditions in which they would be applied in practice. In general, PCTs are distinguished from explanatory trials by three key attributes: 1) an intent to inform decision makers (e.g., patients, policymakers, clinicians) rather than elucidating a biological or social mechanism, 2) an intent to enroll a representative population relevant to the decision in practice, and 3) either an intent to streamline procedures and data collection or measuring a broad range of outcomes.[5] Few clinical trials can be fully pragmatic, but efforts to prioritize features of pragmatic trials when feasible should be considered by emergency medicine investigators given the increasing emphasis placed on dissemination and implementation of interventions with proven benefit in real-world settings.[6–9]

As the pendulum continues to swing more toward value-based care, Medicare and other payers, as well as the National Institutes of Health (NIH), have shown an increasing interest in the conduct of PCTs. For example, Medicare and United Healthcare insurance coverage decisions are being influenced by PCTs of older adults receiving acupuncture for low back pain and neutralizing antibodies for COVID-19.[10,11] The NIH Health Care Systems Collaboratory, supported by the NIH Common Fund and multiple institutions, is a particular example of a broader initiative aiming to improve the ways in which PCTs are conducted. To support the design and rapid execution of PCTs across health care systems, the NIH Collaboratory has also developed the electronic Living Textbook including expert consensus, standard approaches, best practices, peer-reviewed articles, and guidance documents related to the conduct of PCTs.[12] Despite the increased interest in PCTs, investigators are being increasingly asked to "do more with less." Adjusted for inflation, the average R01-equivalent grant award amount has remained essentially stagnant over the past 25 years.[13] With limited funds available, investigators may encounter challenges meeting desired enrollment numbers, which runs counter to the often-needed greater enrollment within PCTs given their variability and flexibility of intervention development and assessment. To discover nonneutral results of novel interventions and treatments, ED investigators must balance the possibility of it simply being too hard to find a "signal" of treatment difference that overwhelms the "noise" of potential reduced sample size and increased variation inherently present within PCTs.[14]

To overcome these challenges and conduct an efficient trial, the planning phase of PCTs is central to its success and requires somewhat different approaches to study design, data collection, and ethical oversight compared to traditional clinical trials.[15] However, a guide for planning and conducting ED-based PCTs is lacking despite the increasing call for their completion. This article reviews concepts related to ED-based PCTs, including: 1) the PRagmatic Explanatory Continuum Indicator Summary 2 (PRECIS-2) framework, 2) additional key study considerations, and 3) illustrative ED-based examples. The goal of this article is to provide a resource for emergency medicine trialists to reference in the study design phase when planning ED-based PCTs given intervention characteristics and available resources.

THE EXPLANTORY–PRAGMATIC SPECTRUM

When establishing the research question to evaluate an intervention, investigators must consider the explanatory–pragmatic spectrum. Trials are rarely completely explanatory or completely pragmatic in nature. Frequently, components of both are present, suggesting that a trial may fall somewhere on a continuum based on decisions made by investigators. Explanatory trials are conducted under ideal settings to standardize intervention delivery often with research staff on hand to enroll and administer the intervention itself. Explanatory trials tend to have fairly restrictive inclusion and exclusion criteria, often resulting in a small proportion of individuals, with comparable clinical and demographic characteristics, at risk for the condition of interest to be enrolled. Explanatory trials answer the question, "Does intervention X work under ideal conditions?" and typically is aligned with clinical efficacy research. Conversely,

pragmatic trials study an intervention's effect in real-world clinical practice by capturing a more heterogeneous population and tolerating some flexibility in the delivery of the intervention. PCTs answer the question "Does intervention X work in routine clinical practice?" and is more frequently aligned with clinical effectiveness research. Resultantly, explanatory trials generally exhibit a high degree of internal validity while pragmatic trials prioritize generalizability and external validity.[16]

With widespread concern that explanatory trials may be poor predictors of real-world effectiveness, the original PRECIS tool was developed to help investigators work through study design decisions to avoid designing a trial that did not meet their own intentions.[17] In 2015, the PRECIS-2 wheel diagram was developed and offers a framework guide for trialists to consider the pragmatic or explanatory nature of their trial across nine domains, including eligibility, recruitment, setting, organization, flexibility—delivery, flexibility—adherence, follow-up, primary outcome, and primary analysis.[17,18] Domains are rated from 1 (very explanatory) to 5 (very pragmatic). ED investigators should strongly consider the use of the PRECIS-2 wheel diagram during the study design phase to increase chances that results of the trial are generalizable across the intended practice settings (Figure 1).

PRECIS-2 CONSIDERATIONS IN ED-BASED PCT DESIGN

Unique challenges exist when conducting PCTs in the ED setting that span the PRECIS-2 domains. For trials involving time-sensitive conditions in the ED, one characteristic of note is the brief therapeutic window for decisions surrounding critical procedures.[14] The PreVent trial is an example ED-based study of bag–mask ventilation during tracheal intubation that addressed pragmatic approaches to time-sensitive conditions by strategically placing randomization envelopes, ensuring broad eligibility criteria, paring down the data collection form to the minimum information required, and having a nonproceduralist that was familiar with the procedure complete the form.[14,19] Second, the broader ED population may also include patients who are not regular visitors to that health system, thereby limiting electronic health record (EHR) data available at the point of care, in comparison to PCTs performed in the primary care or outpatient cardiology settings, as examples. Third, and to be discussed further, is the added complexity surrounding capacity and consent for certain vulnerable ED populations with mental illness or substance use disorder.

Researchers face multiple decisions when conducting clinical trials, with the planning phase of the trial serving as a critical step to optimize robust and meaningful study findings. Lack of foresight to address potential pitfalls a priori or rushing into study enrollment can leave the investigator searching for ways to "save" the trial post hoc. Detailed below, the nine domains within the PRECIS-2 framework are expanded upon to identify decision points for investigators to consider in the design of ED-based clinical trials.[17,20–24] Each domain is not meant to be a "yes" or "no" in regard to a pragmatic approach, but instead the PRECIS-2 framework is rather a series of decisions to fit a trial design to the purpose intended by the investigators. We additionally provide illustrative examples of studies across disciplines that have mapped the PRECIS-2 framework domains and

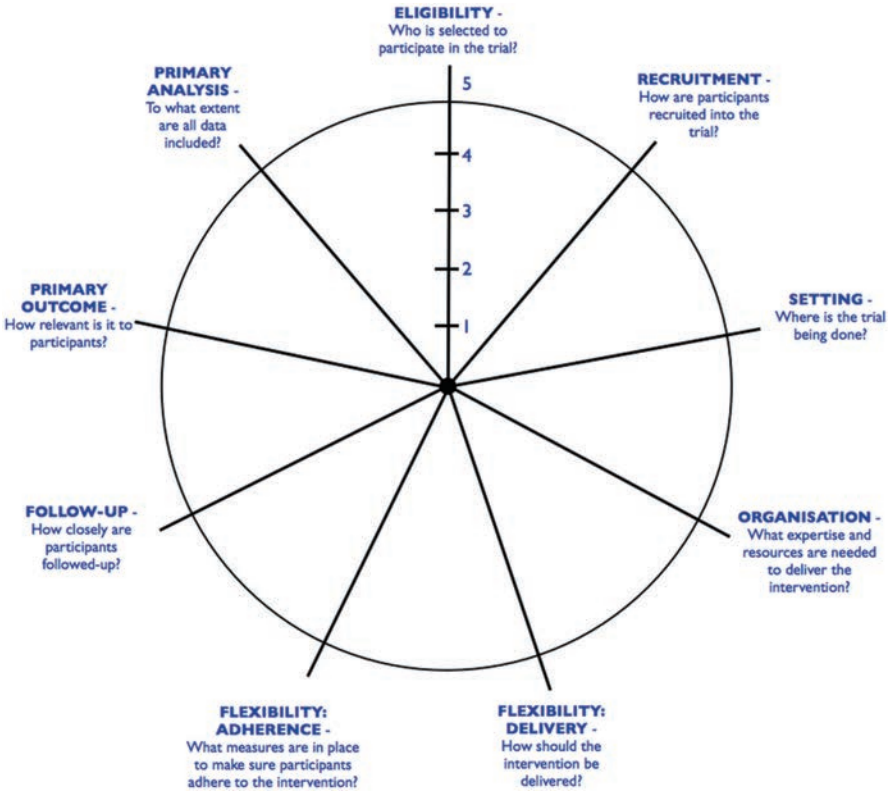


FIGURE 1 The PRagmatic-Explanatory Continuum Indicator Summary 2 (PRECIS-2) wheel.
Note: Used with permission from the Loudon et al.[15]

TABLE 1 Pragmatic and explanatory characteristics with illustrative "very pragmatic" study examples of the PRECIS-2 domains

PRECIS-2 domain	Pragmatic	Explanatory	Very pragmatic example ED study
Eligibility criteria—Who is selected to participate in the trial?	Broad eligibility criteria	Limited inclusion criteria	Martel et al.[25]
Recruitment—How are participants recruited into the trial?	Recruitment through usual care	Extra effort beyond usual care (e.g., targeted letters, endorsement from professional societies)	Wang et al.[26]
Setting—Where is the trial being done?	Diverse patient recruitment locations that reflect intended broader population	Single-center or only specialized academic medical center EDs	Schmucker et al.[27]
Organization—What expertise and resources are needed to deliver the intervention?	Uses existing clinical workflow	Relies on increased staffing, training requirements, or other resources	Yiadom et al.[28]
Flexibility (delivery)—How should the intervention be delivered?	Nonrigid prescription of intervention delivery or implementation	Strict protocol with measures to improve intervention compliance	Bernstein et al.[29]
Flexibility (adherence)—What measures are in place to ensure participants adhere to the intervention?	Allows interaction with the intervention as able by the end user	Monitors patient compliance or adherence to the intervention, with potential exclusion	Self et al.[30]
Follow-up—How closely are participants followed up?	No increased follow-up in comparison to usual care	More frequent and/or longer follow-up visits	Melnick et al.[31]
Primary outcome—How relevant is it to participants?	Inherent importance to key stakeholders	Surrogate outcomes, composite outcomes, or outcomes measured at an earlier time than usual care	Tan et al.[32]
Primary analysis—To what extent are all data included?	Intention-to-treat analyses	Consideration of per-protocol analyses	Hess et al.[33]

Abbreviation: PRECIS-2, PRagmatic Explanatory Continuum Indicator Summary 2.

were considered to have taken a "very pragmatic" approach within a specific domain (Table 1).

1. Eligibility criteria—Enrolling all individuals in a clinical setting with minimal eligibility criteria is a hallmark of PCTs. If an intervention has been previously shown to not be beneficial for children, then a reasonable pragmatic design decision would be to exclude only that population. However, if there remains clinical equipoise regarding an intervention's efficacy across populations of interest, then the most pragmatic approach would be to have very limited exclusion criteria.
2. Recruitment—Pragmatic approaches to recruitment might consider only patients who present for health care with a condition of interest, without overt recruitment effort. ED trials may be more likely to fall toward the pragmatic end of the spectrum for this domain, insofar as investigators opportunistically enroll participants who present for acute clinical care rather than relying on healthy volunteers or individuals seen in clinic settings with prediagnosed conditions (e.g., cancer).
3. Setting—Trialists must consider the match between the trial setting and where the results will be applied. At a glance, it may seem that all ED trials would be pragmatic with respect to the setting. However, conducting a trial in highly resourced EDs with immediate access to neuro- and cardiac-intensivists as well as trauma may limit the pragmatic nature if the findings are intended for a broader population. Conducting the trial in diverse ED settings will increase the trial's PRECIS-2 pragmatic score.
4. Organization—A central tenet to this domain is the expertise and resources needed to deliver the intervention, with the goal of a PCT being to add the "successful" intervention to the usual care of the condition or population of interest without additional resources or health care staff. This domain offers a high degree of risk for trialists to move away from the pragmatic end of the spectrum, given a potential reliance on a greater number of staff, increased training requirements, or additional resources.
5. Flexibility (delivery)—Real-world clinical practice suggests that clinicians will implement an intervention with varying fidelity to the protocol. A pragmatic approach to this domain would be for investigators to not rigidly prescribe intervention implementation. Specific mention in the study protocol of allowance of cointerventions, managing side effects of the intervention, or tightly defining intervention delivery to maximize effect suggests a more explanatory approach.
6. Flexibility (adherence)—Variation in adherence to clinical interventions is a reality of designing trials. Highly pragmatic studies have limited compliance or engagement requirements and instead allow the end user to modify the intervention within certain constraints. This includes observing variation in compliance and not excluding nonadherent patients to capture whether the

intervention will have an impact on outcomes despite the presence of this real-world challenge.

7. Follow-up—Compared to usual care, a pragmatic approach to this domain involves trials that have no more follow-up than usual care and do not rely on additional data collection. ED trialists are typically interested in the brief follow-up time period immediately after receipt of the intervention. If follow-up does not differ in care management compared to usual care, then trialists may gather short-term outcomes without undue burden to patients or compromise to the pragmatic nature of the domain.
8. Primary outcome—A general principle of PCTs is that there should be relatively few endpoints, and the ones chosen should be easily measured and salient to stakeholders. If unknown at the time of investigation, outcomes of significance to stakeholders may be determined by gathering qualitative data through interviews or focus groups. Pragmatic measurement of outcomes may also be completed by EHR surveillance, as opposed to the intensive follow-up assessments often done with more explanatory trials. Surrogate outcomes (e.g., process outcomes or other lead metrics) can be misleading due to potential nonlinear relationships between some surrogate endpoints and the clinical outcome of interest. However, these may be more readily available and thus appropriate for interpreting the impact of identified results. Composite endpoints, combining at least two component outcomes of interest, may be difficult to interpret given the wide variability of clinical importance with regards to individual component endpoints.[34] Resultantly, the pragmatic nature of a study may be decreased when investigators select surrogate outcomes only important to clinicians (e.g., blood test), use a composite outcome with component parts not important to patients, or measure short-term outcomes when longer term outcomes are more relevant.[35]
9. Primary analysis—Using all available data in an intention-to-treat analysis is generally recognized as a more pragmatic approach. Excluding nonadherent participants (per-protocol analysis) or study sites that recruit a lower volume would suggest a more explanatory approach to this domain. However, a per-protocol analysis may be considered as a secondary or explanatory analysis for future hypothesis generation.

While ED-based PCTs are available and discussed further in the following sections, few have explicitly mapped their study to the PRECIS-2 wheel diagram. CRASH2 trial is one well-known example of an ED-based pragmatic trial that started with the end in mind. Researchers first thought how the outcome data would be used in clinical practice, forcing investigators to shift along the spectrum from a more explanatory approach to a more pragmatic approach to allow clinicians to use antifibrinolytic therapy in "patients who the treating clinician thinks are bleeding or at significant risk of bleeding." As a PCT, CRASH2 benefited from its broad inclusion criteria, generalizability of findings, and a reduced risk of bias incorporation.[16,36] Another example of an ED-based pragmatic trial is the Resuscitation Outcomes Consortium Pragmatic Airway Resuscitation Trial (PART) comparing the effectiveness of initial laryngeal tube insertion and endotracheal tube insertion on outcomes in adult out-of-hospital cardiac arrest. The PART benefited from a "very pragmatic" approach in the eligibility, recruitment, setting, organization, flexibility (adherence), and follow-up domains of the PRECIS-2 framework.[26]

ADDITIONAL STUDY CONSIDERATIONS FOR PCTs

Aside from domains within the PRECIS-2 framework, ED-based investigators must also consider several other decision points during the design phase. While not completely exhaustive, this list includes randomization, human subjects concerns, and EHR integration.

Randomization

Randomization is the process of assigning participants by chance to groups that receive different treatments, without influence by the investigator. Randomization offers two critical benefits, namely, the reduction of selection bias and the avoidance of known and, more importantly, unknown confounders. Bias in clinical trials is described as a systematic error that encourages one result or outcome over another based on methodological errors. While chance imbalance may arise,[37] selection bias includes systematic differences in baseline characteristics between the compared groups and can generally be overcome with adequate randomization techniques and trial sample size.

Specific techniques of randomization depend on the randomization unit, including individual participants or clusters (e.g., entire EDs, health systems) with cluster randomization particularly relevant to PCTs. Individual randomization results in each participant receiving an experimental intervention, a comparator therapy (e.g., usual care), or a placebo when no interaction among trial participants is anticipated after randomization. Cluster randomization involves groups of participants being assigned to the same study conditions in which interaction among participants within the same clustered is anticipated or unavoidable before and after randomization. Key questions for investigators to consider in determining if individual or cluster randomization is appropriate include[38]: Is the phenomenon of interest something that takes place primarily at the level of the individual patient or study participant (e.g., response to an experimental drug)? Is the proposed intervention delivered at the level of the group (e.g., adoption of a new treatment guideline)? If randomized at the individual level, would it be difficult for clinicians to avoid contamination? Cluster randomized trials are preferred when there is potential for significant treatment contamination—i.e., participants randomized to the control group may receive aspects of the intervention. Contamination has the potential to reduce the study's reliability and can also decrease the observed differences between groups, therefore biasing results toward the null hypothesis.[39] To be discussed below, two key examples of cluster-randomized trials include parallel and stepped-wedge designs.[40] Of note, clinician-level

randomization may be considered in either individual or cluster-randomized techniques depending on the target of the intervention and the objective of the trial.

From an analytic perspective, it is important to note that outcomes of participants within a cluster may be correlated. The intraclass correlation coefficient (ICC) is a standard measure of how similar outcomes of individuals within the same cluster are likely to be.[41] With ICC values ranging from 0 to 1, individual outcomes within clusters are increasingly correlated with one another as the ICC increases. Intuitively, if the ICC is zero, then a cluster-randomized trial will often behave just like an individual-level RCT. In the most extreme case when the ICC is 1, a cluster-randomized trial will behave like an RCT with a sample size equal to the number of clusters because the outcomes will simply be identical within each cluster. In practice, the reported ICC values are usually in the 0.05–0.1 range. However, it has been noted that even a small ICC can lead to a substantial sample size inflation relative to an individual level RCT, especially when the cluster size becomes large.[42] If considering cluster randomization, the goal should be to determine the randomization unit that yields the most clusters, while still being feasible to deliver the intervention without risk of contamination. A greater number of clusters provides improved statistical properties when analyzing cluster-level data by increasing power and reducing potential imbalances of baseline characteristics. Additionally, there are several approaches to balance between-cluster differences, including pair matching, stratification, and constrained randomization.[39,43] Prior to planning the study, emergency investigators should partner with biostatisticians to optimize randomization approaches and determine a preliminary ICC to facilitate accurate sample size estimation.[44]

Human subjects concerns in PCTs

Ethical review from institutional review boards (IRBs) is a key component in the conduct of all clinical research,[45] with the implementation of PCTs often accompanied by unique challenges.[46] With PCTs often recruiting across several institutions, lead investigators typically will use a single, centralized IRB to eliminate redundant reviews of the same material across multiple sites with reliance agreements between local intuitional IRBs and the central IRB. These reliance agreements defer to the centralized IRB's protocol determinations thereby overcoming the need for local amendment determinations across study sites and substantially increasing efficiency of the protection of human subjects review process.[45] To further streamline IRB review, investigators may benefit from convening calls with IRB administrators before the protocol is submitted at each institution to help understand up-front concerns that may delay participant enrollment.

Regulatory boards often recommend that any human subject provide written informed consent. However, proponents of PCTs have noted that existing standards for informed consent are incompatible with a pragmatic study's nature and intent.[47] Waivers of informed consent are becoming increasingly common, with investigators required to meet four criteria of the Common Rule (45 Code of Federal Regulations [CFR] 46.116)[48,49]: 1) research must present no greater than minimal risk to subjects, 2) waiver or alteration will not adversely affect the rights and welfare of the subjects, 3) research could not practicably be carried out without the waiver or alteration, and 4) whenever appropriate, the subjects will be provided with additional pertinent information after participation. To be considered by ED investigators, public approval has also been expressed for exception from informed consent (EFIC) in emergency clinical trials.[50] As an ED-based example, Martel et al.[25] used an EFIC to study the treatment of acute agitation given previously described limitations with obtaining informed consent among intoxicated ED patients.[51] In collaboration with the institutional IRB, the authors additionally performed "community consultation," by reviewing the study protocol with local detoxification facilities, acute psychiatric treatment facilities, and residents of facilities for homeless patients with alcohol use disorder.[25]

It is important to note that the third criterion (i.e., waiver or modification of planned consent) is not met by inconvenience, cost, time, speed, or other resource demands to the team of investigators. In the case of ED-based PCTs, "practicality" does not need to refer solely to the ability to conduct the study *at all*, but also includes if the study's goals are unlikely to be achieved.[52] Counter to the goal of PCTs, nonroutine workflow procedures associated with informed consent process can hinder recruitment, introduce selection bias, and impact generalizability. In determining whether a waiver or modification of planned consent may be appropriate for ED-based PCTs, three questions pertaining to practicality outlined previously by Kalkman et al.[48] should be considered. First, "Is the loss of data quality caused by selection bias or inability to recruit a sufficient number of patients, and can this loss be reasonably attributed to regulatory consent?" Second, "Is the expected loss of data quality so extreme that the trial is not worth doing?" Finally, "Apart from modifying regulatory consent, are there no means or resources available that may be reasonably expected to counter the loss of data quality?" Aside from a waiver of consent, additional options for consent in ED-based PCTs may include broadcast notification,[53] opt-out consent,[54] a "short-form" consent,[55] or electronic consent.[56]

EHR integration

An additional key study consideration is the role of the EHR in ED-based PCTs. Integration of PCT interventions within existing clinical workflows and data streams is essential to limit the costs and burden of ongoing intervention testing.[57] EHR integration requires early engagement with local information technology staff to reach consensus on protocol details and to build the necessary technological infrastructure for the trial. Previous PCT studies have used the EHR for eligibility determination, intervention delivery, and outcome measurement, yet prior work has identified challenges. Primarily, these limitations have been related to a lack of EHR data interoperability across

heterogeneous systems, a lack of usable interfaces and standard workflows, and the inadequate ability to collect clinically appropriate outcome data.[58] Aside from more widespread incentives needed to invest in interoperability and the routine collection of patient-reported outcomes, investigators conducting ED-based PCTs would benefit from working with their health care organization to promote EHR standardization (e.g., developing algorithms or computable phenotypes for patient identification) and avoiding overly customized digital interventions at individual sites in multicenter PCTs (Figure 2).[57,59] Phenotyping allows for the identification of patients with concern for particular conditions or potential for improvement after a specific treatment or intervention. Maximizing the phenotyping potential of the EHR may be particularly critical in the time-sensitive ED environment when investigators or clinicians are attempting to overcome treatment response heterogeneity.[14,60]

PCT STUDY DESIGN TYPES AND EXAMPLES

The various study design types discussed within this section offer unique advantages and disadvantages to be considered by emergency investigators conducting PCTs. Historically, the parallel-group trial design type has been favored, but other types such as crossover, factorial, and stepped-wedge designs are becoming more common.[61] Advantages, disadvantages, and rationale for the common randomized PCT study design types with illustrative examples are provided in Table 2. In

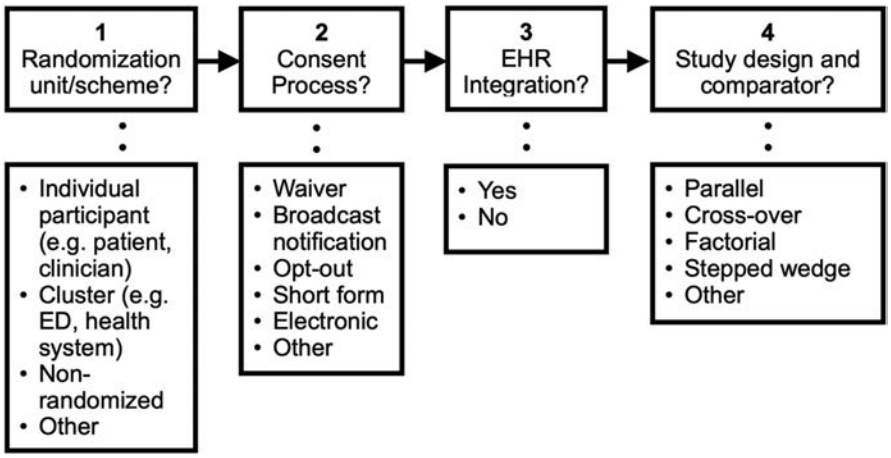


FIGURE 2 Additional key decision points to be considered by investigators conducting ED-based pragmatic clinical trials. Note: Options are bulleted below each decision point. The order of boxes is not the only order in which factors can be considered

TABLE 2 Study selection considerations and illustrative examples for common randomized PCT study design types

Study type	Pros/cons/rationale	Example studies
Parallel	Pro—No inadvertent contamination by unplanned interventions or crossover Con—Often require larger sample sizes due to between-subject variation, which may increase cost and resource utilization Rationale—Most common study type, appropriate if concerns regarding crossover may be present or if the disease or condition being studies may progress over time.	Balamuth et al.,[62] Melnick et al.,[63] Alshahrani et al.,[64] Hess et al.[65]
Crossover	Pro—Comparison of treatment effects both *within and between* participants can increase power Con—Risk of contamination if the intervention cannot be turned "on" and "off" without residual practices being carried over from one period to the next; duration of follow-up generally longer Rationale—Appropriate if concerns exist regarding temporal confounders or significant population variation that may prevent balanced distribution between groups	Wang et al.,[26] Anderson et al.,[67] Self et al.[30]
Factorial	Pro—Efficient, in that multiple research questions may be answered with limited sample sizes Con—Complex in design and statistical analysis; difficulty meeting inclusion criteria for both intervention(s) or components Rationale—Allows assessment of several intervention(s)/components and even interactions between them, often providing information whether varying levels or doses of an intervention affects different populations in different ways	Drummond et al.,[70] Bernstein et al.,[71] Abernethy et al.,[72] Charlesworth et al.,[73] Rickard et al.[74]
Stepped wedge	Pro—All participants receive the intervention; comparison of treatment effects both *within* and *between* participants can increase power Con—Increased complexity may require additional statistical expertise and resources; may be subject to temporal confounding Rationale—Developed to address feasibility and ethical concerns that all participants should eventually receive the intervention within the study time frame when the intervention is anticipated to produce a positive outcome.	Grudzen et al.,[76] Kerber et al..[77] Vaillancourt et al.[78]

Abbreviation: PCT, pragmatic clinical trial.

parallel-group trials, participants are randomized to different study arms and stay within their assigned group for the duration of the study. Parallel-group trials in the ED have studied pediatric septic shock treatment,[62] computerized clinical decision support to streamline the workflow for ED-initiated buprenorphine/naloxone and referral,[63] ketamine infusion for acute sickle vaso-occlusive crisis,[64] and shared decision making in low-risk chest pain.[65]

Crossover trial design types overcome the ethical limitations of a placebo control group by allowing all participants to receive the intervention, just in a different order. Of note, an adequate washout period must be included to eliminate the carryover effect from the intervention that was first delivered.[66] Previous crossover trials in the ED have assessed optimal airway management in out-of-hospital cardiac arrest,[26] head positioning in acute stroke,[67] and fluid resuscitation in noncritically ill patients.[30]

Factorial (2 × 2 or more) design trials allow study of multiple interventions or multilevel interventions in varying combinations. Although this design type allows for comparisons between smaller sample sizes, investigators conducting full factorial design studies often need to ensure the trial functions as two independent RCTs without interaction between the two treatments.[68] In the presence of interaction between two treatments, variations in treatment effect estimands exist and more sophisticated modeling techniques should be considered to account for such an interaction effect.[69] Factorial designs have rarely been employed in the ED setting to date,[70,71] but studies from other disciplines and settings that could be adapted to the ED setting have assessed palliative care educational outreach,[72] peer support and reminiscence therapy for persons living with dementia and their caregivers,[73] and dressing and securement of peripherally inserted central catheters.[74]

Stepped-wedge designs may be considered in multicenter studies, particularly when all participants (or clusters) can begin in the control group. As time passes, participants (or clusters) are randomized to the treatment arm in a sequential fashion, with all participants (or clusters) receiving the intervention within the study period. Illustrative ED-based pragmatic trials using stepped-wedge designs have addressed multidisciplinary primary palliative care,[75,76] treatment strategies for benign paroxysmal positional vertigo,[77] and paramedic assessment of low-risk trauma patients to selectively transport without immobilization.[78] Group allocation across the PCT study timeframe for common randomized study design types is depicted in Figure 3.

CONCLUSIONS

There is growing interest in ED-based pragmatic clinical trials for their ability to test interventions in real-world settings and develop evidence that is more widely translatable than traditional explanatory research. To ensure the collection of meaningful results, emergency investigators must consider use of the PRECIS-2 tool as well as additional decisions within the study design phase. This article can serve as a guide to navigate the unique advantages and disadvantages of various design types to be considered by investigators pursuing pragmatic clinical trials.

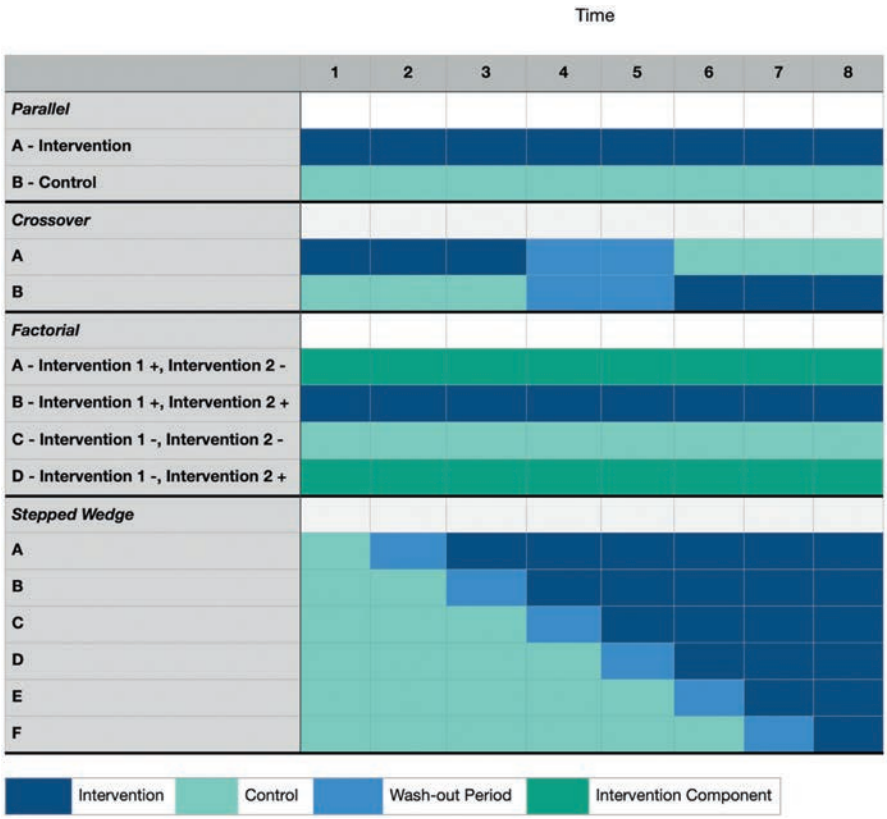


FIGURE 3 Group allocation across the study timeframe for randomized PCT study design types.
Note: Groups are delineated by A–F. PCT, pragmatic clinical trial

AUTHOR CONTRIBUTIONS

Drafting of the manuscript—Cameron J. Gettel, Edward R. Melnick. Critical revision of the manuscript for intellectual content—Cameron J. Gettel, Maame Yaa A.B. Yiadom, Steven L. Bernstein, Corita R. Grudzen, Bidisha Nath, Fan Li, Ula Hwang, Erik P. Hess, Edward R. Melnick. Acquisition of funding—Edward R. Melnick.

CONFLICT OF INTEREST

The authors report no conflicts of interest.

ORCID

Cameron J. Gettel https://orcid.org/0000-0002-6249-1023
Edward R. Melnick https://orcid.org/0000-0002-6509-9537

REFERENCES

1. Broderick KB, Ranney ML, Vaca FE, et al. Study designs and evaluation models for emergency department public health research. *Acad Emerg Med*. 2009;16(11):1124-1131.
2. Califf RM, Zarin DA, Kramer JM, Sherman RE, Aberle LH, Tasneem A. Characteristics of clinical trials registered in Clinicaltrials.gov, 2007-2010. *JAMA*. 2012;307(17):1838-1847.
3. Ioannidis JPA. Why most published research findings are false. *PLoS Med*. 2005;2(8):e124.
4. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis*. 1967;20(8):637-648.
5. Califf RM, Sugarman J. Exploring the ethical and regulatory issues in pragmatic clinical trials. *Clin Trials*. 2015;12(5):436-441.
6. Ford I, Norrie J. Pragmatic trials. *N Engl J Med*. 2016;375(5):454-463.
7. Pragmatic clinical studies. Patient-Centered Outcomes Research Institute. 2016. Accessed October 20, 2021. https://www.pcori.org/research-results/pragmatic-clinical-studies
8. Pragmatic clinical trials: testing treatments in the real world. National Institute on Aging. 2017. Accessed October 20, 2021. https://www.nia.nih.gov/research/blog/2017/06/pragmatic-clinical-trials-testing-treatments-real-world
9. Bernstein SL, Stoney CM, Rothman RE. Dissemination and implementation research in emergency medicine. *Acad Emerg Med*. 2015;22(2):229-236.
10. Summary of Q&A from the technical assistant videocast "HEAL Initiative: pragmatic randomized controlled trial of acupuncture for management of chronic low back pain in older adults (UG3/UH3 clinical trial required)". National Center for Complementary and Integrative Health. 2019. Accessed October 20, 2021. https://www.nccih.nih.gov/news/events/heal-acupuncture-2019-summary
11. Lilly and UnitedHealth Group partner on pragmatic study of neutralizing antibody bamlanivimab (LY-CoV555) for COVID-19. Eli Lilly and Company. December 4, 2020. Accessed October 20, 2021. https://investor.lilly.com/news-releases/news-release-details/lilly-and-unitedhealth-group-partner-pragmatic-study
12. Rethinking clinical trials: a living textbook of pragmatic clinical trials. NIH Collaboratory Living Textbook of Pragmatic Clinical Trials. Accessed October 20, 2021. https://rethinkingclinicaltrials.org/
13. R01-equivalents grants: average funding in current and constant dollars. National Institutes of Health. 2021. Accessed March 24, 2022. https://report.nih.gov/nihdatabook/report/158
14. Casey JD, Semler MW. Embedding pragmatic trials into emergency and critical care. 2021. Accessed March 24, 2022. https://dcricollab.dcri.duke.edu/sites/NIHKR/KR/GR-Slides-10-29-21.pdf
15. O'Rourke PP, Carrithers J, Patrick-Lake B, et al. Harmonization and streamlining of research oversight for pragmatic clinical trials. *Clin Trials*. 2015;12(5):449-456.
16. Coats TJ. Future research in emergency medicine: explanation or pragmatism? Large or small? Simple or complex? *Emerg Med J*. 2011;28(12):1004-1007.
17. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ*. 2015;350:h2147.
18. Johnson KE, Neta G, Dember LM, et al. Use of PRECIS ratings in the National Institutes of Health (NIH) Health Care Systems Research Collaboratory. *Trials*. 2016;17:32.
19. Casey JD, Janz DR, Russell DW, et al. Bag-mask ventilation during tracheal intubation of critically ill adults. *N Engl J Med*. 2019;380(9):811-821.
20. Kanzler KE, McGeary DD, McGeary C, et al. Conducting a pragmatic trial in integrated primary care: key decision points and considerations. *J Clin Psychol Med Settings*. 2022;29(1):185-194.
21. Chan CL, Taljaard M, Lancaster GA, Brehaut JC, Eldridge SM. Pilot and feasibility studies for pragmatic trials have unique considerations and areas of uncertainty. *J Clin Epidemiol*. 2021;138:102-114.
22. Gordon KS, Peduzzi P, Kerns RD. Designing trials with purpose: pragmatic clinical trials of nonpharmacological approaches for pain management. *Pain Med*. 2020;21(Suppl 2):S7-S12.
23. Quiñones AR, Mitchell SL, Jackson JD, et al. Achieving health equity in embedded pragmatic trials for people living with dementia and their family caregivers. *J Am Geriatr Soc*. 2020;68(Suppl 2):S8-S13.
24. Norton WE, Loudon K, Chambers DA, Zwarenstein M. Designing provider-focused implementation trials with purpose and intent: introducing the PRECIS-2-PS tool. *Implement Sci*. 2021;16(1):7.
25. Martel ML, Driver BE, Miner JR, Biros MH, Cole JB. Randomized double-blind trial of intramuscular droperidol, ziprasidone, and lorazepam for acute undifferentiated agitation in the emergency department. *Acad Emerg Med*. 2021;28(4):421-434.
26. Wang HE, Schmicker RH, Daya MR, et al. Effect of a strategy of initial laryngeal tube insertion vs endotracheal intubation on 72-hour survival in adults with out-of-hospital cardiac arrest: a randomized clinical trial. *JAMA*. 2018;320(8):769-778.
27. Schmucker AM, Flannery M, Cho J, Goldfeld KS, Grudzen C, EMPallA Investigators. Data from emergency medicine palliative care access (EMPallA): a randomized controlled trial comparing the effectiveness of specialty outpatient versus telephonic palliative care of older adults with advanced illness presenting to the emergency department. *BMC Emerg Med*. 2021;21(1):83.
28. Yiadom MYAB, Domenico H, Byrne D, et al. Randomised controlled pragmatic clinical trial evaluating the effectiveness of a discharge follow-up phone call on 30-day hospital readmissions: balancing pragmatic and explanatory design considerations. *BMJ Open*. 2018;8(2):e019600.
29. Bernstein SL, Rosner J, DeWitt M, et al. Design and implementation of decision support for tobacco dependence treatment in an inpatient electronic medical record: a randomized trial. *Transl Behav Med*. 2017;7(2):185-195.
30. Self WH, Semler MW, Wanderer JP, et al. Balanced crystalloids versus saline in noncritically ill adults. *N Engl J Med*. 2018;378(9):819-828.
31. Melnick ER, Nath B, Ahmed OM, et al. Progress report on EMBED: a pragmatic trial of user-centered clinical decision support to implement EMergency department-initiated BuprenorphinE for opioid use Disorder. *J Psychiatr Brain Sci*. 2020;5:e200003.
32. Tan PJ, Khoo EM, Chinna K, et al. Individually-tailored multifactorial intervention to reduce falls in the Malaysian Falls Assessment and Intervention Trail (MyFAIT): a randomized controlled trial. *PLoS One*. 2018;13(8):e0199219.
33. Hess EP, Homme JL, Kharbanda AB, et al. Effect of the head computed tomography choice decision aid in parents of children with minor head trauma: a cluster randomized trial. *JAMA Netw Open*. 2018;1(5):e182430.

34. McCoy CE. Understanding the use of composite endpoints in clinical trials. *West J Emerg Med*. 2018;19(4):631-634.
35. Choosing and specifying endpoints and outcomes, section 2, meaningful endpoints. NIH Pragmatic Trials Collaboratory: Rethinking Clinical Trials. Accessed October 20, 2021. https://rethinkingclinicaltrials.org/chapters/design/choosing-specifying-end-points-outcomes/meaningful-endpoints/
36. CRASH-2 Trial Collaborators. Effects of tranexamic acid on death, vascular occlusive events, and blood transfusion in trauma patients with significant haemorrhage (CRASH-2): a randomized, placebo-controlled trial. *Lancet*. 2010;376(9734):23-32.
37. Zeng S, Li F, Wang R, Li F. Propensity score weighting for covariate adjustment in randomized clinical trials. *Stat Med*. 2021;40(4):842-858.
38. Experimental designs and randomization schemes, section 5, choosing between cluster and individual randomization. NIH Pragmatic Trials Collaboratory: Rethinking Clinical Trials. Accessed November 10, 2021. https://rethinkingclinicaltrials.org/chapters/design/experimental-designs-randomization-schemes-top/considerations-when-choosing-between-a-crt-individual-randomization-designs/
39. Cook AJ, Delong E, Murray DM, Vollmer WM, Heagerty PJ. Statistical lessons learned for designing cluster randomized pragmatic clinical trials from the NIH Health Care Systems Collaboratory Biostatistics and Design Core. *Clin Trials*. 2016;13(5):504-512.
40. Turner EL, Li F, Gallis JA, Prague M, Murray DM. Review of recent methodological developments in group-randomized trials: Part 1 – Design. *Am J Public Health*. 2017;107(6):907-915.
41. Analysis plan, section 2, intraclass correlation. NIH Pragmatic Trials Collaboratory: Rethinking Clinical Trials. Accessed October 20, 2021. https://rethinkingclinicaltrials.org/chapters/design/analysis-plan-top/intraclass-correlation/
42. Murray DM. *Design and Analysis of Group-randomized Trials*. Oxford University Press Inc; 1998.
43. Li F, Lokhnygina Y, Murray DM, Heagerty PJ, DeLong ER. An evaluation of constrained randomization for the design and analysis of group-randomized trials. *Stat Med*. 2016;35(10):1565-1579.
44. Wears RL. Advanced statistics: statistical methods for analyzing cluster and cluster-randomized data. *Acad Emerg Med*. 2002;9(4):330-341.
45. Marquis-Gravel G, Robertson H, Jones WS, et al. Streamlining the institutional review board process in pragmatic randomized clinical trials: challenges and lessons learned from the Aspirin Dosing: a Patient-centric Trial Assessing Benefits and Long-term Effectiveness (ADAPTABLE) trial. *Trials*. 2021;22(1):90.
46. Roberts MK, Fisher DM, Parker LE, et al. Ethical and regulatory concerns in pragmatic clinical trial monitoring and oversight. *Ethics Hum Res*. 2020;42(5):29-37.
47. Murdoch B, Caulfield T. Pragmatic clinical trials and the consent process. *Res Ethics*. 2018;14(2):1-14.
48. Kalkman S, Kim SYH, van Thiel GJMW, Grobbee DE, van Delden JJM. Ethics of informed consent for pragmatic trials with new interventions. *Value Health*. 2017;20(7):902-908.
49. 45 CFR 46. U.S. Department of Health & Human Services, Office for Human Research Protections. 2018. Accessed October 20, 2021. https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html#
50. Feldman WB, Hey SP, Franklin JM, Kesselheim AS. Public approval of Exception from Informed Consent in emergency clinical trials: a systematic review of community consultation surveys. *JAMA Netw Open*. 2019;2(7):e197591.
51. Martel ML, Klein LR, Miner JR, et al. A brief assessment of capacity to consent instrument in acutely intoxicated emergency department patients. *Am J Emerg Med*. 2018;36(1):18-23.
52. Kim SYH, Miller F. Waivers and alterations to consent in pragmatic clinical trials: respecting the principle of respect for persons. The Hastings Center. Accessed February 24, 2017. October 20, 2021. https://www.thehastingscenter.org/irb_article/waivers-and-alterations-to-consent-in-pragmatic-clinical-trials-respecting-the-principle-of-respect-for-persons/
53. Huang SS, Septimus E, Kleinman K, et al. Targeted versus universal decolonization to prevent ICU infection. *N Engl J Med*. 2013;368(24):2255-2265.
54. Coronado GD, Vollmer WM, Petrik A, et al. Strategies and opportunities to STOP colon cancer in priority populations: design of a cluster-randomized pragmatic trial. *Contemp Clin Trials*. 2014;38(2):344-349.
55. McKinney RE, Beskow LM, Ford DE, et al. Use of altered informed consent in pragmatic clinical research. *Clin Trials*. 2015;12(5):494-502.
56. Use of electronic informed consent: questions and answers. U.S. Department of Health and Human Services. December 2016. Accessed October 20, 2021. https://www.hhs.gov/ohrp/regulations-and-policy/guidance/use-electronic-informed-consent-questions-and-answers/index.html#_ftn1
57. Richesson RL, Marsolo KS, Douthit BJ, et al. Enhancing the use of EHR systems for pragmatic embedded research: lessons from the NIH Health Care Systems Research Collaboratory. *J Am Med Inform Assoc*. 2021;28(12):2626-2640. doi:10.1093/jamia/ocab202
58. Al Rifai M, Itchhaporia D, Virani SS. Pragmatic clinical trials – ready for prime time? *JAMA Netw Open*. 2021;4(12):e2140212.
59. Sinsky CA, Bavafa H, Roberts RG, Beasley JW. Standardization vs customization: finding the right balance. *Ann Fam Med*. 2021;19(2):171-177.
60. NIH Pragmatic Trials Collaboratory. Electronic health records-based phenotyping. Accessed March 24, 2022. https://rethinkingclinicaltrials.org/chapters/conduct/electronic-health-records-based-phenotyping/definitions/
61. Nair B. Clinical trial designs. *Indian Dermatol Online J*. 2019;10(2):193-201.
62. Balamuth F, Kittick M, McBridge P, et al. Pragmatic pediatric trial of balanced versus normal saline fluid in sepsis: the PRoMPT BOLUS randomized controlled trial pilot feasibility study. *Acad Emerg Med*. 2019;26(12):1346-1356.
63. Melnick ER, Jeffery MM, Dziura JD, et al. User-centred clinical decision support to implement emergency department-initiated buprenorphine for opioid use disorder: protocol for the pragmatic group randomised EMBED trial. *BMJ Open*. 2019;9(5):e028488.
64. Alshahrani MS, AlSulaibikh AH, ElTahan MR, et al. Ketamine administration for acute painful sickle cell crisis: a randomized control trial. *Acad Emerg Med* 2022;29(2):150-158.
65. Hess EP, Hollander JE, Schaffer JT, et al. Shared decision making in patients with low risk chest pain: prospective randomized pragmatic trial. *BMJ*. 2016;355:i6165.
66. Li F, Forbes AB, Turner EB, Preisser JS. Power and sample size requirements for GEE analyses of cluster randomized crossover trials. *Stat Med*. 2019;38(4):636-649.
67. Anderson CS, Arima H, Lavados P, et al. Cluster-randomized, crossover trial of head positioning in acute stroke. *N Engl J Med*. 2017;376(25):2437-2447.
68. McAlister FA, Straus SE, Sackett DL, Altman DG. Analysis and reporting of factorial trials: a systematic review. *JAMA*. 2003;289(19):2545-2553.
69. Tian Z, Esserman D, Tong G, et al. Sample size calculation in hierarchical 2 x 2 factorial trials with unequal cluster sizes. *Stat Med*. 2022;41(4):645-664.
70. Drummond C, Deluca P, Coulton S, et al. The effectiveness of alcohol screening and brief intervention in emergency departments: a multicentre pragmatic cluster randomized controlled trial. *PLoS One*. 2014;9(6):e99463.

71. Bernstein SL, Dziura J, Weiss J, et al. Tobacco dependence treatment in the emergency department: a randomized trial using Multiphase Optimization Strategy. *Contemp Clin Trials*. 2018;66:1-8.
72. Abernethy AP, Currow DC, Hunt R, et al. A pragmatic 2 x 2 x 2 factorial cluster randomized controlled trial of educational outreach visiting and case conferencing in palliative care – methodology of the Palliative Care Trial [ISRCTN 81117481]. *Contemp Clin Trials*. 2006;27(1):83-100.
73. Charlesworth G, Burnell K, Crellin N, et al. Peer support and reminiscence therapy for people with dementia and their family carers: a factorial pragmatic randomised trial. *J Neurol Neurosurg Psychiatry*. 2016;87(11):1218-1228.
74. Rickard CM, Marsh NM, Webster J, et al. Peripherally InSerted CEntral catheter dressing and securement in patients with cancer: the PISCES trial. Protocol for a 2x2 factorial, superiority randomised controlled trial. *BMJ Open*. 2017;7(6):e015291.
75. Chung FR, Turecamo S, Cuthel AM, Grudzen CR, Investigators PRIM-ER. Effectiveness and reach of the Primary Palliative Care for Emergency Medicine (PRIM-ER) pilot study: a qualitative analysis. *J Gen Intern Med*. 2021;36(2):296-304.
76. Grudzen CR, Brody AA, Chung FR, et al. Primary palliative care for emergency medicine (PRIM-ER): protocol for a pragmatic, cluster-randomised, stepped wedge design to test the effectiveness of primary palliative care education, training and technical support for emergency medicine. *BMJ Open*. 2019;9(7):e030099.
77. Kerber KA, Damschroder L, McLaughlin T, et al. Implementation of evidence-based practice for benign paroxysmal positional vertigo in the emergency department: a stepped-wedge randomized trial. *Ann Emerg Med*. 2020;75(4):459-470.
78. Vaillancourt C, Charette M, Taljaard M, et al. Pragmatic strategy empowering paramedics to assess low-risk trauma patients with the Canadian C-spine rule and selectively transport them without immobilization: protocol for a stepped-wedge cluster randomized trial. *JMIR Res Protoc*. 2020;9(6):e16966.

How to cite this article: Gettel CJ, Yiadom MY, Bernstein SL, et al. Pragmatic clinical trial design in emergency medicine: Study considerations and design types. *Acad Emerg Med*. 2022;29:1247-1257. doi: 10.1111/acem.14513

Received: 13 March 2022 | Revised: 31 May 2022 | Accepted: 21 June 2022

DOI: 10.1111/acem.14553

RESEARCH LETTER

Laboratory trends in severe MIS-C

Multisystem inflammatory syndrome in children (MIS-C) has been recognized as a rare illness associated with a recent coronavirus disease 2019 (COVID-19) infection and consists of a fever, laboratory evidence of inflammation, and multisystem (≥2) involvement.[1,2] A spectrum of illness has been appreciated, and those with severe MIS-C may exhibit significant cardiac disease, shock, and death.[3]

Given the wide clinical spectrum, identifying children at risk for severe disease is critical to optimizing care. Those at risk for severe MIS-C have been shown to more likely be non-Hispanic black, obese, male, 13–18 years of age, and have higher peak laboratory markers (troponin, brain natriuretic peptide [BNP], ferritin, C-reactive protein [CRP], interleukin 6 [IL-6]).[4] While studies have attempted to identify laboratory predictors of severe MIS-C, they have not accounted for the patients' day of illness (DOI).[5] It is unclear when each of the laboratory markers rise. The objective of this study is to describe the temporal relationship between serial laboratory markers and symptom onset for severe and non-severe MIS-C. Our overall hypothesis was that that DOI biomarker trajectories would differ between groups, with greater overall concentrations in severe versus non-severe cases.

We performed a retrospective chart review of all patients with MIS-C treated at a single institution from April 1, 2020, to February 15, 2021. Patients met published Centers for Disease Control (CDC) criteria for MIS-C, and severe MIS-C was defined as cardiovascular shock (need for vasopressors), cardiac arrhythmia requiring intervention, or respiratory failure (need for positive pressure ventilation). DOI was defined as the onset of any symptom related to MIS-C, with the first day defined as day 1. Symptom onset was identified using the emergency department (ED) note and was identified on all patients. If symptom onset was not explicitly mentioned in the ED note, the hospital admission note was used. Patients who met criteria for severe MIS-C at any point during the hospital course were included in analysis. Descriptive data for the overall cohort are presented as medians with inter-quartile range (IQR) or proportions. Laboratory values on the day of ED presentation were compared using the Wilcoxon rank-sum test. Linear mixed-effects models were used to examine the relationship between repeated measures of laboratory values (outcome) in those with moderate versus severe MIS-C. Based on model diagnostics and visual inspection of residual plots, all lab values, except albumin, were log transformed. Model outputs for transformed variables are interpreted as the percentage difference in marginal mean lab value in severe versus moderate MIS-C (calculated as one-parameter estimate × 100). Covariates were selected based on review of prior studies as well as biologic and/or clinical plausibility and included DOI, age, race, and DOI-MIS-C category interaction (DOI × MIS-C); a random intercept was included in all models. Biologic sex was considered as a covariate but was non-significant for all models and worsened model fit so was excluded. Model fit was assessed using Bayesian Information Criteria and covariance structure was selected using the null model likelihood ratio test. A two-tailed significance level was set at 0.05. Data analysis was performed using SAS version 9.4 (SAS Institute). This study was approved by the Institutional Review Board of Central Michigan University.

During the study period, we identified 32 patients with severe and 29 patients with non-severe MIS-C with single ED visits. Those with severe disease were older (7.6 vs. 4.9 years) and more often black (59.4% vs. 42.9%), with roughly half (54%) of patients in both cohorts being female. There was no difference in median days of symptoms at ED presentation (4.0 [IQR 3.0–5.0 days] in severe vs. 5.0 [IQR 2.5–7.0 days] for non-severe, $p = 0.46$). Criteria for severe MIS-C were met in the ED for 10 patients, with 15 meeting criteria on hospital day 1, 3 on hospital day 2, and 3 on hospital day 3, and 1 on hospital day 6. Median DOI for meeting severe criteria was 6 (IQR 4–6). In the severe group, 31 patients were treated with intravenous gamma globulin (IVIG) and 22 with steroid administration. In non-severe patients, 20 were treated with IVIG and 2 with steroids. The median DOI was similar for IVIG administration in severe (5.0, IQR 4.0–6.0 days) versus moderate (6.0, 3.0–8.0 days) groups, $p = 0.29$. A total of 24 patients received at least one dose of steroids, with 22 in the severe and 2 in the moderate group. Median DOI for first steroid dose was 6.0 (IQR 3.0–8.0 days) for severe versus 8.0 days for moderate.

While not all subjects had biomarkers checked on each hospital day, any patient with at least one measured value of the biomarker of interest was included in the model for that outcome. Description of missing biomarker data is provided in the supplement; there were no missing values for any covariates. On day of ED presentation, patients who developed severe MIS-C had greater median troponin (29 ng/ml, IQR 11–183 vs. 11 ng/ml, IQR 4–45.5, $p = 0.027$); ferritin (311.8 μg/L, IQR 188.3–452.7 vs. 137.9 μg/L, IQR 116.2–246.2, $p = 0.03$); and d-dimer (2.44 ng/ml, IQR 1.38–32.9 vs. 1.38 ng/ml, IQR 0.74–2.32, $p = 0.034$).

For the mixed-effects model analysis, only mean BNP and CRP varied significantly by DOI (DOI × MIS-C) between non-severe and severe MIS-C. For CRP, severe MIS-C patients had greater initial CRP that declined faster than non-severe MIS-C. For BNP, initial levels were greater and subsequently declined in severe MIS-C, whereas

Supervising Editor: Dr. Elizabeth Alpern

 © 2022 Society for Academic Emergency Medicine.

non-severe patients had a slow increase (Figure 1). For other laboratory parameters, trends by DOI were similar for all patients, but the severe MIS-C group had greater mean troponin (185%, 95% confidence interval [CI] 46–456, $p = 0.002$), ferritin (46%, 95% CI −1.1 to 117, $p = 0.05$), and d-dimer (116%, 95% CI 46–218, $p = 0.0002$), whereas albumin was 0.6 g/dl lower (95% CI −0.8 to −0.4, $p < 0.001$).

Patients with severe MIS-C have been shown to have higher peak laboratory markers compared to non-severe MIS-C and severe COVID-19.[4] Our data demonstrate this same trend upon ED presentation with a higher median troponin, d-dimer, and ferritin. However, when the laboratory trends of each patient were compared by DOI, only the BNP and CRP showed a significant difference. The severe cohort had an earlier increase in both the BNP and CRP, and then a steady decline over the course of the illness. Patients with non-severe disease presented with a lower BNP early in illness, which slowly rose over time. BNP is secreted by cardiac myocytes in the ventricles in response to wall stretch or mechanical load, and levels typically rise reliably with cardiac dysfunction.[6] Thus, the etiology of the decline in BNP levels in patients with severe cardiac disease is unclear. This could represent the natural history of the disease, recognition of cardiac dysfunction with appropriate treatment (fluid restriction or diuresis), or residual confounding. The influence of treatments, including IVIG and steroids, on biomarker concentrations and trajectories remains unclear as we did not adjust for these given our small sample size and the complexities of modeling multiple time-varying treatments and their interaction. Presentation

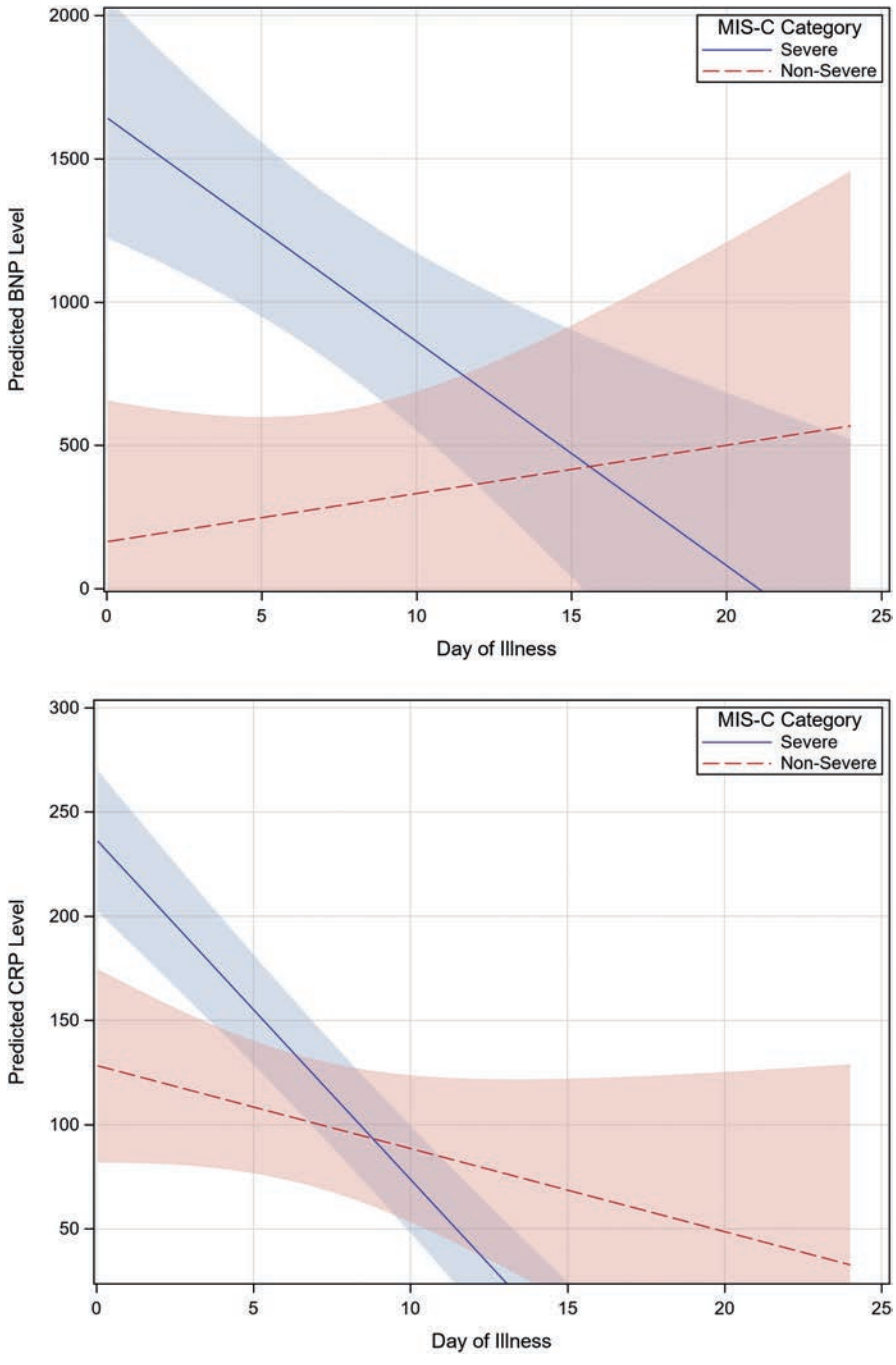


FIGURE 1 Interaction plot of DOI with MIS-C category showing that the trajectory of BNP and CRP over the course of illness is different for those with severe versus non-severe MIS-C. Fit was computed at mean age (7 years) and for black race.

DOI and initiation of treatments were similar, and while one would expect severe patients to have greater concentrations of inflammatory markers, and treatment-associated declines in all patients, the BNP trajectory in moderate patients is less intuitive. Determination of whether this represents natural disease course, confounding due to treatments administered, or other patient-level factors requires further study.

Our study is limited by sample size and risks inherent to a retrospective study from a single site. We did not adjust for multiple comparisons (of investigated biomarkers) as our goal was to minimize false negatives and crude adjustment methods (e.g., Bonferroni) drastically reduce power.[7] While this can increase the false positive rate, we felt the overall benefit outweighed the potential harms given that these biomarkers are routinely obtained but there remains a paucity of data on temporal laboratory trends in MIS-C. Follow-up studies from larger data sets are needed to confirm our findings.

In conclusion, we found that patients with severe MIS-C had higher levels of laboratory markers when compared to non-severe MIS-C, with similar trends across DOIs in both groups for most biomarkers. However, the CRP and BNP levels were significantly greater among patients with severe MIS-C earlier in the disease course. Further investigation may reveal that these two markers may help identify patients at risk for severe MIS-C prior to clinical deterioration.

AUTHOR CONTRIBUTIONS

Study concept and design (Curt Stankovic, Amy DeLaroche, Rajan Arora, Priya Spencer), acquisition of data (Priya Spencer), data analysis and interpretation (Curt Stankovic, Robert R. Ehrman, Amy DeLaroche, Robert R. Ehrman), drafting of the manuscript (Curt Stankovic, Robert R. Ehrman, Rajan Arora, Amy DeLaroche, Priya Spencer), statistical expertise (Robert R. Ehrman), funding acquisition (Curt Stankovic).

CONFLICT OF INTEREST

We have no conflict of interest.

PRESENTATION

This manuscript will be presented at the Pediatric Academic Society meeting on April 22, 2022.

GRANT MONEY FOR INVESTIGATOR INITIATED RESEARCH

C.S. reports grant money to Central Michigan University to conduct research conceived and written by Curt Stankovic from the Children's Foundation.

Curt Stankovic MD[1]
Amy DeLaroche MBBS[1]
Rajan Arora MD[1]
Priya Spencer MPH[1]
Robert R. Ehrman MD, MS[2]

[1]*Central Michigan University, Discipline of Pediatrics, Children's Hospital of Michigan, Detroit, Michigan, USA*
[2]*Wayne State University, Department of Emergency Medicine, Detroit, Michigan, USA*

Correspondence
Curt Stankovic, MD, Central Michigan University, Discipline of Pediatrics, Children's Hospital of Michigan, Detroit, MI 48201, USA.
Email: cstankov@dmc.org

ORCID

Curt Stankovic https://orcid.org/0000-0001-7886-6988
Amy DeLaroche https://orcid.org/0000-0001-5425-6771
Rajan Arora https://orcid.org/0000-0001-9816-8140
Robert R. Ehrman https://orcid.org/0000-0001-9730-8471

REFERENCES

1. Centers for Disease Control and Prevention Emergency preparedness and response: information for healthcare providers about multisystem inflammatory syndrome in children. Accessed 2/3/2022. https://www.cdc.gov/mis/mis-c/hcp/index.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fmis%2Fhcp%2Findex.html
2. Feldstein LR, Tenforde MW, Friedman KG, et al. Characteristics and outcomes of US children and adolescents with multisystem inflammatory syndrome in children (MIS-C) compared with severe acute COVID-19. *JAMA*. 2021;325(11):1074-1087. doi:10.1001/jama.2021.2091
3. Sethuraman U, Kannikeswaran N, Ang J, et al. Multisystem inflammatory syndrome in children associated with novel coronavirus, SARS-CoV-2, presentations to a pediatric emergency Department in Michigan. *Am J Emerg Med*. 2020;164-167. doi:10.1016/j.ajem.2020.10.035
4. Abrams JY, Oster ME, Godfred-Cato SE, et al. Factors linked to severe outcomes in multisystem inflammatory syndrome in children (MIS-C) in the USA: a retrospective surveillance study. *Lancet Child Adolesc Health*. 2021;5(5):323-331. doi:10.1016/S2352-4642(21)00050-X. PMID: 33711293; PMCID: PMC7943393
5. Delaroche A, Stankovic C, Ehrman R, et al. ED screening for multisystem inflammatory syndrome in children. *Am J Emerg Med*. 2020. doi:10.1016/.j.ajem.2020.09.076
6. Nasser N, Perles Z, Rein AJJT, Nir A. NT-proBNP as a marker for persistent cardiac disease in children with history of dilated cardiomyopathy and myocarditis. *Pediatr Cardiol*. 2006;27(1):87-90. doi:10.1007/s00246-005-1027-z. PMID: 16132296
7. Greenland S. Analysis goals, error-cost sensitivity, and analysis hacking: essential considerations in hypothesis testing and multiple comparisons. *Paediatr Perinat Epidemiol*. 2021;35:8-23. doi:10.1111/ppe.12711

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Received: 3 May 2022 | Revised: 21 June 2022 | Accepted: 12 July 2022

DOI: 10.1111/acem.14562

RESEARCH LETTER

Novel quality assessment methodology in focused cardiac ultrasound

Cardiac point-of-care ultrasound (POCUS) has become a fundamental component of the evaluation of patients in the emergency department (ED) to diagnose cardiac pathology.[1] Quality assessment (QA) is one of the six required elements of diagnostic POCUS examinations per the American College of Emergency Physicians (ACEP).[2] QA is routinely performed to ensure standard of patient care is met and to assess for competency, particularly at the trainee level. It has recently been described that 82% of EM residency programs report use of QA as an assessment tool.[3] It is imperative to provide a reliable scoring system to limit inconsistencies in the way we are measuring clinical skill as it relates to ultrasound.

The current grading scale used for QA that is endorsed by ACEP was developed from a consensus report of emergency ultrasound leaders indicating a need for a systematic method to report and communicate POCUS findings.[2] The ACEP grading scale is a nonspecific grading classification that applies regardless of the type of study performed.[2] This contrasts with other QA grading systems that have been described in an organ-specific manner. Examples of these systems include focused cardiac ultrasound assessment demonstrated by Kimura et al.[4] and focused gynecological emergency ultrasound examination by Salomon.[5] The goal of this study was to determine whether a similar, organ-specific grading scale would be a more reliable method of assessment with improved interobserver agreement. Furthermore, we sought to determine whether an organ-specific grading scale had more variance in the scores that are chosen and therefore more focused feedback to the sonographers performing the studies.

We conducted a prospective analysis of the first 200 cardiac POCUS studies performed in our ED that were submitted for QA into our image database in the year of 2020. Four reviewers, who were either emergency ultrasound fellowship trained or current emergency ultrasound fellows with at least 9 months of QA experience, scored each of the studies. Two reviewers used the current ACEP grading scale: 1 = no recognizable structures; 2 = minimally recognizable structures but insufficient for diagnosis; 3 = minimal criteria met for diagnosis, recognizable structures but with some technical or other flaws; 4 = minimal criteria met for diagnosis, all structures imaged well; and 5 = minimal criteria met for diagnosis, all structures imaged with excellent image quality.[2] The two other reviewers used a cardiac-specific grading scale as previous described by Kimura et al.: 0 = no image obtained; 1 = only cardiac motion detected; 2 = chambers and valves grossly resolved; 3 = endocardium and wall thickness seen but incomplete; and 4 = greater than 90% of endocardium and valve motion seen.[3] The primary outcomes were the level of agreement between the reviewers, indicating the reliability of the scoring system, and the variability of the scores given to the studies. This study was approved by the VCU institutional review board.

The ACEP score was on a scale of 1–5, while the organ-specific score was on a scale of 0–4. For equal comparison, we added 1 to the organ-specific grading. For the primary outcome, the intraclass correlation coefficient (ICC) based on two-way random-effect model with a single rater for each grading scale was computed. Ten thousand bootstrapped ICCs were generated to construct 95% confidence intervals (CIs) for both grading systems, and a two-sided one-sample t-test was used to determine if there were differences in the bootstrapped ICCs between the two grading systems. The ICC between reviewers for the ACEP grading scale was 0.54 (95% CI 0.410–0.555) indicating moderate agreement, while the ICC between reviewers using the organ-specific grading scale was 0.75 (95% CI 0.600–0.769) indicating good agreement. These findings were statistically significant with one-sample t-test p-value of <0.0001. The ACEP grading scale mean (±SD) was 3.15 (±0.693) versus 4.16 (±0.967) for the organ-specific grading system. A 95% CI for the variance ratio was constructed to determine whether there were differences in the variability between the two grading systems. The variance of scores using the organ-specific grading scale was found to be more than 1.95 times greater when compared to the scores using ACEP grading scale. The variance among each group of scores using a 95% CI of the bootstrapped variance ratio (1.49–2.51), the organ-specific grading scale was found to have significantly more variability than the ACEP scoring system. The summary figure for the ACEP and organ-specific grading scale are found in Figure 1.

This study is a quantitative prospective analysis comparing the current ACEP recommended QA method to an organ-specific method. We found that there was an increased interobserver agreement between reviewers and increased variability of the scores given to the studies when using the organ-specific method compared to the ACEP method. This suggests that added detail in

Presented at ACEP Research Forum, October 2021.

Supervising Editor: Dr. Daniel Theodoro.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2022 The Authors. *Academic Emergency Medicine* published by Wiley Periodicals LLC on behalf of Society for Academic Emergency Medicine.

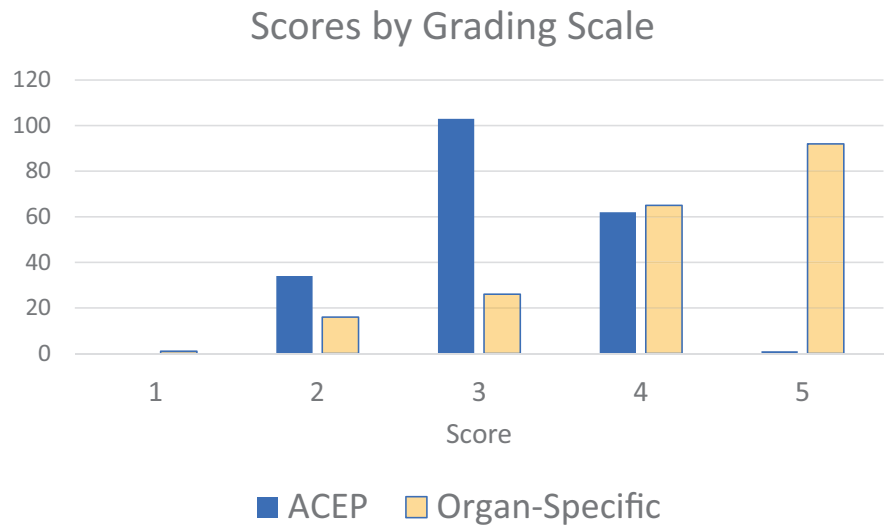


FIGURE 1 Scores given to studies by grading scale

group	min	Q1	median	Q3	max	mean	sd	n	missing
Organ-Specific	1	4	4	5	5	4.155	0.9672917	200	0
ACEP	2	3	3	4	5	3.150	0.6927478	200	0

the guidance to the reviewers led to increased use of the spectrum of the scores. This is particularly important for trainees; they are often the performers of the ultrasound and, as such, the recipient of the QA reports. More direct feedback may expand their knowledge base to be implemented in future studies. For example, the top score on the ACEP grading scale states in a nonspecific manner, "all structures imaged with excellent image quality" versus the cardiac organ-specific grading scale states "greater than 90% of endocardium and valve motion seen." This allows the learner to understand why the top score was given and the structures needed to meet this requirement. The opposite is also true, the organ-specific grading scale presents a detailed response and as such suggests direction for improvement on subsequent studies. The use of a standardized focused scoring system may also improve the quality of the images themselves. Salomon[4] demonstrated that the implementation of an organ specific gynecological emergency ultrasound may improve the quality level of ultrasound examinations performed.

Our data highlight the importance of QA methodology that promotes objectivity demonstrated by high interobserver agreement. POCUS examinations in the ED continue to rise, and as they do there is a need for a standardized method to perform QA that has been thoroughly validated prior to implementation. This single-center study is limited by the images being obtained and reviewed at an academic ED by ultrasound-trained faculty limiting the generalizability.

This investigation demonstrates an organ-specific quality assurance of focused cardiac ultrasound procedures performed in the ED may give the providers performing the studies a wider range of feedback with increased inter-rater agreement among those doing the reviews. Given these findings, there may be a benefit to moving toward an organ-specific QA scale for POCUS studies obtained in the ED.

KEYWORDS

cardiac point of care ultrasound, cardiac ultrasound, emergency department, quality assessment

ACKNOWLEDGMENT

The authors thank Chen Wang and Martin Lavallee with the VCU Biostatistical Consulting Laboratory, which is partially supported by award ULT1R002649 from the National Institutes of Health's National Center for Advancing Translational Science, Department of Biostatistics, VCU School of Medicine.

Stephanie DeMasi MD[1]
Lindsay A. Taylor MD[1]
Adam Weltler MD[1]
John C. Wiggins MD[1]
Jacob Wayman MD[1]
Chen Wang MS[2]
David P. Evans MD[1]
Jessica R. Balderston MD[1]

[1]*Department of Emergency Medicine, Virginia Commonwealth University, Richmond, Virginia, USA*
[2]*Department of Biostatistics, Virginia Commonwealth University, Richmond, Virginia, USA*

Correspondence
Jessica R. Balderston, MD, Department of Emergency Medicine, Virginia Commonwealth University, Richmond, VA 23298, USA.
Email: jessica.balderston@vcuhealth.org

ORCID

Stephanie DeMasi https://orcid.org/0000-0002-4149-8955
Lindsay A. Taylor https://orcid.org/0000-0001-7397-8665
Jessica R. Balderston https://orcid.org/0000-0002-5315-3604

REFERENCES

1. Pattock AM, Kim MM, Kersey CB, et al. Cardiac point-of-care ultrasound publication trends. *Echocardiography*. 2022;39(2):240-247.
2. Liu RB, Blaivas M, Moore C, Sivitz AB, Flannigan M, Tirado A., et. al. *Emergency Ultrasound Standard Reporting Guidelines*. ACEP Emergency Ultrasound Guidelines. June 2018.
3. Amini R, Adhikari S, Fiorello A. Ultrasound competency assessment in emergency medicine residency programs. *Acad Emerg Med*. 2014;21:799-801.
4. Kimura BJ, Gilcrease GW 3rd, Showalter BK, Phan JN, Wolfson T. Diagnostic performance of a pocket-sized ultrasound device for quick-look cardiac imaging. *Am J Emerg Med*. 2012;30(1):32-36.
5. Salomon LJ. A Score-based Method to Improve the Quality of Emergency Gynaecological Ultrasound Examinations. *Eur J Obstet Gynecol Reprod Biol*. 2009;143(2):116-120.

Received: 14 June 2022 | Revised: 17 July 2022 | Accepted: 26 July 2022

DOI: 10.1111/acem.14575

SPECIAL CONTRIBUTION

Global emergency medicine: A scoping review of the literature from 2021

Sean M. Kivlehan MD, MPH[1,2] | Braden J. Hexom MD[3] | Joseph Bonney MBChB, MPH, MScDM, MGCPS[4,5] | Amanda Collier MD, DTMH[6] | Benjamin D. Nicholson MD[7] | Nana Serwaa A. Quao MD, MGCS[8] | Megan M. Rybarczyk MD, MPH[9] | Anand Selvam MD, MSc, DTM&H[10] | Chris A. Rees MD, MPH[11] | Charlotte M. Roy MD, MPH[12] | Nidhi Bhaskar BA[13] | Torben K. Becker MD, PhD[14] | on behalf of the Global Emergency Medicine Literature Review (GEMLR) Group

[1]Department of Emergency Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA

[2]Harvard Humanitarian Initiative, Cambridge, Massachusetts, USA

[3]Department of Emergency Medicine, Rush University Medical Center, Chicago, Illinois, USA

[4]Department of Emergency Medicine, Komfo Anokye Teaching Hospital, Kumasi, Ghana

[5]Global Health and Infectious Disease Research Group, Kumasi Center for Collaborative Research in Tropical Medicine, Kumasi, Ghana

[6]Department of Emergency Medicine, Queen's University, Kingston, Ontario, Canada

[7]Department of Emergency Medicine, Virginia Commonwealth University, Richmond, Virginia, USA

[8]Department of Emergency Medicine, Accident and Emergency Centre, Korle Bu Teaching Hospital, Accra, Ghana

[9]Department of Emergency Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA

[10]Department of Emergency Medicine, Yale University, New Haven, Connecticut, USA

[11]Division of Pediatric Emergency Medicine, Emory University School of Medicine, Atlanta, Georgia, USA

[12]Department of Emergency Medicine, Loma Linda University, Loma Linda, California, USA

[13]Brown University, Providence, Rhode Island, USA

[14]Department of Emergency Medicine, University of Florida, Gainesville, Florida, USA

Correspondence
Sean M. Kivlehan, Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA.
Email: smkivlehan@bwh.harvard.edu

Abstract

Objective: The objective was to identify the most important and impactful peer-reviewed global emergency medicine (GEM) articles published in 2021. The top articles are summarized in brief narratives and accompanied by a comprehensive list of all identified articles that address the topic during the year to serve as a reference for clinicians, researchers, and policy makers.

Methods: A systematic PubMed search was carried out to identify all GEM articles published in 2021. Title and abstract screening was performed by trained reviewers and editors to identify articles in one of three categories based on predefined criteria: disaster and humanitarian response (DHR), emergency care in resource-limited settings (ECRLS), and emergency medicine development (EMD). Included articles were each scored by two reviewers using established rubrics for original (OR) and review

Global Emergency Medicine Literature Review (GEMLR) Group members are listed in Appendix A.

Supervising Editor: Dr. Michael S. Runyon

Presented at the Society for Academic Emergency Medicine Annual Meeting, New Orleans, LA, May 2022.

 © 2022 Society for Academic Emergency Medicine.

(RE) articles. The top 5% of articles overall and the top 5% of articles from each category (DHR, ECRLS, EMD, OR, and RE) were included for narrative summary.
Results: The 2021 search identified 44,839 articles, of which 444 articles screened in for scoring, 25% and 22% increases from 2020, respectively. After removal of duplicates, 23 articles were included for narrative summary. ECRLS constituted the largest category ($n = 16$, 70%), followed by EMD ($n = 4$, 17%) and DHR ($n = 3$, 13%). The majority of top articles were OR ($n = 14$, 61%) compared to RE ($n = 9$, 39%).
Conclusions: The GEM peer-reviewed literature continued to grow at a fast rate in 2021, reflecting the continued expansion and maturation of this subspecialty of emergency medicine. Few high-quality articles focused on DHR and EMD, suggesting a need for further efforts in those fields. Future efforts should focus on improving the diversity of GEM research and equitable representation.

INTRODUCTION

Global emergency medicine (GEM) is a rapidly growing academic subspecialty of emergency medicine focused on the strengthening of emergency care around the world.[1] This work takes many forms, such as direct patient care in broadly diverse contexts, humanitarian response, and system development. The maturation of the subspecialty of GEM has been accompanied by an ever-increasing quantity of academic publications of varying quality. The Global Emergency Medicine Literature Review (GEMLR) was created in 2005 to assist clinicians, researchers, and policy makers in identifying key current literature most likely to influence the practice of GEM through its annual review.[2–17]

The GEMLR Group has grown along with GEM to serve as a mentoring and networking hub for GEM practitioners around the world.[18] The team changes and grows annually, and term-limits force turnover in leadership positions to allow for an increasingly diverse team. In addition to the annual review, the GEMLR group produces an annual systematic review on an understudied topic and maintains a website for data and literature dissemination.[19–22]

Since the inception of the annual review, the number of articles screened for inclusion has grown from under one hundred to tens of thousands. To keep up with this growth, structured technical, communications, and planning committees meet twice annually to revisit and revise the approach to identify the top articles each year. This includes defining GEM, to which the GEMLR group takes an inclusive approach leading to frequent modifications to the search strategy. This iterative approach to article identification and scoring rubrics has been used by other groups.[23,24] This scoping review of the GEM literature seeks to provide an objective and accessible list of the top articles published in 2021 along with an interpretation of current GEM publication trends.

METHODS

The GEMLR procedures used in the 2021 review build upon the approach to identifying top articles established in previous year's reviews. Modifications to the strategy occur most years as a result of committee work, which this year included elimination of the gray literature search and minor scoring rubric changes. The GEMLR methods are outlined in detail in the annually updated procedure manual (Data Supplement S1.) The Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) checklist was used for this review (Data Supplement S2).

The GEMLR team is selected each year through a competitive process after an annual call for applications. Term limits for leadership positions were recently introduced to promote turnover and enhance diversity. This year, the team consisted of 42 people from 10 countries. Team roles included an editor-in-chief; managing editor; assistant managing editor; technical editor; six editors; two assistant editors; 24 reviewers; and six assistant reviewers based in Australia, Brazil, Canada, England, Ethiopia, The Gambia, Ghana, Scotland, South Africa, and the United States (Appendix A). The GEMLR is unfunded and all team members are volunteers. Conflicts of interest can arise when a team member was an author on a publication identified in the search and are managed through a process outlined in the procedure manual.

The search, screening, and scoring process occurs in two blocks to facilitate the quantity of work required. The first search covered January 1 to August 31 and the second search covered September 1 to December 31, 2021. The search string was developed with the support of a librarian to optimize efficiency. Minor overlap with articles published in 2020 and 2022 occurs due to journal publication and indexing variation. The search string is designed to identify articles in English, Spanish, and French that address emergency care delivery in lower resourced and humanitarian settings. All articles published in *African Journal of Emergency Medicine*, *Bulletin of the World Health Organization*, *The Lancet*, and *Prehospital and Disaster Medicine* undergo title and abstract screening due to their historic high rate of included articles (hand search). Case reports, commentaries, editorials, and other similar narrative publications were excluded. The full PubMed search string is in Data Supplement S3.

Historically, a small team of reviewers would perform a manual gray literature search of high yield non–peer-reviewed publications and guideline websites. This year the gray literature search was

discontinued after a review of prior years revealed an extremely low yield of high-impact GEM articles.

Articles retrieved in the PubMed search undergo title and abstract screening for relevance to GEM per predefined criteria outlined in the procedures manual. All reviewers perform practice screening and are given feedback to improve consistency. Two team members, a reviewer and an editor, review each title and abstract followed by a third level of review by both the assistant managing and the managing editor.

Screened-in articles are categorized as either original research (OR) or a review article (RE). Original research articles are further categorized as quantitative or qualitative or mixed methods for scoring purposes. All articles are also classified as one of three potential GEM article categories: disaster and humanitarian response (DHR), emergency care in resource-limited settings (ECRLS), or emergency medicine development (EMD). All screened-in articles each undergo independent full-text review and scoring by two reviewers.

Three scoring rubrics are used for each of the three types of articles: OR quantitative, OR qualitative, and RE. The scoring for each ranges from 1 to 20 points, with scores spread across four subcategories: clarity (RE only), ethics (OR only), rigor, importance, and impact. This year, changes were made to the importance category to clarify scoring related to generalizability and affected population. The full templates are available in Data Supplement S4.

The top 5% of all scored articles as well as the top 5% of each article category (OR, RE, DHR, ECRLS, and EMD) are narratively summarized. These summaries are a single page in length and follow a templated structure that includes a single-sentence summary; a brief and objective summary of the article's methods and findings; and an interpretation of the strengths, limitations, and importance to GEM. The summaries are prepared by reviewers and edited by editors. They are meant to serve as an accessible brief summary of the top GEM literature for 2021 (Figure 1).

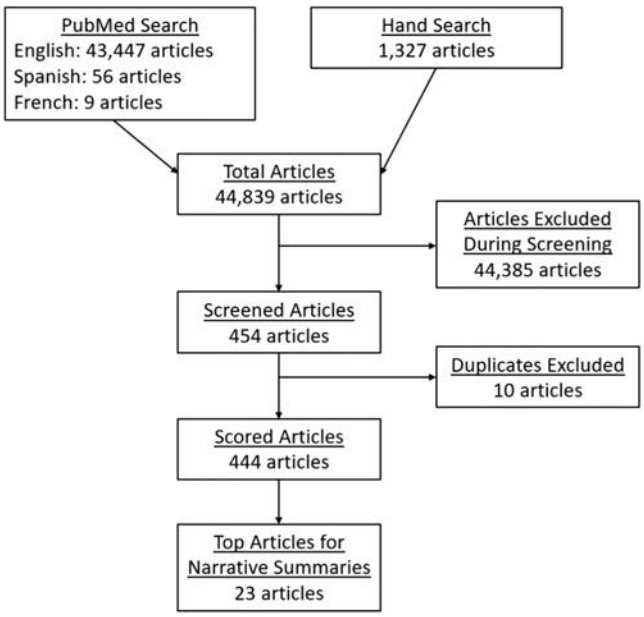


FIGURE 1 Flow diagram of articles selected for review

RESULTS

The PubMed search retrieved 44,839 articles, which included 1327 from the hand search. A total of 444 articles met the inclusion criteria and went on to have a full-text review and scoring (Table 1 and Data Supplement S5). After duplicates were removed, 23 articles were selected for narrative summarization after being identified as a top article by our predefined criteria (Table 2). Of the selected articles, 14 (61%) were OR and nine (39%) were RE. Of the OR articles, 13 (93%) were quantitative and one (7%) was qualitative. Three (13%) were DHR, 16 (70%) were ECRLS, and four (17%) were EMD. Two of the selected top articles were from the hand search journals and two were COVID-19 related. The narrative summaries of each of the 23 top articles are provided in Data Supplement S6.

DISCUSSION

The quantity of GEM research has continued to grow at a fast rate, reflecting the continued growth of the subspecialty. This year's search identified 44,839 articles, a 25% increase from the 35,970 identified in 2020. A total of 444 articles underwent full-text review, a 22% increase from last year's 364. These increases correspond to the 15% increase in articles that underwent a full narrative summary this year. This precipitous growth rate is likely related to the increased recognition of emergency medicine as a specialty around the world, increased funding opportunities for GEM research, and a growing number of journals focusing on global health.[25–28]

Last year's search utilized two distinct search strings to separate COVID-19–related articles from all others; however, this year we did not seek to separately identify articles related to the coronavirus pandemic. While last year's search identified a significant quantity of COVID-19 articles in the search (15%) and among screened-in articles (9.6%), none of the identified top articles were COVID-19 related. This was thought to be related to the relatively lower quality of COVID-19–related articles meeting GEMLR inclusion criteria.[29,30] This year, two COVID-19–related articles were among the top articles, which may reflect an improvement in the quality of COVID-19 research in non–high-resourced settings.

The GEMLR editorial board consisting of all current and many former editors meets twice yearly to evaluate and improve policies and procedures, supported by three different committees that are assigned tasks to deliberate and evaluate in-between board meetings. Changes were implemented for this year's search affecting the search, scoring, and structural domains. No changes were made to search terms; however, COVID-19–specific selectors were removed. Revision of the hand-search journals occurs on a regular basis to both add promising new journals not immediately indexed in PubMed and to remove journals when all relevant articles from the prior 2 years appear in the core PubMed search. However, no changes were made this year. Minor scoring revisions were made to both the importance and impact subsections to better reflect the generalizability and affected populations of a study.

TABLE 1 Summary statistics for article scoring

		Scores				
	Number (%)	Minimum	25th Percentile	Median	75th Percentile	Maximum
Total	444	3.5	12	13.5	15.5	20
Article category						
DHR	87 (19.6)	5.5	11.5	13	15	18
ECRLS	282 (63.5)	6	12	13.5	15.5	20
EMD	75 (16.9)	3.5	12	13.5	15.5	18.5
Type of research article						
OR	365 (82.2)	3.5	12	13.5	15	19.5
RE	79 (17.8)	7	13	15	16.5	20
Type of original research article						
Quantitative study	272 (74.5)	6	12	13.5	15	19.5
Qualitative study	86 (23.6)	3.5	12	13.5	14.5	18.5
Mixed-methods study	7 (1.9)	13.5	13.5	15	16.5	16.5

Abbreviations: DHR, disaster and humanitarian response; ECRLS, emergency care in resource-limited settings; EMD, emergency medicine development; OR, original research; RE, research article.

The parallels between colonization and GEM are increasingly recognized and the GEMLR group has looked inward this year to discuss authorship equity of both included articles and the review itself.[31–33] The GEMLR reviewers, who are listed as contributing authors in Appendix A, are more geographically diverse than the listed authors. GEMLR authorship positions are determined as outlined in the procedures manual (Data Supplement S1), with all authors required to have previously served as a reviewer. Author-level role term limits were tightened this year to accelerate turnover from the increasingly diverse reviewer group. Additionally, GEMLR annual recruitment efforts have broadened to advertise via several emergency medicine groups outside of North America in an effort to promote membership diversity. A review of the GEMLR's included papers over the past several years has clearly demonstrated authorship inequity between high- and low-and-middle-income countries.[34] Changes to scoring rubrics are planned for next year to address local authorship representation, and future committee work will evaluate the role of funding channels in GEM research equity.

The papers reviewed each year by GEMLR are broadly heterogenous and not amenable to a formal systematic review or meta-analysis. The scoping review format allows for the annual identification of a diverse set of original research and reviews, and the application of a rigorously developed scoring system provides confidence in article selection. The topics and importance of this year's top articles are briefly discussed below. The full list of scored articles is presented in Data Supplement S5 as a searchable and sortable spreadsheet and the full narrative summaries can be found in Data Supplement S6.

Disaster and humanitarian response

Cavaliere and colleagues[35] reviewed publicly available data from the Global Terrorism Database and identified 901 incidents in 74 countries targeting health care facilities from 1970 to 2018. This work brings attention to this persistent and widespread threat to health care workers and patients around the world. In a systematic review of the quality of health systems in humanitarian settings covering 2010 to 2019, Jordan and colleagues[36] found that a disproportionate emphasis is placed on system foundations rather than competency of care, population health, or economic or societal benefits. Emphasizing the importance of quality of care, they report the common finding that absence of a health care provider led to better outcomes than the presence of a "noncompetent" one. The third and final DHR article included is a systematic review of health system responses to disasters during the COVID-19 pandemic by Sohrabizadeh and colleagues.[37] Their findings demonstrate the challenge of balancing individual and public health while responding to acute events such as earthquakes, floods, and wildfires in addition to highlighting preparedness and response strategies.

Emergency care in resource-limited settings

As in prior years, ECRLS constituted the majority of top articles, many of which were pediatric oriented. In a randomized controlled trial of children with severe pneumonia in Africa, Maitland and colleagues[38] demonstrated a mortality reduction with high-flow nasal therapy (HFNT) compared to low-flow oxygen and permissive hypoxemia. While this is the largest multicenter controlled trial of oxygen therapy conducted in a pediatric population, it was terminated early due to ethical concerns around withholding oxygen in one arm. Notably, their data showed that HFNT with or without oxygen reduced mortality. Sessions and colleagues[39] evaluated continuous positive airway pressure (CPAP) for nonneonatal pediatric respiratory distress in low- and middle-income countries (LMICs). A meta-analysis was performed on three of the included studies which

TABLE 2 Top-scoring GEM literature review 2021 articles

Category	First author	Title	Journal	OR or RE	Summary
DHR	Cavaliere[35]	Terrorist attacks against healthcare facilities: a review	*Health Secur*	RE	Healthcare facilities are key infrastructure that are increasingly targets for terrorist attacks around the world.
	Jordan[36]	Quality in crisis: a systematic review of the quality of health systems in humanitarian settings	*Confl Health*	RE	Health system quality in humanitarian settings is currently evaluated in terms of workforce and tools, but further evaluations need to address competency and population appropriateness.
	Sohrabizadeh[37]	A systematic review of health sector responses to the coincidence of disasters and COVID-19	*BMC Public Health*	RE	Disaster responses during the COVID-19 pandemic highlighted the challenge of balancing individual and public health while providing robust aid to a community. The authors emphasize the need for unique multidisciplinary preparation, vigorous coordination, and robust logistics in order to respond to disasters during a pandemic.
ECRLS	Ahmed[44]	Effect of 3 days of oral azithromycin on young children with acute diarrhea in low resource settings	*JAMA Netw Open*	OR (QUANT)	This randomized clinical trial found no survival benefit associated with the addition of a 3-day course of azithromycin to standard case management of acute nonbloody watery diarrhea in children.
	Al Hussain[40]	Validity of pneumonia severity assessment scores in Africa and South Asia: A systematic review and meta-analysis	*Healthcare (Basel)*	RE	The CURB-65 and CRB-65 pneumonia severity scores perform well and are valid for use in LMICs.
	Ayede[42]	Management of possible serious bacterial infection in young infants where referral is not possible in the context of existing health system structure in Ibadan, south-west Nigeria	*PLoS One*	OR (QUANT)	Outpatient management of possible serious bacterial infection in young infants can successfully reduce neonatal morbidity and mortality within existing health system infrastructure, but requires focused investment to ensure treatment availability, training and motivation of health workers, and community integration and involvement.
	Bandyopadhyay[46]	Traumatic brain injury-related pediatric mortality and morbidity in low- and middle-income countries: A systematic review	*World Neurosurg*	RE	Road traffic accidents account for the majority of pediatric traumatic brain injury in LMICs with young males at greatest risk.
	Barry[41]	Derivation and internal validation of a score to predict dehydration severity in patients over 5 years with acute diarrhea	*Am J Trop Med Hyg*	OR (QUANT)	The Novel, Innovative Research for Understanding Dehydration in Adults and Kids (NIRUDAK) score was developed in Bangladesh to identify dehydration severity among patients aged ≥5 years with acute diarrheal illness.
	Beza[50]	Acute coronary syndrome treatment delay in low to middle-income countries: A systematic review	*Int J Cardiol Heart Vasc*	RE	Multiple individual and systemic risk factors for delay in treatment of acute coronary syndrome exist within LMICs.
	Charle-Cuéllar[45]	Effectiveness and coverage of treatment for severe acute malnutrition delivered by community health workers in the Guidimakha Region, Mauritania	*Children (Basel)*	OR (QUANT)	Community health workers may help lessen the financial and geographic barriers to severe acute malnutrition treatment, compared to traditional health facility-based care.
	Leclerc[51]	A simplified fluid resuscitation formula for burns in mass casualty scenarios: Analysis of the consensus recommendation from the WHO Emergency Medical Teams Technical Working Group on Burns	*Burns*	OR (QUANT)	Initial fluid rate of 100 mL/kg/24 h either orally or intravenously may be useful for initial resuscitation, particularly adult patients, in disaster scenarios.
	Maitland[38]	Randomized controlled trial of oxygen therapy and high-fow nasal therapy in African children with pneumonia	*Intensive Care Med*	OR (QUANT)	Use of high-flow nasal therapy may be an effective and feasible tool in low-resource settings for reduction of mortality in children with severe pneumonia.
	Reis[53]	Effect of early treatment with hydroxychloroquine or lopinavir and ritonavir on risk of hospitalization amount patients with COVID-19: The TOGETHER randomized clinical trial	*JAMA Netw Open*	OR (QUANT)	This randomized control trial compared hydroxychloroquine and lopinavir-ritonavir against placebo for the treatment of COVID-19 in high-risk outpatients in Brazil and found that neither drug significantly reduced 90-day hospitalization or death.

TABLE 2 (Continued)

Category	First author	Title	Journal	OR or RE	Summary
	Roberts[52]	A high-dose 24-h tranexamic acid infusion for the treatment of significant gastrointestinal bleeding: HALT-IT RCT	*Health Technol Assess*	OR (QUANT)	This multicenter, randomized, placebo-controlled, double-blinded trial demonstrated that tranexamic acid does not reduce deaths or improve clinical outcomes in patients with significant gastrointestinal bleeding.
	Rondy[49]	Field evaluation of the performance of two rapid diagnostic tests for meningitis in Niger and Burkina Faso	*Microorganisms*	OR (QUANT)	Rapid diagnostic tests offer public health authorities an important tool in detecting meningococcal meningitis epidemics and surveying its most common variants.
	Salahuddin[48]	A shorter post-exposure prophylaxis regimen for rabies, Pakistan	*Bull World Health Organ*	OR (QUANT)	This paper demonstrates that an abridged, 1-week rabies postexposure prophylaxis regimen with low immunoglobulin use is noninferior to a standard 4-week regimen with high immunoglobulin use in an urban Pakistani hospital, with improved cost-effectiveness and patient adherence.
	Sessions[39]	Continuous positive airway pressure for children in resource-limited settings, effect on mortality and adverse events: systematic review and meta-analysis	*Arch Dis Child*	RE	This systematic review and meta-analysis found inconclusive and low certainty evidence for any mortality benefit from CPAP use for pediatric respiratory distress in LMICs. Adverse events were similarly inconclusive and of low certainty.
	Thomson[43]	Effects of antibiotic resistance, drug target attainment, bacterial pathogenicity and virulence, and antibiotic access and affordability on outcomes in neonatal sepsis: an international microbiology and drug evaluation prospective substudy (BARNARDS)	*Lancet Infect Dis*	OR (QUANT)	Based on this study, the combination of ampicillin-gentamicin for neonatal sepsis (as per WHO recommendation) might need to be reconsidered in LMICs because of, among other factors, the high prevalence of resistance. Ceftazidime-amikacin is a potential replacement.
	Whitaker[47]	Assessing trauma care systems in low-income and middle-income countries: a systematic review and evidence synthesis mapping the Three Delays framework to injury health system assessments	*BMJ Glob Health*	RE	This systematic review and evidence synthesis of the assessment of LMIC trauma systems applies the three-delays framework, more commonly used to assess maternal health care, alongside rapid assessment principles to identify gaps in the current literature and where future priorities should focus.
EMD	Hands[55]	Emergency Triage Assessment and Treatment Plus (ETAT+): adapting training to strengthen quality improvement and task-sharing in emergency pediatric care in Sierra Leone	*J Glob Health*	OR (QUANT)	Implementation of ETAT+ using a mentored pragmatic approach decreased pediatric mortality and demonstrated improved quality of emergency care.
	Ong[56]	International multi-center real world implementation trial to increase out-of-hospital cardiac arrest survival with a dispatcher-assisted cardio-pulmonary resuscitation package (Pan-Asian resuscitation outcomes study phase 2)	*Resuscitation*	OR (QUANT)	The article evaluates an implementation of dispatcher-assisted bystander CPR for management of OHCA across multiple countries in the Asia-Pacific region and suggests that its implementation is both feasible and beneficial in the management of OHCA.
	Orkin[54]	Emergency care with lay responders in underserved populations: a systematic review	*Bull World Health Organ*	RE	Emergency care by trained layperson responders in underserved populations is understudied, but a systematic review of the limited data available suggests potential for benefit in health outcomes.
	Waydhas[57]	Survey on worldwide trauma team activation	*Eur J Trauma Emerg Surg*	OR (QUAL)	This paper supports the suggested criteria used in the activation of trauma teams regardless of a country's income level.

Abbreviations: DHR, disaster and humanitarian response; ECRLS, emergency care in resource-limited settings; EMD, emergency medicine development; GEM, global emergency review; LMICs, low- and middle-income countries; OHCA, out-of-hospital cardiac arrest; OR, original research; QUAL, qualitative original research; QUANT, quantitative original research; RE, review article.

found inconclusive and low-certainty evidence for any mortality benefit from CPAP. The authors' findings cast doubt on the efficacy of CPAP for pediatric respiratory distress in LMICs, and further point out that while an oxygen concentrator can only provide CPAP for one child at a time, it could instead deliver low-flow oxygen to five children simultaneously.

Al Hussain and colleagues[40] performed a systematic review demonstrating ability of two pneumonia severity scores, CURB-65 and CRB-65, to predict mortality among adults in LMICs. Notably, the CURB-65 tool was superior but the CRB-65 tool was still useful in settings where blood chemistry testing is unavailable. As most widely used risk prediction scores have not been validated in LIMCs, this work provides practitioners in LMICs with an evidenced-based tool for triage. In another article focused on severity scoring, Barry and colleagues[41] performed a secondary analysis of data from the Novel, Innovative Research for Understanding Dehydration in Adults and Kids (NIRUDAK) study, which evaluated acute diarrhea management in Bangladesh, to derive and internally validate a nurse-led 13-point scoring tool to predict the severity of diarrhea in patients over 5 years of age. The discriminatory values of the tool were acceptable and should now be evaluated externally.

Ayede and colleagues[42] implemented the World Health Organization (WHO) outpatient treatment guidelines for neonates 0–59 days old with possible serious bacterial infection (PSBI) when inpatient referral is not possible at eight outpatient primary health care centers in southwest Nigeria. Over a 16-month period, trained community resource persons identified 1214 newborns with PSBI of which 91.5% declined hospital referral and were treated at home. All patients with pneumonia were successfully treated at home and 98% of those with clinical severe infection were also cured with home antibiotic treatment. This important study demonstrates that novel outpatient treatment guidelines can be successfully implemented within existing health system infrastructure resulting in reduced infant morbidity and mortality when social barriers to care are considered and addressed in a contextually relevant and practically applicable manner. The use and efficacy of empiric antibiotic therapy for neonatal sepsis in seven countries was evaluated by Thomson and colleagues[43] in an observational study. Although the WHO first-line recommendation of ampicillin and gentamicin was among the most commonly used antibiotics, 379 (97.2%) and 274 (70.3%) of 390 Gram-negative isolates were resistant to ampicillin and gentamicin, respectively. These findings suggest the need to revisit the current guidelines in LMICs.

In a randomized controlled trial of treatment for pediatric acute nonbloody watery diarrhea in seven countries, Ahmed and colleagues[44] found no survival benefit to the addition of a 3-day course of azithromycin to standard case management. These findings add valuable evidence to support adherence to current WHO case management guidelines including zinc, rehydration, and nutritional counseling for the treatment of pediatric acute watery diarrhea in low- and middle-income settings. Focusing on malnutrition in children, a study performed by Charle-Cuéllar and colleagues[45] in Mauritania found that children with severe acute malnutrition (SAM) treated by community health workers (CHWs) were found to have similar SAM cure and follow-up failure rates as those treated by traditional outpatient health facilities in a nonrandomized controlled study. This suggests that CHWs can help bridge the gap between access and delivery of treatment for SAM, particularly in resource-limited settings.

Road traffic accidents were found to the be the most common cause of pediatric traumatic brain injury (pTBI) in a systematic review of the causes of pTBI performed by Bandyopadhyay and colleagues.[46] Their review covered 32 LMICs and demonstrated that pTBI patients in LMICs are younger than those in high-income countries and more often pedestrians than vehicle occupants. The authors also highlighted a gap in pTBI knowledge as data were only available from 26% of 135 LMICs. In an evaluation of trauma systems, Whitaker and colleagues[47] used the three-delays framework, previously used to characterize barriers to maternal health care, to perform a systematic review of the literature on LMIC trauma care systems. Delay 1 considers barriers to seeking care, Delay 2 looks at barriers to reaching care, and Delay 3 concentrates on barriers to receiving care. The authors found that Delay 3 was the most studied, with facility assessments being the most common tool, and concluded that in addition to whole health system injury research, increased research on Delays 1 and 2 are needed.

Rabies and meningococcal meningitis remain major public health issues in the world. Salahuddin and colleagues[48] showed that a shortened 1-week rabies postexposure prophylaxis (PEP) schedule was equivalent to the standard 4-week schedule in Pakistan. Despite an increased number of patients requiring PEP during the study period, there was less use of rabies vaccine and immunoglobulin vials, cost savings, increased patient adherence, and no reported deaths. Rondy and colleagues[49] compared the performance characteristics of two rapid diagnostic tests (RDTs) compared with traditional PCR-based tests to detect meningitis caused by *Neisseria meningitidis* and *Streptococcus pneumoniae* in Burkina Faso and Niger. They found that the RDTs met WHO acceptable standards, providing evidence to support this diagnostic tool.

Risk factors for delays in the diagnosis and treatment of acute coronary syndrome (ACS) in LMICs were identified in a systematic review by Beza and colleagues,[50] identifying many targets for future work in this area. Individual risk factors included lower socioeconomic status; female gender; older age; and comorbid conditions of diabetes, hypertension, and dyslipidemia, while systemic risk factors included rural location, lack of transportation, ambulance availability, and lack of education surrounding ACS for both patients and clinical staff. Turning attention to the fluid resuscitation of burns, Leclerc and colleagues[51] evaluated the efficiency of simplified formulas that could be helpful in mass casualty and disaster scenarios when resources may be scarce. In a mathematical analysis, Leclerc and colleagues compared the WHO Technical Working Group for Burns recommendation of 100 mL/kg over 24 h either orally or intravenously for burns larger than 20% total body surface area to existing guidelines. They concluded that although it underestimates fluid volumes in small children and in people with large burns, this simpler formula should be appropriate in mass burn casualty situations.

The final two ECRLS articles are negative studies regarding treatment for gastrointestinal (GI) bleeding and COVID-19. In a multicenter randomized controlled trial of the effectiveness of tranexamic acid (TXA) or placebo for patients acute GI bleeding, Roberts and colleagues[52] showed that TXA did not reduce mortality and also that it increased the risk of venous thromboembolic events and seizures. This large study, performed in 15 countries spanning income levels, has produced strong evidence that TXA should not be used in patients with acute GI bleeding. A large randomized controlled trial conducted in Brazil on COVID-19 treatment early in the pandemic evaluated two medications against placebo, hydroxychloroquine and lopinavir-ritonavir. Reis and colleagues[53] found that neither medication reduced 90-day hospitalization or death, impacting global COVID-19 treatment recommendations.

Emergency medicine development

The first of the four included EMD articles is a systematic review assessing the effects of layperson emergency care training in low-resource areas by Orkin and colleagues.[54] They included 34 papers and concluded that while layperson training for acute interventions is widespread and logical, there is limited evidence for a health benefit in this context in the existing literature. Hands and colleagues[55] demonstrated a decrease in pediatric mortality at health care facilities in Sierra Leone after implementation of Emergency Triage Assessment and Treatment Plus (ETAT+). The authors used a pragmatic, mentored approach to ETAT+ implementation, which is critical to sustained use and impact. This paper adds to the growing body of evidence that improving the process of emergency care leads to improved patient outcomes, including reduced mortality.

Ong and colleagues[56] performed a large-scale controlled implementation of dispatcher-assisted bystander CPR for out-of-hospital cardiac arrest (OHCA) in the Asia-Pacific region. While they did demonstrate a significantly higher increase in bystander CPR rates and favorable neurologic outcome in the intervention arm, a large majority of data was from high-income counties with less than 10% of recorded OHCA events having occurred in an LMIC, reducing the generalizability of the findings. The final EMD article is a cross-sectional observational study of physician level of agreement on conditions requiring trauma team activation regardless of country income level by Waydhas and colleagues.[57] Through purposive sampling and using Delphi methodology, they found that criteria related to abnormal vital signs and need for certain emergency invasive procedures such as airway management had the highest levels of agreement across country income levels while acknowledging that some criteria may not be readily available in all settings.

LIMITATIONS

There are limitations to the GEMLR's approach to identifying top articles each year. The search terms are reviewed and revised each year to be as inclusive as possible; however, it is possible that some relevant articles are missed due to the broad and heterogenous nature of GEM. Extensive efforts are taken each year to maintain and update objective and complete scoring rubrics to identify top articles; however, some subjectivity in assigning points remains. Each year the GEMLR technical committee revisits these processes based on input from reviewers; details of this can be found in Data Supplement S1.

CONCLUSIONS

The continuously expanding quantity and quality of global emergency medicine literature is an indicator of the maturation of this subspecialty. The Global Emergency Medicine Literature Review has continued to evolve to keep up with this growth by diversifying its team and improving its search, screening, and scoring procedures. Emergency care in resource-limited settings continues to be the largest area of global emergency medicine research, and the lower quantity of disaster and humanitarian response and emergency medicine development studies suggests a need for increased research in those areas. Future Global Emergency Medicine Literature Review work should align with efforts to diversify and decolonize global emergency medicine while serving as a hub for academic mentorship and networking.

AUTHOR CONTRIBUTIONS

Study concept and design: Sean M. Kivlehan, Braden J. Hexom, Torben K. Becker. Acquisition of data: Nidhi Bhaskar. Analysis and interpretation of data: Sean M. Kivlehan, Braden J. Hexom, Joseph Bonney, Amanda Collier, Benjamin D. Nicholson, Nana Serwaa A. Quao, Megan M. Rybarczyk, Anand Selvam, Chris A. Rees, Charlotte M. Roy. Drafting of the manuscript: Sean M. Kivlehan. Critical revision of the manuscript for important intellectual content: Sean M. Kivlehan, Braden J. Hexom, Joseph Bonney, Amanda Collier, Benjamin D. Nicholson, Nana Serwaa A. Quao, Megan M. Rybarczyk, Anand Selvam, Chris A. Rees, Charlotte M. Roy, Nidhi Bhaskar, Torben K. Becker. Statistical analyses: Sean M. Kivlehan. Acquisition of funding: none.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

ORCID

Sean M. Kivlehan https://orcid.org/0000-0002-4857-4316
Megan M. Rybarczyk https://orcid.org/0000-0003-1936-8473
Torben K. Becker https://orcid.org/0000-0002-5287-8056

REFERENCES

1. Patiño AM, Chen J, DeVos EL, et al. Emergency medicine around the world: analysis of the 2019 American College of Emergency Physicians international ambassador country reports. *J Am Coll Emerg Physicians Open*. 2022;3(2):e12681.

2. Levine AC, Gadiraju S, Goel A, Johar S, King R, Arnold K. International emergency medicine: a review of the literature. *Acad Emerg Med*. 2007;14:182-183.
3. Levine AC, Goel A, Keay CR, et al. International emergency medicine: a review of the literature from 2006. *Acad Emerg Med*. 2007;14:1190-1193.
4. Levine AC, Becker J, Lippert S, Rosborough S, Arnold K. Emergency medicine resident association international emergency medicine literature review group. International emergency medicine: a review of the literature from 2007. *Acad Emerg Med*. 2008;15:860-865.
5. Lippert S, Levine AC, Becker J, et al. International emergency medicine: a review of the literature from 2008. *Acad Emerg Med*. 2009;16:1335-1340.
6. Foran M, Levine A, Lippert S, et al. International emergency medicine: a review of the literature from 2009. *Acad Emerg Med*. 2011;18:86-92.
7. Aschkenasy M, Arnold K, Foran M, et al. International emergency medicine: a review of the literature from 2010. *Acad Emerg Med*. 2011;18:872-879.
8. Schroeder ED, Jacquet G, Becker TK, et al. Global emergency medicine: a review of the literature from 2011. *Acad Emerg Med*. 2012;19:1196-1203.
9. Jacquet GA, Foran M, Bartels S, et al. Global emergency medicine: a review of the literature from 2012. *Acad Emerg Med*. 2013;20:835-843.
10. Becker TK, Jacquet GA, Marsh R, et al. Global emergency medicine: a review of the literature from 2013. *Acad Emerg Med*. 2014;21:810-817.
11. Becker TK, Bartels S, Hansoti B, et al. Global emergency medicine: a review of the literature from 2014. *Acad Emerg Med*. 2015;22:976-984.
12. Becker TK, Hansoti B, Bartels S, et al. Global emergency medicine: a review of the literature from 2015. *Acad Emerg Med*. 2016;23:1183-1191.
13. Becker TK, Hansoti B, Bartels S, et al. Global emergency medicine: a review of the literature from 2016. *Acad Emerg Med*. 2017;24:1150-1160.
14. Becker TK, Trehan I, Hayward AS, et al. Global emergency medicine: a review of the literature from 2017. *Acad Emerg Med*. 2018;25:1287-1298.
15. Trehan I, Osei-Ampofo M, Balhara KS, et al. Global emergency medicine: a review of the literature from 2018. *Acad Emerg Med*. 2019;26:1186-1196.
16. Trehan I, Kivlehan SM, Balhara KS, et al. Global emergency medicine: a review of the literature from 2019. *Acad Emerg Med*. 2021;28:117-128.
17. Trehan I, Kivlehan SM, Balhara KS, et al. Global emergency medicine: a scoping review of the literature from 2020. *Acad Emerg Med*. 2021;28(11):1328-1340.
18. Lee JA. *A History of the Global Emergency Medicine Literature Review*. SAEM Pulse; 2022.
19. Balhara KS, Silvestri DM, Tyler Winders W, et al. Impact of nutrition interventions on pediatric mortality and nutrition outcomes in humanitarian emergencies: a systematic review. *Trop Med Int Health*. 2017;22:1464-1492.
20. Balhara KS, Bustamante ND, Selvam A, et al. Bystander assistance for trauma victims in low- and middle-income countries: a systematic review of prevalence and training interventions. *Prehosp Emerg Care*. 2019;23:389-410.
21. Winders WT, Bustamante ND, Garbern SC, et al. Establishing the effectiveness of interventions provided to first responders to prevent and/or treat mental health effects of response to a disaster: a systematic review. *Disaster Med Public Health Prep*. 2021;15:115-126.
22. Winders WT, Garbern SC, Bills CB, et al. The effects of mobile health on emergency care in low- and middle-income countries: a systematic review and narrative synthesis. *J Glob Health*. 2021;11:04023.
23. Sarin RR, Hick JL, Livinski AA, et al. Disaster medicine: a comprehensive review of the literature from 2016. *Disaster Med Public Health Prep*. 2019;13(5–6):946-957.
24. Rybarczyk MM, Ludmer N, Broccoli MC, et al. Emergency medicine training programs in low- and middle-income countries: a systematic review. *Ann Glob Health*. 2020;86(1):60. doi:10.5334/aogh.2681 PMID: 32587810; PMCID: PMC7304456.
25. Razzak J, Beecroft B, Brown J, Hargarten S, Anand N. Emergency care research as a global health priority: key scientific opportunities and challenges. *BMJ Glob Health*. 2019;4(Suppl 6):e001486.
26. Hansoti B, Aluisio AR, Barry MA, et al. Global emergency medicine think tank clinical research working group. Global Health and emergency care: defining clinical research priorities. *Acad Emerg Med*. 2017;24(6):742-753.
27. Bruijns SR. Looking back over a decade with the African journal of emergency medicine. *Afr J Emerg Med*. 2019;9(4):163-164.
28. Hansoti B, Levine A, Ganti L, et al. Funding global emergency medicine research-from seed grants to NIH support. *Int J Emerg Med*. 2016;9(1):27.
29. Zdravkovic M, Berger-Estilita J, Zdravkovic B, Berger D. Scientific quality of COVID-19 and SARS CoV-2 publications in the highest impact medical journals during the early phase of the pandemic: a case control study. *PLoS One*. 2020;15(11):e0241826.
30. Jung RG, Di Santo P, Clifford C, et al. Methodological quality of COVID-19 clinical research. *Nat Commun*. 2021;12:943.
31. Kelaher M, Ng L, Knight K, Rahadi A. Equity in global health research in the new millennium: trends in first-authorship for randomized controlled trials among low- and middle-income country researchers 1990-2013. *Int J Epidemiol*. 2016;45(6):2174-2183.
32. Rees CA, Ali M, Kisenge R, et al. Where there is no local author: a network bibliometric analysis of authorship parasitism among research conducted in sub-Saharan Africa. *BMJ Glob Health*. 2021;6(10):e006982.
33. Rees CA, Lukolyo H, Keating EM, et al. Authorship in paediatric research conducted in low- and middle-income countries: parity or parasitism? *Trop Med Int Health*. 2017;22(11):1362-1370.
34. Garbern SC, Hyuha G, Gonzalez-Marques C, et al. Authorship representation in global emergency medicine: a bibliometric analysis from 2016–2020. *BMJ Glob Health*. 2022;7(6):e009538.
35. Cavaliere GA, Alfalasi R, Jasani GN, Ciottone GR, Lawner BJ. Terrorist attacks against healthcare facilities: a review. *Health Secur*. 2021;19(5):546-550.
36. Jordan K, Lewis TP, Roberts B. Quality in crisis: a systematic review of the quality of health systems in humanitarian settings. *Confl Health*. 2021;15(1):7.
37. Sohrabizadeh S, Yousefian S, Bahramzadeh A, Vaziri MH. A systematic review of health sector responses to the coincidence of disasters and COVID-19. *BMC Public Health*. 2021;21(1):709.
38. Maitland K, Kiguli S, Olupot-Olupot P, et al. Randomised controlled trial of oxygen therapy and high-flow nasal therapy in African children with pneumonia. *Intensive Care Med*. 2021;47(5):566-576.
39. Sessions KL, Smith AG, Holmberg PJ, et al. Continuous positive airway pressure for children in resource-limited settings, effect on mortality and adverse events: systematic review and meta-analysis. *Arch Dis Child*. 2022;107(6):543-552.
40. Al Hussain SK, Kurdi A, Abutheraa N, et al. Validity of pneumonia severity assessment scores in Africa and South Asia: a systematic review and meta-analysis. *Healthcare (Basel)*. 2021;9(9):1202.
41. Barry MA, Qu K, Gainey M, et al. Derivation and internal validation of a score to predict dehydration severity in patients over 5 years with acute diarrhea. *Am J Trop Med Hyg*. 2021;105(5):1368-1375.
42. Ayede AI, Ashubu OO, Fowobaje KR, et al. Management of possible serious bacterial infection in young infants where referral is not possible in the context of existing health system structure in Ibadan, south-West Nigeria. *PLoS One*. 2021;16(3):e0248720.

43. Thomson KM, Dyer C, Liu F, et al. Effects of antibiotic resistance, drug target attainment, bacterial pathogenicity and virulence, and antibiotic access and affordability on outcomes in neonatal sepsis: an international microbiology and drug evaluation prospective substudy (BARNARDS). *Lancet Infect Dis*. 2021;21(12):1677-1688.
44. Antibiotics for Children With Diarrhea (ABCD) Study Group, Ahmed T, Chisti MJ, et al. Effect of 3 days of Oral azithromycin on young children with acute diarrhea in low-resource settings: a randomized clinical trial. *JAMA Netw Open*. 2021;4(12):e2136726.
45. Charle-Cuéllar P, Lopez-Ejeda N, Toukou Souleymane H, et al. Effectiveness and coverage of treatment for severe acute malnutrition delivered by community health Workers in the Guidimakha Region, Mauritania. *Children (Basel)*. 2021;8(12):1132.
46. Bandyopadhyay S, Kawka M, Marks K, et al. Traumatic brain injury-related pediatric mortality and morbidity in low- and middle-income countries: a systematic review. *World Neurosurg*. 2021;153:109-130.e23.
47. Whitaker J, O'Donohoe N, Denning M, et al. Assessing trauma care systems in low-income and middle-income countries: a systematic review and evidence synthesis mapping the three delays framework to injury health system assessments. *BMJ Glob Health*. 2021;6(5):e004324.
48. Salahuddin N, Ansari N, Gohar MA. A shorter post-exposure prophylaxis regimen for rabies. *Pakistan Bull World Health Organ*. 2021;99(7):506-513.
49. Rondy M, Tamboura M, Sidikou F, et al. Field evaluation of the performance of two rapid diagnostic tests for meningitis in Niger and Burkina Faso. *Microorganisms*. 2021;9(4):832.
50. Beza L, Leslie SL, Alemayehu B, Gary R. Acute coronary syndrome treatment delay in low to middle-income countries: a systematic review. *Int J Cardiol Heart Vasc*. 2021;19(35):100823.
51. Leclerc T, Potokar T, Hughes A, et al. A simplified fluid resuscitation formula for burns in mass casualty scenarios: analysis of the consensus recommendation from the WHO emergency medical teams technical working group on burns. *Burns*. 2021;47(8):1730-1738.
52. Roberts I, Shakur-Still H, Afolabi A, et al. A high-dose 24-hour tranexamic acid infusion for the treatment of significant gastrointestinal bleeding: HALT-IT RCT. *Health Technol Assess*. 2021;25(58):1-86.
53. Reis G, Moreira Silva EADS, Medeiros Silva DC, et al. Effect of early treatment with hydroxychloroquine or lopinavir and ritonavir on risk of hospitalization among patients with COVID-19: the TOGETHER randomized clinical trial. *JAMA Netw Open*. 2021;4(4):e216468.
54. Orkin AM, Venugopal J, Curran JD, et al. Emergency care with lay responders in underserved populations: a systematic review. *Bull World Health Organ*. 2021;99(7):514-528H.
55. Hands C, Hands S, Verriotis M, et al. Emergency triage assessment and treatment plus (ETAT+): adapting training to strengthen quality improvement and task-sharing in emergency paediatric care in Sierra Leone. *J Glob Health*. 2021;18(11):04069.
56. Ong MEH, Shin SD, Ko PC, et al. International multi-center real world implementation trial to increase out-of-hospital cardiac arrest survival with a dispatcher-assisted cardio-pulmonary resuscitation package (pan-Asian resuscitation outcomes study phase 2). *Resuscitation*. 2022;171:80-89.
57. Waydhas C, Trentzsch H, Hardcastle TC, et al. Survey on worldwide trauma team activation requirement. *Eur J Trauma Emerg Surg*. 2021;47(5):1569-1580.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kivlehan SM, Hexom BJ, Bonney J, et al. Global emergency medicine: A scoping review of the literature from 2021. *Acad Emerg Med*. 2022;29:1264-1274. doi: 10.1111/acem.14575

APPENDIX A

Members of the 2021 Global Emergency Medicine Literature Review

Global Emergency Medicine Literature Review (GEMLR) Group (in alphabetical order) members:

Kamna S. Balhara, MD, MA, Department of Emergency Medicine, Johns Hopkins University, Baltimore, MD

Norkamari Shakira Bandolin, MD, Department of Emergency Medicine, University of California Davis Health, Sacramento, CA

Holly Bannon-Murphy, MBBS, BA, Emergency and Trauma Centre, Alfred Health, Melbourne, Australia

Torben K. Becker, MD, PhD, Department of Emergency Medicine, University of Florida, Gainesville, FL

Nidhi Bhaskar, BA, Warren Alpert Medical School, Brown University, Providence, RI

Corey B. Bills, MD, MPH, School of Medicine, Department of Emergency Medicine, University of Colorado, Aurora, CO

Joseph Bonney, MBChB, MPH, MScDM, MGCPS, Department of Emergency Medicine, Komfo Anokye Teaching Hospital, Kumasi, Ghana, and Global Health and Infectious Disease Research Group, Kumasi Center for Collaborative Research in Tropical Medicine, Kumasi, Ghana

Addie E. Boone, MD, Department of Emergency Medicine, John Peter Smith Hospital, Fort Worth, TX

Morgan C. Broccoli, MD MPH MSc, Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA

Andrew D. I. Charlton, MBChB, Emergency Department, Bradford Teaching Hospital Trust, Bradford, West Yorkshire, UK

Chuan-Jay Jeffrey Chen, MD, Department of Emergency Medicine, Adventist Health St Helena, St Helena, CA

Daniel K. Cho, Brown University, Providence, RI

Joseph D. Ciano, DO, MPH, MS, Emergency Department, Northwell Health- Long Island Jewish Medical Center, New Hyde Park, NY

Cassandra Clay, DO, MHA, Department of Emergency Medicine, University of Florida, Gainesville, FL

Amanda Collier MD, DTMH, Department of Emergency Medicine, Queen's University, Kingston, Ontario, Canada

Nanaba A. Dawson-Amoah, BSc, MBChB, MBA, Emergency Medicine Department, Greater Accra Regional Hospital, Accra, Ghana

Julia Dixon MD, MPH, School of Medicine, Department of Emergency Medicine, University of Colorado, Aurora, CO

Jonathan W. Dyal, MD, MPH, Department of Emergency Medicine, McGovern Medical School at the University of Texas Health Science Center at Houston, Houston, TX

Stephanie Chow Garbern, MD, MPH, Department of Emergency Medicine, Brown University, Providence, RI,

Melaku Getachew, MD, Department of Emergency and Critical care Medicine, Haramaya University, Harar, Ethiopia

Mohammad Adrian Hasdianda, MD, MSc, MMSc, Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA

Alison S. Hayward, MD, MPH, Department of Emergency Medicine, Brown University, Providence, RI

Braden J. Hexom, MD, Department of Emergency Medicine, Rush University Medical Center, Chicago, IL

Carolyn Hunter, MB BCh BAO, BSc, DTMH, MSc, NHS Scotland, Glasgow, Scotland

Ashley A. Jacobson, MD, Department of Emergency Medicine, Mayo Clinic Health System, Eau Claire, WI

Jennifer Jones, MBBS, Department of Emergency Medicine, Stanford University, Palo Alto, CA

Vinay N. Kampalath, MD, DTM&H, Division of Emergency Medicine, Children's Hospital of Philadelphia, Philadelphia, PA

Sean M Kivlehan, MD, MPH, Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA

Colleen E. Laurence, MD, MPH, Department of Emergency Medicine, University of Cincinnati, Cincinnati, OH

Joseph Leanza, MD, Department of Emergency Medicine, University of Pennsylvania, Philadelphia, PA

Elizabeth Ledger, BM, BCh, MA, MRCPCH, MSc, DTM&H, Children's Emergency Department, Bristol Royal Hospital for Children, Bristol, UK

J. Austin Lee, MD, MPH, Department of Emergency Medicine, Brown University, Providence, RI

Adam C. Levine, MD, MPH, Department of Emergency Medicine, Brown University, Providence, RI

Gideon Loevinsohn, MD, PhD, Department of Emergency Medicine, Massachusetts General Hospital, Boston, MA

Katie Main, MD, Department of Emergency Medicine, University of Washington, Seattle, WA

Sarah McCuskee, MD, MPH, Department of Emergency Medicine, NYU Langone Health and Bellevue Hospital Center, New York, NY

Benjamin Nicholson, MD, Department of Emergency Medicine, Virginia Commonwealth University, Richmond, VA

Gerard M. O'Reilly, MBBS, FACEM, MPH, MBiostat, PhD, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia

Amelia Y. Pousson, MD, MPH, Department of Emergency Medicine, Johns Hopkins University, Baltimore, MD

Nana S. A. Quao, MD, MGCS, Accident and Emergency Centre, Korle Bu Teaching Hospital, Accra, Ghana

Chris A. Rees, MD, MPH, Division of Pediatric Emergency Medicine, Emory University School of Medicine, Atlanta, GA

Pryanka Relan, MD, MPH, Department of Emergency Medicine, Mount Sinai St. Luke's and West Hospitals, New York, NY

Daniel Roberts, MB ChB, Emergency Department, Kent Surrey and Sussex Deanery, Eastbourne District Hospital, Eastbourne, East Sussex, UK

Charlotte M. Roy, MD, MPH, Department of Emergency Medicine, Loma Linda University, Loma Linda, CA

Megan M. Rybarczyk, MD, MPH, Department of Emergency Medicine, University of Pennsylvania, Philadelphia, PA

Gabriel Lucca de Oliveira Salvador, MD, Radiology Department, Brazilian Red Cross, Curitiba, Parana, Brazil

Megan L. Schultz, MD, MA, Departments of Pediatrics and Emergency Medicine, Medical College of Wisconsin, Milwaukee, WI

Anand Selvam, MD, MSc, DTM&H, Department of Emergency Medicine, Yale University, New Haven, CT

Daniel R. Still, MBChB, DIPPEC (RSA), Emergency Department, Karl Bremer Hospital, Cape Town, South Africa

Jonathan Strong, MD, MPH, Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA

Indi Trehan, MD, MPH, DTM&H, Departments of Pediatrics, Global Health, and Epidemiology, University of Washington, Seattle, WA

Alex H. Wang, MD, Department of Emergency Medicine, University of Connecticut, Hartford, CT

W. Tyler Winders, MD, MPH, Department of Emergency Medicine, Hannibal Regional Hospital, Hannibal, MO

Received: 29 June 2022 | Accepted: 1 July 2022

DOI: 10.1111/acem.14556

HOT OFF THE PRESS

Hot off the press: Low-risk, recurrent abdominal pain in the emergency department

Justin Morgenstern MD[1] | Kirsty Challen BSc, MBChB, MRes, PhD[2] | Corey Heitz MD[3] | William K. Milne MD[4]

[1]The University of Toronto, Toronto, Ontario, Canada

[2]Lancashire Teaching Hospitals, Lancashire, UK

[3]Virginia Tech Carilion School of Medicine, Roanoke, Virginia, USA

[4]University of Western Ontario, Goderich, Ontario, Canada

Correspondence
Justin Morgenstern, MD, Department of Emergency Medicine, Markham Stouffville Hospital, 381 Church Street, Markham, Ontario, L3P 7P3, Canada.
Email: justin.morgenstern@gmail.com

Keywords: abdominal pain, chronic pain, emergency medicine

BACKGROUND

The Society for Academic Emergency Medicine has launched an initiative called GRACE, which stands for Guidelines for Reasonable and Appropriate Care in the Emergency Department. The second iteration of these guidelines examine the issue of low-risk, recurrent abdominal pain in the emergency department (ED).[1] Abdominal pain is a very common chief complaint in emergency medicine, and its evaluation is relatively resource intensive, with imaging ordered in more than 50% of patients.[2] There is evidence that we overuse imaging in this population, as the rate of CT increased from 10% in 2001 to 22.5% in 2005, while the rate of hospital admission and diagnosis of important pathology such as appendicitis and diverticulitis remained unchanged.[3] There is a small population of patients who make frequent visits to the ED for recurrent, undifferentiated abdominal pain.[4] The cumulative risk of ionizing radiation from repeated CTs increases the risk of harm in this population, but that harm must be weighed against the possibility of a new acute diagnosis on any given ED visit.[5,6] Patients with repeated visits to the ED for the same complaint may warrant different management plans, as it seems irrational to repeat the same, previously fruitless investigations and treatments time and again. Emergency clinicians need evidence-based guidance on the management of these complex patients. With the publication of the GRACE-2 guidelines, Broder et al.[1] attempt to provide us with the best evidence around four key questions in the management of patients with low-risk, recurrent abdominal pain in the ED.

ARTICLE SUMMARY

This is a guideline developed using the GRADE (Grading of Recommendations Assessment Development and Evaluation) methodology. This is a structured framework for guideline development that involves performing systematic reviews of identified questions, assessment of the evidence using specific GRADE criteria, and development of recommendations using a specific "evidence-to-decision" framework. The GRACE-2 writing group consisted of emergency physicians from across the United States and Canada as well as a patient representative and a psychiatrist with specialization in pain management. These guidelines focused on low-risk patients, defined as patients between 18 and 65 years of age, without history of acute trauma, pregnancy, active cancer, inflammatory bowel disease, organ transplantation, immunosuppression, previous bowel obstruction, or severe active psychiatric disease. They were interested in recurrent abdominal pain, defined as two or more ED visits for a similar complaint within a 12-month period, but excluding repeat visits within the first 30 days (because those may represent the same acute pathology). The group members identified four priority questions which they put through this guideline development process:

DiscussingBroder JS, Oliveira JE, Silva L, et al. Guidelines for Reasonable and Appropriate Care in the Emergency Department 2 (GRACE-2): Low-risk, recurrent abdominal pain in the emergency department. *Acad Emerg Med*. 2022;29(5):526–560.

Supervising Editor: Jeffrey A. Kline, MD.

 © 2022 Society for Academic Emergency Medicine.

1. Should adult ED patients with low-risk, recurrent, and previously undifferentiated abdominal pain receive a repeat CTAP after a negative CTAP within the past 12 months?
2. Should adult ED patients with low-risk, recurrent, and previously undifferentiated abdominal pain with a negative CTAP receive additional imaging with abdominal ultrasound?
3. Should adult ED patients with low-risk, recurrent, and previously undifferentiated abdominal pain receive screening for depression/anxiety?
4. Should adult ED patients with low-risk, recurrent, and previously undifferentiated abdominal pain receive nonopioid and/or nonpharmacologic analgesics?

QUALITY ASSESSMENT

This is a well-done guideline that addressed an important topic in emergency medicine. They performed a thorough and well-defined literature search. They employed a rigorous approach evaluating the available literature (in following the GRADE methodology). They explicitly discuss their choice of outcomes and their relative values and discuss both harms and benefits of interventions. They tried to be inclusive of multiple perspectives, including involving a patient advocate on their panel.

As with all endeavors in evidence-based medicine, there are limitations to consider. These guidelines were designed for application in a setting with advanced imaging, laboratory testing, and specialty referral, and so many not be applicable to all clinicians worldwide. Pediatric patients were also explicitly excluded from this guideline. This group identified four priority questions, but there are many other questions that clinicians might ask when caring for these patients, which this guideline does not address. The involvement of a patient representative provides a valuable perspective, but a single patient cannot possibly represent the myriad of opinions held by our patients. Most importantly, systematic reviews and resultant guidelines can only be as strong as the underlying literature, and this guideline is ultimately limited by the paucity of research in ED patients with low-risk, recurrent abdominal pain.

KEY RESULTS

The guideline panel identified four key questions of interest, and after a systematic review of the literature and evidence synthesis, they make four recommendations.

Recommendation 1: In adult ED patients with low-risk, recurrent, undifferentiated abdominal pain and prior negative CT of the abdomen and pelvis within 12 months, there is insufficient evidence to accurately identify populations in whom repeat imaging can be safely avoided or routinely recommended in the ED. (No recommendation) [No evidence]

Recommendation 2: In adult ED patients with low-risk, recurrent, undifferentiated abdominal pain and a negative CTAP with IV contrast in the ED, we suggest against ultrasound unless there is concern for pelvic or biliary pathology. (Conditional recommendation, against) [Very low certainty of evidence]

Recommendation 3: In adult ED patients with low-risk, recurrent, undifferentiated abdominal pain, we suggest screening for depression and/or anxiety may be performed during the ED evaluation. (Conditional recommendation, either) [Very low certainty of evidence]

Recommendation 4: In adult ED patients with low-risk, recurrent, undifferentiated abdominal pain, we suggest an opioid-minimizing strategy for pain control. (Conditional recommendation) [Consensus, no evidence]

AUTHOR'S COMMENTS

The paucity of evidence for such a common and complex problem in emergency medicine is somewhat disappointing. Given the tremendous uncertainty that remains about the ideal management plans for these patients, it is essential to involve our patients in shared decision making, and ensure that the chosen management plan aligns with their individual values.

TOP SOCIAL MEDIA COMMENTARY

There was some excellent discussion of these guidelines on the SGEM blog. Among other concerns, Sara says, “I understand the value of a ‘mental health’ history but I feel it minimises and devalues these women's symptoms that may very well be physical and that depression/anxiety may be a response to genuine ongoing pain.” Joshua Broder, the lead author of these guidelines, responds, “The GRACE-2 Writing Group shared your concerns about potential stigma of mental health screening or dismissing undiagnosed causes of pain such as endometriosis, and this is discussed in the full guideline document. The patient representative on the panel provided insights as well, and we did our best to balance the potential benefits of identifying a concurrent (though not necessarily causal) relationship between depression, anxiety, and pain.”

He adds that this issue is addressed in the patient information document: “I'M HERE FOR ABDOMINAL PAIN. WHY IS THE DOCTOR ASKING ABOUT DEPRESSION? Depression affects as many as 25% (one in four) of people visiting EDs for abdominal pain. Sometimes mental health conditions such as depression cause physical symptoms such as abdominal pain. Sometimes having a long-term painful condition such as abdominal pain can lead to or make depression worse. When your doctor asks about depression, they aren't saying that your abdominal pain is ‘just in your head.’ They are hoping to find ways to help you feel better. Your doctor knows that you could have a serious cause of abdominal pain, whether or not you have symptoms of depression.”

Dan Eickmeier raises some of the problems we face when trying to care for these patients. “When we are losing beds to patient boarding

in the ER and decreased nursing staffing, can we really be going down the rabbit hole of repeat investigation for what is most likely (especially after one round of comprehensive negative investigations) functional abdominal pain/IBS. I'd love to have a dietician give some counsel on fiber, etc., to all the abdo pain patients we see (also good in diverticulitis, hemorrhoids, fissures, etc) but that ain't happening anytime soon." To which Joshua Broder responds, "I'm excited to see this response, which really poses hypotheses for future research: is some or all recurrent undifferentiated abdominal pain functional pain/irritable bowel syndrome? Would dietary changes be an effective treatment for some patients? We need more studies to understand the potentially diverse causes of recurrent undifferentiated abdominal pain."

The conversation on Twitter seemed to focus on problems of overinvestigation, and frustration at the lack of available science to guide us:

Josh Davis (@MedFactChecks): We really need a way to reduce imaging. we all know 20 CT per year is unnecessary, but no one wants to be the last one to miss something. Care Plans really help with this. Even Harder when they have diagnosis of organic (often IBD) disease even though no records to support it.

Ryan Radecki, MD MS (@emlitofnote): These summaries are just softballs to aspiring early career investigators – "here's a huge important problem, here are the massive holes in the evidence". Great opportunities to do some retrospective work and build to small grants or Ks.

Josh Davis (@MedFactChecks): So true. Lots of work and no good science— but lots of opportunities!

TAKE-TO-WORK POINTS

Given the lack of evidence available to guide us, there is tremendous uncertainty in the most appropriate management plan for these patients. We should be open about that uncertainty with our patients and involve them in shared decision-making to ensure that the chosen management plan matches their personal values. Although the current lack of evidence is somewhat disappointing, uncertainty is a core skill of emergency physicians.

ORCID

Justin Morgenstern https://orcid.org/0000-0001-7969-1382
Corey Heitz https://orcid.org/0000-0002-5040-3217

REFERENCES

1. Broder JS, Silva L OJE, Bellolio F, et al. Guidelines for reasonable and appropriate Care in the Emergency Department 2 (GRACE-2): low-risk, recurrent abdominal pain in the emergency department. *Acad Emerg Med*. 2022;29(5):526-560.
2. Meltzer AC, Pines JM, Richards LM, Mullins P, Mazer-Amirshahi M. US emergency department visits for adults with abdominal and pelvic pain (2007-13): trends in demographics, resource utilization and medication usage. *Am J Emerg Med*. 2017;35(12):1966-1969.
3. Pines JM. Trends in the rates of radiography use and important diagnoses in emergency department patients with abdominal pain. *Med Care*. 2009;47(7):782-786.
4. Jiang HJ, Weiss AJ, Barrett ML. Characteristics of emergency department visits for super-utilizers by payer, 2014: statistical brief #221. In: Healthcare Cost and Utilization Project (HCUP) Statistical Briefs. Agency for Healthcare Research and Quality; 2006. Accessed June 15, 2022. http://www.ncbi.nlm.nih.gov/books/NBK442038
5. Griffey RT, Sodickson A. Cumulative radiation exposure and cancer risk estimates in emergency department patients undergoing repeat or multiple CT. *AJR Am J Roentgenol*. 2009;192(4):887-892.
6. Sodickson A, Baeyens PF, Andriole KP, et al. Recurrent CT, cumulative radiation exposure, and associated radiation-induced cancer risks from CT of adults. *Radiology*. 2009;251(1):175-184.

How to cite this article: Morgenstern J, Challen K, Heitz C, Milne WK. Hot off the press: Low-risk, recurrent abdominal pain in the emergency department. *Acad Emerg Med*. 2022;29:1275-1277. doi: 10.1111/acem.14556

Received: 14 May 2022 | Revised: 22 May 2022 | Accepted: 23 May 2022

DOI: 10.1111/acem.14534

COMMENTARY - UNSOLICITED

The bridge

Rebekah Mannix MD, MPH | **Amanda M. Stewart MD, MPH** | **Joel Hudgins MD, MPH**

Department of Medicine, Division of Emergency Medicine Boston Children's Hospital, Harvard Medical School, Boston, MA, USA

Correspondence
Rebekah Mannix, Division of Emergency Medicine, Boston Children's Hospital, 300 Longwood Avenue, Boston, MA 02115, USA.
Email: rebekah.mannix@childrens.harvard.edu

There is only so much stress a bridge can take before it fails.

Rarely does the stress come in the form of a single catastrophic event—a force so great that the steel twists, bends, and finally crumbles. Rather, it is more likely that the stress comes in small increments—no one individual load greater than the bridge's ability to withstand it. But over time these repetitive injuries silently and insidiously weaken the bridge, until one day—a day that seems like any other—the bridge comes down.

Structural engineers have a name for these silent injuries—fatigue. Bridge fatigue is a progressive process in which structural damage accumulates due to the repetitive application of external loadings such as the burden of too many or too heavy vehicles. These applied loadings may each individually be well below the structural resistance capacity—such that a conventional structural stress analysis might lead to an erroneous conclusion of safety on any given day. Yet, the burden over time is unsustainable. This is because the calculus of risk of damage requires a consideration of stress over time, not just a one-time assessment related to a singular event.

It's much more difficult to define structural fatigue when it comes to the practice of medicine—though this is the state we are in with the current mental and behavioral health crisis, with children and adolescents waiting days and weeks in emergency departments (EDs) to receive definitive psychiatric care. Despite the lack of concrete definitions, the concept of structural fatigue applies—the daily forces that slowly weaken the system to the point that it can no longer serve the purpose for which it was designed. Our providers, our structures, and our system cannot survive these forces much longer, and without intervention, they may crumble.

Historically, in medicine, we have focused on singular events as the source of structural fatigue—a provider suffering from "burnout" or a "difficult" patient. Implicit in the concept of burnout is an individual design flaw, an intrinsic weakness—a bridge not built to withstand the known strain accompanying its functional load, a human whose resilience is inadequate. Similarly, the "difficult" patient can be written off as a fluke or an aberration—an individual truck whose load was accidentally too heavy for the bridge. This focus on singular events distracts from the systemic analysis that is needed to stabilize our structures.

The science of bridge fatigue suggests that we need to ask different questions about the load being imposed on our EDs. ED visits among children and adolescents with mental and behavioral health complaints are up over 250% over the past decade, with roughly half of those patients boarding in the ED while waiting on definitive placement.[1,2] Recently, in our 52 bed ED, our providers cared for 39 patients with behavioral health complaints over 24 h, 12 of whom were admitted or discharged and 27 of whom were forced to remain in our ED awaiting inpatient psychiatric placement. The adverse patient care outcomes that follow such an influx of long-stay behavioral health patients is well known: long wait times for medical and behavioral health patients to be evaluated (8 h+), delayed care for high-acuity patients, increased rates of left without being seen. On this particular day, despite the adverse patient care outcomes and the high provider load, the waiting room was empty by the morning, for some a sign that the bridge had held.

But the bridge is silently, inexorably failing. Our EDs are not designed to house large numbers of behavioral health patients for days at a time. The load is too great and the burden over time is unsustainable—and the cracks in the system are beginning to show. Our patients are not getting the care they need and our providers are strained with the moral injury of providing inadequate care. We can wait for the day when disaster strikes—a preventable code in the waiting room, a mass exodus of nurses—or we can apply some of the basics of bioengineering to understand how the chronic load weakens the system and apply mitigation strategies to prevent the inevitable disaster. There is only so much stress a bridge can take before it fails.

ORCID

Rebekah Mannix https://orcid.org/0000-0003-3608-2792

Supervising Editor: Dr. Kamna Balhara

 © 2022 Society for Academic Emergency Medicine.

REFERENCES

1. Foster AA, Porter JJ, Monuteaux MC, Hoffmann JA, Hudgins JD. Pharmacologic restraint use during mental health visits in pediatric emergency departments. *J Pediatr*. 2021;236:276-283.e2. doi:10.1016/j.jpeds.2021.03.027
2. McEnany FB, Ojugbele O, Doherty JR, McLaren JL, Leyenaar JAK. Pediatric mental health boarding. *Pediatrics*. 2020;146:4, e20201174. doi:10.1542/peds.2020-1174

How to cite this article: Mannix R, Stewart AM, Hudgins J. The bridge. *Acad Emerg Med*. 2022;29:1278–1279. doi: 10.1111/acem.14534

Received: 11 June 2022 | Revised: 11 June 2022 | Accepted: 12 June 2022

DOI: 10.1111/acem.14549

CORRESPONDENCE - UNSOLICITED LETTERS TO THE EDITOR

How to treat venous aneurysm in emergency department?

To the Editor:

Venous aneurysms are extremely rare. At present, approximately 500 cases have been reported in the literature, and the first symptom is often deep vein thrombosis (DVT) or pulmonary embolism,[1] so they have become a very special disease in the emergency department (ED). Venous aneurysms can occur at any age, with the same sex distribution, and can occur in any vein, but they are most common in popliteal vein (45%), portal vein (30%), and superior mesenteric vein (10%). In the process of screening the etiology of venous thromboembolism (VTE), venous aneurysms of the lower limbs is one of the reasons, and popliteal venous aneurysms are the most common. Cystic vein aneurysm is also considered as a high-risk factor for DVT. Surprisingly, anticoagulation alone does not seem to reduce the risk of thromboembolic events (VTE), because venous aneurysms provide an environment for the formation and spread of thrombosis.[2] At least 60% of portal or superior mesenteric vein aneurysms are at risk of complications, including portal hypertension, rupture and fistulation, and subsequent gastrointestinal bleeding.[3]

For symptomatic patients or patients with venous aneurysms with signs of thromboembolism, surgical repair is recommended. There is no consensus on the management of asymptomatic patients. The low incidence rate of venous aneurysms makes it difficult to assess the outcome of interventional therapy, long-term treatment, and the incidence of postthrombotic sequelae. At present, the management of different centers is not unified, and multiagency cooperation should be encouraged. However, since most patients appear in the ED with their first symptoms, multidisciplinary cooperation to treat this disease in the ED should be promoted.

Surgery is the main treatment for venous aneurysms. Endovascular treatment of venous aneurysms has not been determined, although endovascular intervention may be suitable for anatomically challenging venous aneurysms. Cystic venous aneurysms are usually treated by resection of venous aneurysms and repair of lateral veins. Fusiform vein aneurysms can be treated by the following methods: resection and end-to-end anastomosis, intermediate transplantation and bypass, and proximal and distal venous ligation.[4] Superficial venous aneurysms can be treated by venous ligation. Surgical risks include hematoma, wound infection, thrombophlebitis, nerve injury, and recurrence. Individuals usually remain anticoagulant for 3–6 months after surgery. The patency rate was estimated to be between 40% and 93%, and venous aneurysm resection and lateral vein repair were associated with the best long-term patency rate.[5] The emergency intervention decision of venous aneurysm is based on the risk balance, considering the risk of rupture and the incidence rate of VTE related to the intervention. Although it may not be possible to conduct a standard randomized controlled trial in the ED, it emphasizes the need to register data and combine multicenter retrospective studies to help determine individualized treatment.

Zhoupeng Wu MD
Yukui Ma MD

Department of Vascular Surgery, West China Hospital, Sichuan University, Chengdu, Sichuan Province, China
Email: hxyymyk@126.com

ORCID

Zhoupeng Wu https://orcid.org/0000-0002-8704-4908
Yukui Ma https://orcid.org/0000-0001-6986-5238

REFERENCES

1. Teter KA, Maldonado TM, Adelman MA. A systematic review of venous aneurysms by anatomic location. *J Vasc Surg Venous Lymphat Disord*. 2018;6:408-413.
2. Patel R, Hanish S, Baril D, Woo K, Lawrence P. Contemporary management of lower extremity venous aneurysms. *J Vasc Surg Venous Lymphat Disord*. 2019;7:860-884.
3. Gabriel S, Eisenberg N, Kim D, Jaberi A, Roche-Nagle G. Primary venous aneurysms: a 20-year retrospective analysis. *Vascular*. 2020;28:577-582.
4. Sessa C, Nicolini P, Perrin M, Farah I, Magne JL, Guidicelli H. Management of symptomatic and asymptomatic popliteal venous aneurysms: a retrospective analysis of 25 patients and review of the literature. *J Vasc Surg*. 2000;32:902-912.
5. Steiner JE, Drolet BA. Classification of vascular anomalies: an update. *Semin Intervent Radiol*. 2017;34:225-232.

© 2022 by the Society for Academic Emergency Medicine.

Received: 16 May 2022 | Accepted: 22 May 2022

DOI: 10.1111/acem.14535

REFLECTION

Out of Time

Fabrizio Elia MD | **Fabrizio Vallelonga MD**

Emergency Medicine, San Giovanni Bosco Hospital, Torino, Italy

Correspondence
Fabrizio Elia, MD, Piazza del Donatore di Sangue, 3, Turin 10154, Italy.
Email: fa.elia@libero.it

First day, 11.12 a.m.
Crashing noises. Maybe bombs?
A white wall. An old clock on it.
People bustling around. Maybe soldiers?
A stabbing pain without definite cause.
Then a blade penetrating my chest.
Blood taste in my mouth.
It's freezing cold.
Am I dead or alive?

Mean blood pressure always at lower limits with the highest dose of vasopressors.
Continuous fever; still abundant pus from chest drainage.
Things get bad.

Third day, 11.12 a.m.
Noises. This time more distant.
Still that white wall. The old clock staring at me.
It seems to me that time has passed but the clock tells me otherwise.
Dazzling lights are blinding me. Maybe friendly lighthouses?
A dull pain in my back. Someone is tearing off a piece of my skin.
I cannot even move my fingertips.
Alcohol smell in the air.
Cold water on my face.
What kind of a place am I in?

Look at his eyes; sounds like hypokinetic delirium. Who knows if he will come to his senses again?
He does not even lift a finger. At least he does not appear to be suffering.

Fifth day, 11.12 a.m.
Still here. Familiar voices. Maybe my children?
The usual wall. Time seems to be definitely stopped. Or maybe I am going crazy.



Cables coming out of my arms; a large tube from my side. Who tied me up?
Itching on the whole body. An unfamiliar figure is stroking my hair.
Hands finally move. I can at least scratch my abdomen.
Dryness in my mouth, but thirst reminds me I'm alive.
It's very hot.
My head is exploding, nothing makes sense to me.

First day without norepinephrine. Procalcitonin is decreasing. We are on the right path.

Seventh day, 11.12 a.m.
Noises of serving trays.
I recognize the wall and the clock. My only certainties.
Familiar faces help me to eat. Maybe nurses?
A slight tingle on my legs; they are very faint.
But I'm sitting on a bed; who knows if I can stand?
The taste of the sauce brings back memories of my family.
I feel better.
Maybe I'm in the hospital.

I did not think he could sit on the bed so quickly.
How lovely it's to see his children helping him eat.

Supervising Editor: Dr. Brian Zink

 © 2022 Society for Academic Emergency Medicine.

Ninth day, 11.12 a.m.
At last, silence. Only my wife's voice next to me.
My wife sitting by my bed, crying.
Still 11:12 a.m. But should not she be working at this time?
She kisses me. I try too.
We hug each other tightly.
I recognize her perfume.
I no longer know if it's cold or hot.
But I'm alive and I'm sure of it.

Several days later, 11.12 a.m.
"*I'm really glad you came through and can get out of here*".
Thank you, Doctor. I know I was very sick. I have few and vague memories, but my wife told me about how close to death I was. I sincerely thank everyone for what you have done for me.
"*That's our job. Is there anything we can still do?*"
Yes doctor, although it's probably not so important for you.
Please change the batteries to that damned old clock!

ACKNOWLEDGMENTS

To all the patients who shared their thoughts. To everyone who put up with our disregard for details. A special gratitude to Fabio Ventrice for the photograph.

ORCID

Fabrizio Elia https://orcid.org/0000-0001-7804-9937
Fabrizio Vallelonga https://orcid.org/0000-0002-4628-6767

How to cite this article: Elia F, Vallelonga F. Out of Time. *Acad Emerg Med.* 2022;29:1281-1282. doi: 10.1111/acem.14535

Received: 19 May 2022 | Accepted: 23 May 2022

DOI: 10.1111/acem.14538

REFLECTION

Just another day

Richard J. Cunningham MD

Department of Emergency Medicine, Emergency Medicine Resident, PGY-3, Maricopa Medical Center, Phoenix, Arizona, USA

The day starts like any other. I roll in at 6 AM to take sign-out from the overnight team who sport purple bags under their eyes. The board is fairly clean. Nobody is too sick, mostly stable abdominal and chest pain work-ups and a few psych holds pending placement. "Follow up on bed 16s scan, repeat a CK on 20 once he gets his fluids, I'll page surgery after sign-out and circle back, etc., etc." Just another day.

The morning is calm, I sip coffee at my desk when I'm not roaming from room to room, pushing on bellies, palpating pulses, probing wounds…. Soon, I get a call on my phone: "EMS is bringing in a full code in 5." I head up to the resus bay. *Just another day*, I think to myself. The medics arrive and give report: 24-year-old at a detox facility, witnessed seizure-like activity and emesis, then went unresponsive. Received bystander CPR. PEA and asystole at each pulse check, three rounds of epi prior to arrival.

Our team goes to work. Pulse check: PEA, back on the chest. Another epi please, bicarb, calcium. Pulse check: PEA, back on the chest. ET tube in place, positive end-tidal. No pneumo on ultrasound. Pulse check: asystole, back on the chest. Another round of epi. No pericardial effusion. End tidal is dropping. After rounds and rounds of compressions, my H's and T's fail me. "Does anyone have any other ideas?" I announce to the room. I receive silence in return. Pulse check: "Time of death: 9:15 AM." *Just another day*, I tell myself. "Let's have a moment of silence…"

Her mother shows up later that morning. I deliver the news. She looks back at me stone-faced. "I was expecting this," she says. The patient had a long history of mental health and substance abuse problems, not uncommon at our safety-net hospital. It's only toward the end of our conversation when she shows any emotion, choking back tears "At least I can tell her children that their mother was good and that she went to heaven." *Just another day*, I think to myself. Dejected, I leave as the social worker directs her to the morgue.

I pick up the next chart: Ear pain. I scrape out a large ball of cerumen and try to recall the Spanish word for earwax. "Ahhh," she sighs in relief, "Muchas gracias, muy amable." From there on, my shift is smooth sailing. I reassure a "GI bleed" that their hemorrhoids are nothing to fret over. I win a wager on a D-Dimer and send the patient home, both of us in good spirits. Hopping over to fast track, I clinch the diagnosis of mallet finger, splint in extension, and move along. *Just another day*, I think, cheering myself on. Some days it's the little victories that outweigh the big defeats.

"The nurse will bring you your paperwork, then you're good to go." Not the first time I've said it, and certainly not the last. "And don't forget to make an appointment with your primary doc!" I add as I close the curtain behind me. I head back to my computer, hit the discharge button, and check the clock. 2:02 PM, I've crossed the finish line, no more patients this shift. As the evening shift crew trickles in I update my patients on their results, catch up on notes, and clean up my dispos for sign-out.

Just another day, I think to myself. Except today is different. After 3 years in the trenches, today is my last shift at the county emergency department, my home base. Of course there are 2 months of residency left, but I'll be off-service, working at outside community hospitals, one more month in the MICU and another in the ED.

I hand over my patients. I have not told my colleagues, my friends, that today is my last. Given the timing they do not realize it, and it just does not feel right to tell them. Why make a fuss? *It really is just another day*, I convince myself.

I hand in my phone in exchange for car keys and walk out the ambulance doors into the bright afternoon sun. I look back and contemplate the hospital and the people inside it who helped form me into the physician I am. I've worked hundreds of shifts and taken care of thousands of patients, but in this job there's always another patient, always another shift, always another turnover. The cycle goes on. Today, like every other day, was *just another day*.

ORCID

Richard J. Cunningham https://orcid.org/0000-0002-6494-2537

How to cite this article: Cunningham RJ. Just another day. *Acad Emerg Med*. 2022;29:1283-1283. doi: 10.1111/acem.14538

Decision Editor: Dr. Brian Zink

 © 2022 Society for Academic Emergency Medicine.