



Systematic Review Methodology for the Fatigue in Emergency Medical Services Project

P. Daniel Patterson, J. Stephen Higgins, Patricia M. Weiss, Eddy Lang & Christian Martin-Gill

To cite this article: P. Daniel Patterson, J. Stephen Higgins, Patricia M. Weiss, Eddy Lang & Christian Martin-Gill (2018) Systematic Review Methodology for the Fatigue in Emergency Medical Services Project, Prehospital Emergency Care, 22:sup1, 9-16, DOI: 10.1080/10903127.2017.1380096

To link to this article: <https://doi.org/10.1080/10903127.2017.1380096>



Published online: 11 Jan 2018.



Submit your article to this journal [↗](#)



Article views: 231



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 9 View citing articles [↗](#)

FOCUS ON METHODOLOGY

SYSTEMATIC REVIEW METHODOLOGY FOR THE FATIGUE IN EMERGENCY MEDICAL SERVICES PROJECT

P. Daniel Patterson, PhD, NRP, J. Stephen Higgins, PhD, Patricia M. Weiss, MLIS, Eddy Lang, MDCM, CCFP (EM), Christian Martin-Gill, MD, MPH

ABSTRACT

Background: Guidance for managing fatigue in the Emergency Medical Services (EMS) setting is limited. The Fatigue in EMS Project sought to complete multiple systematic reviews guided by seven explicit research questions, assemble the best available evidence, and rate the quality of that evidence for purposes of producing an Evidence Based Guideline (EBG) for fatigue risk management in EMS operations. **Methods:** We completed seven systematic

reviews that involved searches of six databases for literature relevant to seven research questions. These questions were developed *a priori* by an expert panel and framed in the Population, Intervention, Comparison, and Outcome (PICO) format and pre-registered with PROSPERO. Our target population was defined as persons 18 years of age and older classified as EMS personnel or similar shift worker groups. A panel of experts selected outcomes for each PICO question as prescribed by the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) methodology. We pooled findings, stratified by study design (experimental vs. observational) and presented results of each systematic review in narrative and quantitative form. We used meta-analyses of select outcomes to generate pooled effects. We used the GRADE methodology and the GRADEpro software to designate a quality of evidence rating for each outcome. **Results:** We present the results for each systematic review in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA). More than 38,000 records were screened across seven systematic reviews. The median, minimum, and maximum inter-rater agreements (Kappa) between screeners for our seven systematic reviews were 0.66, 0.49, and 0.88, respectively. The median, minimum, and maximum number of records retained for the seven systematic reviews was 13, 1, and 100, respectively. We present key findings in GRADE Evidence Profile Tables in separate publications for each systematic review. **Conclusions:** We describe a protocol for conducting multiple, simultaneous systematic reviews connected to fatigue with the goal of creating an EBG for fatigue risk management in the EMS setting. Our approach may be informative to others challenged with the creation of EBGs that address multiple, inter-related systematic reviews with overlapping outcomes. **Key words:** systematic reviews; evidence-based guidelines; fatigue; EMS

Received July 14, 2017 from the Department of Emergency Medicine, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania (PDP, CMG); National Highway Traffic Safety Administration, Washington, District of Columbia (JSH); Health Sciences Library System, University of Pittsburgh, Pittsburgh, Pennsylvania; (PMW); Department of Emergency Medicine, Cummins School of Medicine, University of Calgary, Calgary, Alberta, Canada (ESL). Revision received September 5, 2017; accepted for publication September 12, 2017.

Work performed on this manuscript was supported with funding from the U.S. Department of Transportation, National Highway Traffic Safety Administration to the National Association of State EMS Officials (NASEMSO): contract/grant number: DTNH2215R00029. The views contained in this article are those of the authors and not necessarily those of the National Highway Traffic Safety Administration.

All authors contributed to the conception of the design of this work. All authors contributed to acquisition of study data. PDP led the analysis, and all authors contributed to the interpretation of findings, drafting of the manuscript, and providing critically important intellectual content. All authors reviewed and approved the final version and agree to be accountable for all aspects of the work.

Address correspondence to P. Daniel Patterson, PhD, University of Pittsburgh, Emergency Medicine, 3600 Forbes Avenue, Iroquois Bldg., Suite 400A, Pittsburgh, PA 15260, USA. E-mail: pattersonpd@upmc.edu

Published with license by Taylor & Francis © 2018 P. Daniel Patterson, J. Stephen Higgins, Patricia M. Weiss, Eddy Lang, and Christian Martin-Gill

doi: [10.1080/10903127.2017.1380096](https://doi.org/10.1080/10903127.2017.1380096)

PREHOSPITAL EMERGENCY CARE 2018;22:9–16

BACKGROUND

The Fatigue in Emergency Medical Services (EMS) Project is one of several EMS-focused efforts to cre-

ate Evidence Based Guidelines (EBGs) for the purpose of improving the safety of patients and EMS personnel (1). This project began with seven research questions prepared by a panel comprised of experts in sleep medicine, fatigue, EMS, administration, and emergency medicine (1). Questions and outcomes were developed following an iterative process prescribed by the Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) methodology (1, 2). The seven research questions dictated seven distinct systematic reviews of published and unpublished literature from multiple sources.

Systematic reviews involve comprehensive searches of published and unpublished literature guided by *a priori* defined research questions and search parameters (3). They comprise a thorough examination of multiple databases. Systematic reviews are scientific endeavors that differ from narrative reviews, scoping reviews, or rapid reviews (4). Narrative reviews summarize selectively identified literature with a potentially subjective interpretation of findings and hence are at risk of bias related to variable methods for gathering and evaluating evidence (5). Scoping reviews seek to rapidly summarize key concepts or topics of interest within a defined area of research (6). Rapid reviews are more rigorous than narrative reviews and more focused than scoping reviews. They require less time than systematic reviews, yet are less thorough due to numerous methodological shortcuts (7). A systematic review is a type of research in its own right: a specific study design that aims to be transparent and reproducible by adhering to explicit steps for the purposes of compiling all relevant information connected to focused research questions (5).

Systematic reviews are the basis for EBGs (8). The Institute of Medicine (IOM) defines EBGs as “statements that include recommendations intended to optimize patient care that are informed by a systematic review of the evidence and an assessment of the benefits and harms of alternative care options” (8). The development of EBGs in the EMS setting has accelerated in recent history (9–13).

In this paper, we describe the unique methods and protocols common to seven distinct but inter-related systematic reviews registered prospectively with PROSPERO; an international database of systematic review protocols (PROSPERO 2016 registration numbers: CRD42016040097; CRD42016040099; CRD42016040101; CRD42016040107; CRD42016040110; CRD42016040112; CRD42016040114) (1). We describe the detailed procedures for systematically searching the published evidence and our chosen approach to summarizing the evidence. This paper may be useful to others charged with completing multiple systematic reviews for the purpose of developing EBGs. Analytical techniques unique to a particular review, such as a meta-analysis, are reported separately (14–20).

METHODS

Study Design and Protocol

In order to cast the most comprehensive net, it is important to search multiple databases and other sources when performing systematic reviews (21). We searched five databases and one website: PubMed/Medline, the Cumulative Index to Nursing and Allied Health Literature (CINAHL), Scopus, PsycINFO, the Published International Literature on Traumatic Stress (PILOTS), and the publications section of the National Institute of Justice (NIJ) website. We selected these databases and website because the topic of interest (fatigue) spans multiple fields, disciplines, and occupations, and unique literature relevant to our research questions was likely available in multiple repositories.

Types of Participants

We included research that involved persons 18 years of age and older classified as EMS personnel or similar shift worker groups. Shift work refers to “any arrangement of work hours other than standard daylight hours” (22). We defined similar worker groups as shift workers whose job activity requires multiple episodes of intense concentration and attention to detail per shift, with serious consequences resulting from a lapse in concentration (1). These included both health professions such as nurses and other occupational areas such as aviation and the military. Studies that did not include shift workers were excluded during screening or full-text review. The decision to include non-EMS shift workers was based on the belief that all types of shift workers are challenged by fatigue in the workplace and EMS can learn from these experiences. Our use of the GRADE methodology allowed for structured consideration of findings from non-EMS shift workers and downgrading this research for indirectness (23).

Types of Interventions

The type(s) of interventions targeted varied by research question. For our first research question (CRD42016040097), the search strategy focused on articles reporting use of fatigue or sleepiness survey instruments to assess/diagnose fatigue in the EMS workplace or a workplace environment of related shift worker groups (14). The search parameters for our second question (CRD42016040099) targeted comparisons of fatigue or fatigue-related outcomes by different shift durations (e.g., 12-hour vs. 24-hour shifts) (15). For our third research question (CRD42016040101), we retained studies that included multiple comparisons with caffeine as a component part of one or more study arms (e.g., caffeine versus placebo, caffeine plus sleep versus caffeine only versus placebo, and so on)

(16). Our fourth research question focused on napping/sleeping during shift work (CRD42016040107) (17). We retained studies that tested the impact of a scheduled nap/sleep period during shift work as a component of one or more study arms. The fifth research question focused on studies that evaluated the impact of fatigue education and/or training on safety and related outcomes (CRD42016040110) (18). Each study had to include education on fatigue and/or sleep health as a minimum, but was retained if investigators also described use of education on related topics (e.g., general health and wellness). For the sixth research question (CRD42016040112), we retained studies that reported tests or evaluations of the effectiveness of a biomathematical model in the operational setting to address fatigue and fatigue-related risks (19). We excluded research if the aim of the study was to calibrate the biomathematical model rather than test the impact on operational outcomes like safety. The intervention search parameters for our seventh research question (CRD42016040114) targeted studies that evaluated the impact of interventions or programs designed to modify task load (or workload) to mitigate fatigue, mitigate fatigue related risks, and/or to improve sleep for EMS personnel and related shift worker groups (20).

Types of Outcome Measures

Outcomes for each systematic review were selected a priori by the project's expert panel and classified as critical or important based on procedures prescribed by the GRADE methodology (2). We describe the process for outcome selection in a prior publication (1).

Search Methods for Studies

A research librarian (PMW) executed searches for all seven systematic reviews individually using five bibliographic database products and one website: PubMed (National Library of Medicine), Scopus (Elsevier B.V.), PsycINFO (Ovid Technologies), CINAHL (EBSCO Industries, Inc.), and Published International Literature on Traumatic Stress (PILOTS) (ProQuest). National Institute of Justice (NIJ) publications were also searched (<http://www.nij.gov/publications/Pages/welcome.aspx>). Each systematic review search incorporated multiple terms covering concepts outlined in the papers that report on the systematic review's findings (14–20). All searches combined standardized terms drawn from controlled vocabularies (such as Medical Subject Headings for PubMed's MEDLINE database), author-selected keywords, and text words. The PILOTS and NIJ searches, already contextualized in stress and law enforcement, respectively, were simpler text word searches covering primarily the fatigue and intervention concepts. All searches included literature from January 1980

to September 2016. The bibliographies of articles retained for full-text review were searched for additional relevant literature. To view the search strategy for each systematic review, access the Online Supplement Appendix A document referenced in separate publications (14–20).

Data Collection and Selection of Studies

Screening

In keeping with standard practice for systematic reviews, we trained two co-investigators assigned to a particular systematic review to independently screen titles and abstracts of search results and identify studies potentially germane to each systematic review's study objectives. We did not focus on identifying prior experience with screening because experience alone does not equate to mastery or substantially lower error in screening (24). Our training program involved instructing co-investigators to apply inclusion/exclusion criteria independently to several actual records. The principal investigator discussed decision-making with screeners to address potential confusion and improve clarity.

We used DistillerSR by Evidence Partners and trained co-investigators to navigate the tool for efficiency purposes. Based on previous research (25), we estimated one-to-two minutes for screeners to reach a decision of inclusion or exclusion for each title/abstract (record) reviewed. Training and use of electronic tools like Distiller improves screening and agreement between screeners (26). We used the Kappa statistic to determine inter-rater agreement of the initial screening decisions based on review of the title and abstract alone. We completed an additional review of decision-making by screeners by comparing their initial include/exclude decisions for six of seven systematic reviews against the include/exclude decisions of the principal investigator. We did this by selecting a random sample of 50 records (titles/abstracts) from the initial pool of records selected for each systematic review and having the principal investigator generate an include/exclude decision. The principal investigator reached judgment on each record while unaware of the decisions made by screeners. We used these data to calculate the percentage of agreement between the screeners and principal investigator.

Conflict Adjudication

Independent reviews often produce a sub-set of records where reviewers disagree. Two investigators, other than those assigned to the initial screening, worked together (simultaneously) to review disagreements and adjudicate a decision of inclusion/exclusion against the systematic review's criteria. These criteria included: a) the title and/or abstract included a description of

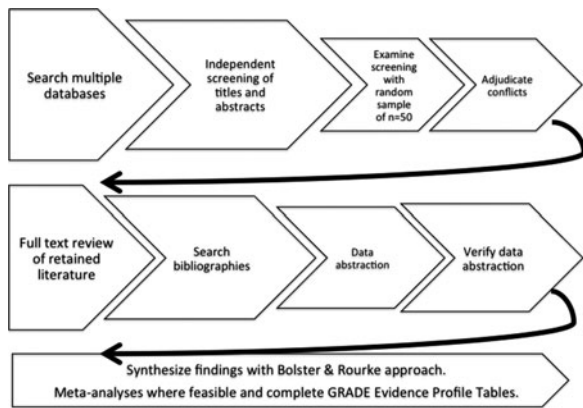


FIGURE 1. Study protocol for systematic reviews.

the population of interest; b) the title and/or abstract described the intervention(s) of interest for the systematic review in question; c) the title and/or abstract described the comparison(s) of interest for the systematic review in question; and/or d) the title and/or abstract described the outcome(s) of interest for the systematic review in question. Following adjudication, we assembled the titles and abstracts for each record retained and then retrieved the full-text articles for further review. Figure 1 is a graphic of our study protocol.

Full-Text Review

We trained co-investigators to use a structured data abstraction form and instructed each to work independently to abstract key information from full-text articles. The key information abstracted included: study design, participant characteristics, intervention characteristics, comparisons, outcome measures, and key findings. Co-investigators reviewed data abstractions completed by their collaborators to verify the information abstracted. Disagreements were adjudicated by discussion between co-investigators and the systematic review's senior author. All articles excluded during the full-text review, and the reasons for exclusion, appear in the Online Supplement Appendix C of each systematic review. We systematically excluded non-scientific journal literature (e.g., newsletters), book chapters, conference abstracts, dissertations, and thesis documents. Co-investigators searched bibliographies of the retained and excluded literature to identify additional, potentially relevant research. Literature identified during bibliography searches was reviewed in full-text.

Risk of Bias Assessment

The three most senior investigators assigned to a systematic review completed a risk of bias assessment for the retained literature. The biases and limitations associated with observational study designs

were summarized with the GRADE tool for evaluating observational research (27). The GRADE tool summarizes risk of bias for non-RCTs across four domains: 1) participant selection; 2) measurement of exposure and outcome; 3) control for confounding; and 4) completeness of follow-up (27). We used the Cochrane Collaboration's Risk of Bias tool to report biases and limitations associated with experimental study designs (28). The Cochrane tool evaluates the risk of bias across six domains: selection bias (i.e., sequence generation and allocation concealment); performance bias (i.e., blinding of participants and personnel); detection bias (i.e., blinding of outcome assessment); attrition bias (i.e., incomplete outcome data); reporting bias (i.e., selective reporting); and other bias (i.e., other sources of bias not addressed in other domains) (28). We used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) tool to assess bias in fatigue and sleepiness assessment survey instruments. The QUADAS-2 assesses the risk of bias across four domains: patient selection, index test, reference standard, and flow and timing (29). Disagreements in the assessment of bias for articles reviewed were resolved by discussion between three senior investigators assigned to a systematic review.

Statistical Analysis

The three most senior investigators for each systematic review worked simultaneously to summarize key findings connected to critical and important outcomes of interest. These investigators were selected based on content knowledge and research experience. They used a categorical system developed by Bolster and Rourke and adapted for purposes of this project (30). Our adaptation permitted categorization of an individual study's findings as favorable, unfavorable, mixed/inconclusive, or no impact for mitigating fatigue, mitigating fatigue-related risks, and/or improving sleep. "Favorable" was assigned when the three senior investigators determined that findings reported in a journal article favored the intervention (e.g., use of naps to mitigate fatigue, use of caffeine to improve alertness, or use of fatigue education and training to improve sleep quality). The category "unfavorable" was assigned when findings did not favor the intervention under study (e.g., when use of naps during shift work did not improve performance). "Mixed/inconclusive" was assigned when the findings of a study suggested the presence of both positive and negative effects of an intervention. The category "mixed/inconclusive" was also assigned when the results reported for a specific outcome were inadequate to draw a definitive conclusion or interpretation. Investigators assigned the category "no impact" when an article's findings were reported to have no statistical or clinically meaningful impact on

outcomes. All decisions and interpretation of findings as favorable, unfavorable, mixed/inconclusive, or no impact were based on consensus by the systematic review's three senior investigators.

When two or more studies for any systematic review used experimental study designs (i.e., randomized, experimental crossover, or quasi-experimental), and reported results for a specific outcome, we pooled data for a meta-analysis. We used RevMan software (version 5.3, Copenhagen, Denmark) to calculate the standardized mean difference (SMD) and 95% confidence intervals (CIs) of a pooled main effect (28). The SMD is the estimated intervention effect of each study relative to the variability in the study (31). An SMD greater than zero indicates the treatment condition is more efficacious than the control/comparison condition. The SMD is non-significant if the corresponding 95% confidence interval is wide and overlaps 0. The I^2 statistic was calculated as a standard measure of heterogeneity (28). The I^2 is the percentage of total variation across the included studies related to heterogeneity and not chance. Values range from 0% to 100% with higher values (e.g., >50%) signifying substantial heterogeneity.

Quality of Evidence

Three senior investigators assigned to a systematic review and the team's GRADE methodologist used the GRADE framework and GRADEpro software to summarize and rate the quality of retained research (quality of evidence or also referred to as "certainty in effect") across each outcome (32). Every row of the GRADE evidence profile table contains key information about the quality or certainty of evidence germane to those outcomes rated as critical and important. Key information includes: number of studies per outcome, judgments about underlying quality of evidence across five domains (i.e., risk of bias, inconsistency, indirectness, imprecision, and other considerations), and statistical results (32). The narrative summary feature of the GRADEpro software was used when pooled effects were not estimable. The GRADEpro software generates an overall quality rating for studies linked to a particular outcome and presented as very low, low, moderate, or high.

Reporting

We present the findings for each systematic review in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) (33).

RESULTS

More than 38,000 records were screened across seven systematic reviews (Table 1). The median, minimum,

and maximum inter-rater agreement (Kappa) between screeners for our seven systematic reviews were 0.66, 0.49, and 0.88, respectively (14–20). The percentage of agreement between the screeners and Principal Investigator was strongly correlated and ranged from 90% to 100% (14–20). The median, minimum, and maximum number of records retained for the seven systematic reviews was 13, 1, and 100, respectively.

DISCUSSION

Systematic reviews collate and evaluate the most relevant and highest quality research (evidence) to answer specific questions about the efficacy or effectiveness of interventions. The synthesis of findings across multiple studies, guided by a specific research question, is used to inform general knowledge, shape conclusions, and guide care strategies and policy decisions. Systematic reviews usually allow us to look beyond one or two studies and take into full consideration the findings from multiple studies involving different study samples (or populations) with possibly differing results (4). Systematic reviews play an increasingly important role in shaping policy and practice. Like any research study, a systematic review is subject to bias and limitations (34). Poorly conducted systematic reviews may offer an incomplete or inaccurate interpretation of current evidence (34).

We describe the methods and procedures common to seven systematic reviews (14–20). We describe the process of decision-making associated with fundamental steps where bias may impact results. We provide a comprehensive and transparent summary of our approach so that others may assess the rigor of our methods and replicate findings. To the best of our knowledge, there is no previously published compilation of systematic reviews germane to our seven research questions, population of interest, interventions of interest, comparisons of interest, or outcomes of interest. We contribute to the literature by simultaneously reviewing the evidence on a variety of intervention questions that inform fatigue risk management for shift workers, including EMS personnel. This research establishes a knowledge base for policy-making and setting priorities for further research.

LIMITATIONS

Several biases are common to systematic reviews. Existing literature and evidence relevant to our research questions may have not been retrieved from the databases selected. Published research germane to one or more of our seven systematic reviews may be indexed in a database other than those searched by our research librarian. We addressed this limitation by including in our protocol for all systematic reviews, searches of the bibliographies of articles reviewed

TABLE 1. Summary of findings for seven systematic reviews

Systematic Review	PROSPERO# Dates searched	Population Intervention Comparison Outcome	Study designs included	Records screened	Kappa on initial screening	Percentage agreement with P.I.	Total studies retained
1	CRD42016040097 1/1/1980-9/30/2016	[P] EMS personnel or similar shift worker groups [I] Use of fatigue and/or sleepiness survey instruments [C] Compared to a gold standard or indirect measure of standard [O] Reliability, validity, sensitivity, and specificity	Observational and experimental	1,257	0.63	90%	34
2	CRD42016040099 1/1/1980-9/30/2016	[P] EMS personnel or similar shift worker groups [I] Shorter shift duration [C] Compared to a longer shift duration [O] Patient safety, personnel safety, personnel performance, acute fatigue, sleep and sleep quality, retention/turnover, long-term health, burnout/stress, cost to system	Observational and experimental	21,674	0.49	100%	100
3	CRD42016040101 1/1/1980-9/30/2016	[P] EMS personnel or similar shift worker groups [I] Use of caffeine [C] Compared to no use of caffeine or use of placebo [O] Patient safety, personnel safety, personnel performance, acute fatigue, sleep and sleep quality, long-term health	Experimental	1,401	0.55	100%	8
4	CRD42016040107 1/1/1980-9/30/2016	[P] EMS personnel or similar shift worker groups [I] Napping/sleeping during shift work [C] Compared to no napping/sleeping during shift work [O] Patient safety, personnel safety, personnel performance, acute fatigue, sleep and sleep quality, retention/turnover, long-term health, cost to system	Experimental	4,660	0.78	100%	13
5	CRD42016040110 1/1/1980-9/30/2016	[P] EMS personnel or similar shift worker groups [I] Use of fatigue education/training [C] Compared to no use of fatigue/sleep education/training [O] Patient safety, personnel safety, personnel performance, acute fatigue, sleep and sleep quality, long-term health, burnout/stress	Observational and experimental	3,817	0.88	100%	18
6	CRD42016040112 1/1/1980-9/30/2016	[P] EMS personnel or similar shift worker groups [I] Implementation of biomathematical fatigue modeling [C] Compared to no use of biomathematical fatigue modeling [O] Patient safety, personnel safety, personnel performance, acute fatigue, sleep and sleep quality, long-term health, cost to system	Experimental	2,777	0.80	n/a*	1
7	CRD42016040114 1/1/1980-9/30/2016	[P] EMS personnel or similar shift worker groups [I] Modification of task load during shift work [C] Compared to no modification of task load during shift work [O] Patient safety, personnel safety, personnel performance, acute fatigue, cost to system	Observational and experimental	3,394	0.66	92%	5

Notes: P.I. = Principal Investigator.

*The percentage agreement between the principal investigator and screeners was not calculated for one systematic review (PROSPERO 2016:CRD42016040112), given that the assigned screeners are experts in the field of biomathematical fatigue modeling.

(i.e., the reference list) in full-text form. The search of bibliographies of retained literature increase the yield of relevant research and improve the completeness of systematic reviews.

Our procedure for screening titles and abstracts has limitations. Systematic reviews begin with thousands of potentially eligible records (articles, book chapters, newsletters, etc.). Screening is used to reduce the total number of relevant articles by applying specific inclusion/exclusion criteria to the title, abstract, or both title and abstract. Screening thousands of records is time-consuming and it is routine to train students, or use algorithm-based electronic techniques for purposes of screening. We completed brief training and orientation of co-investigators to the screening process. We instructed co-investigators to apply explicit criteria when making the initial decision to include or exclude. Despite training, conflict between screeners is common. The median inter-rater agreement (Kappa) between screeners for our seven systematic reviews was 0.66 (14–20), which is similar to inter-rater agreement reported in previous systematic reviews (35–38). The percentages of agreement between the screeners and principal investigator were substantial (range 90% to 100%) (14–20). The percentage agreement between the principal investigator and screeners was not calculated for one systematic review (PROSPERO 2016:CRD42016040112) (19), given that the assigned screeners are experts in the field.

Key information abstracted from retained literature may be imprecise and impact the findings of systematic reviews. Horton and colleagues showed that errors are common during data abstraction for systematic reviews (24). Buscemi and colleagues determined that the occurrence of erroneous data abstraction was reduced with use of multiple reviewers (verifiers) of the same article (39). For our purposes, key information from the final pool of retained literature was abstracted into tables for all seven systematic reviews. For each systematic review, we abstracted descriptive data and stated measures of central tendency reported in the tables, graphs, and text of papers for purposes of meta-analyses and calculation of pooled effects. Co-investigators independently verified abstraction performed by colleagues. They evaluated the key information abstracted to confirm that information was accurate and comprehensive.

CONCLUSIONS

We describe a protocol for conducting multiple, simultaneous systematic reviews to inform the creation of an EBG for fatigue risk management in the EMS setting. Our approach may be informative to others challenged with creation of EBGs and multiple, inter-related systematic reviews with overlapping outcomes.

References

1. Patterson PD, Higgins JS, Lang ES, Runyon MS, Barger LK, Studnek JR, Moore CG, Robinson K, Gainor D, Infanger A, et al. Evidence-based guidelines for fatigue risk management in EMS: formulating research questions and selecting outcomes. *Prehosp Emerg Care*. 2017;21(2):149–56. <https://doi.org/10.1080/10903127.2016.1241329> PMID:27858581.
2. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, Alderson P, Glasziou P, Falck-Yitter Y, Schunemann HJ. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol*. 2011;64(4):395–400. <https://doi.org/10.1016/j.jclinepi.2010.09.012> PMID:21194891.
3. Uman LS. Systematic reviews and meta-analyses. *J Can Acad Child Adolesc Psychiatry*. 2011;20(1):57–9. PMID:21286370.
4. Akobeng AK. Understanding systematic reviews and meta-analysis. *Arch Dis Child*. 2005;90(8):845–8. <https://doi.org/10.1136/adc.2004.058230> PMID:16040886.
5. Garg AX, Hackam D, Tonelli M. Systematic review and meta-analysis: when one study is just not enough. *Clin J Am Soc Nephrol*. 2008;3(1):253–60. <https://doi.org/10.2215/CJN.01430307> PMID:18178786.
6. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol*. 2005;8(1):19–32. <https://doi.org/10.1080/1364557032000119616>.
7. Garrity C, Stevens A, Gartlehner G, King V, Kamel C. Cochrane Rapid Reviews Methods Group to play a leading role in guiding the production of informed high-quality, timely research evidence syntheses. *Syst Rev*. 2016;5(1):184. <https://doi.org/10.1186/s13643-016-0360-z> PMID:27793186.
8. Institute of Medicine. Clinical practice guidelines we can trust. Washington, DC: The National Academies of Sciences; 2011.
9. Lang ES, Spaite DW, Oliver ZJ, Gotschall CS, Swor RA, Dawson DE, Hunt RC. A national model for developing, implementing, and evaluating evidence-based guidelines for prehospital care. *Acad Emerg Med*. 2012;19(2):201–9. <https://doi.org/10.1111/j.1553-2712.2011.01281.x> PMID:22320372.
10. Bulgar EM, Snyder D, Schoelles K, Gotschall C, Dawson D, Lang ES, Sanddal ND, Butler FK, Fallat M, Taillac P, et al. An evidence-based prehospital guideline for external hemorrhage control: American College of Surgeons Committee on Trauma. *Prehosp Emerg Care*. 2014;18(2):163–73. <https://doi.org/10.3109/10903127.2014.896962> PMID:24641269.
11. Gausche-Hill M, Brown KM, Oliver ZJ, Sasson C, Dayan PS, Eschmann NM, Weik TS, Lawner BJ, Sahni R, Falck-Yitter Y, et al. An evidence-based guideline for prehospital analgesia in trauma. *Prehosp Emerg Care*. 2014;18(Suppl 1):25–34. <https://doi.org/10.3109/10903127.2013.844873> PMID:24279813.
12. Thomas SH, Brown KM, Oliver ZJ, Spaite DW, Lawner BJ, Sahni R, Weik TS, Falck-Yitter Y, Wright JL, Lang ES. An evidence-based guideline for the air medical transportation of prehospital trauma patients. *Prehosp Emerg Care*. 2014;18(Suppl 1):35–44. <https://doi.org/10.3109/10903127.2013.844872> PMID:24279767.
13. Brown KM, Macias CG, Dayan PS, Shah MI, Weik TS, Wright JL, Lang ES. The development of evidence-based prehospital guidelines using a GRADE-based methodology. *Prehosp Emerg Care*. 2014;18(Suppl 1):3–14. <https://doi.org/10.3109/10903127.2013.844871> PMID:24279739.
14. Patterson PD, Weaver MD, Fabio A, Teasley EM, Renn ML, Curtis BR, Matthews ME, Kroemer AJ, Xun X, Bizhanova Z, et al.

- Reliability and validity of survey instruments to measure work-related fatigue in the Emergency Medical Services setting: a systematic review. *Prehosp Emerg Care*. 2018;22(S1):17–27.
15. Patterson PD, Runyon MS, Higgins JS, Weaver MD, Teasley EM, Kroemer AJ, Matthews ME, Curtis BR, Flickinger KL, Xun X, et al. Shorter versus longer shift duration to mitigate fatigue and fatigue related risks in Emergency Medical Services: a systematic review. *Prehosp Emerg Care*. 2018;22(S1):28–36.
 16. Temple JL, Hostler D, Martin-Gill C, Moore CG, Weiss PM, Sequeira DJ, Condle JP, Lang ES, Higgins JS, Patterson PD. A systematic review and meta-analysis of the effects of caffeine in fatigued workers: implications for Emergency Medical Services personnel. *Prehosp Emerg Care*. 2018;22(S1):37–46.
 17. Martin-Gill C, Barger LK, Moore CG, Higgins JS, Teasley EM, Weiss PM, Condle JP, Flickinger KL, Coppler PJ, Sequeira DJ, et al. Effects of napping during shift work on sleepiness and performance in Emergency Medical Services personnel and similar shift workers: a systematic review and meta-analysis. *Prehosp Emerg Care*. 2018;22(S1):47–57.
 18. Barger LK, Runyon MS, Renn ML, Moore CG, Weiss PM, Condle JP, Flickinger KL, Divecha AA, Coppler PJ, Sequeira DJ, et al. Effect of fatigue training on safety, fatigue, and sleep in Emergency Medical Services personnel and other shift workers: a systematic review and meta-analysis. *Prehosp Emerg Care*. 2018;22(S1):58–68.
 19. James FO, Waggoner LB, Weiss PM, Patterson PD, Higgins JS, Lang ES, Van Dongen HPA. Does implementation of biomathematical models mitigate fatigue and fatigue related risks in Emergency Medical Services operations? A systematic review. *Prehosp Emerg Care*. 2018;22(S1):69–80.
 20. Studnek JR, Infinger A, Renn ML, Weiss PM, Condle JP, Flickinger KL, Kroemer AJ, Curtis BR, Xun X, Divecha AA, et al. Effect of task load interventions on fatigue in Emergency Medical Services personnel and other shift workers: a systematic review. *Prehosp Emerg Care*. 2018;22(S1):81–88.
 21. Smith V, Devane D, Begley CM, Clarke M. Methodology in conducting a systematic review of systematic reviews of health-care interventions. *BMC Med Res Methodol*. 2011;11(1):15. <https://doi.org/10.1186/1471-2288-11-15> PMID:21291558.
 22. International Agency for Research on Cancer. IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Vol. 98: Painting, Firefighting, and Shiftwork. Lyon, France: World Health Organization. 2010:563.
 23. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, Alonso-Coello P, Falck-Yitter Y, Jaeschke R, Vist G, et al. GRADE guidelines: 8. Rating the quality of evidence—indirectness. *J Clin Epidemiol*. 2011;64(12):1303–10. <https://doi.org/10.1016/j.jclinepi.2011.04.014> PMID:21802903.
 24. Horton J, Vandermeer B, Harling L, Tjosvold L, Klassen TP, Buscemi N. Systematic review data extraction: cross-sectional study showed that experience did not increase accuracy. *J Clin Epidemiol*. 2010;63(3):289–98. <https://doi.org/10.1016/j.jclinepi.2009.04.007> PMID:19683413.
 25. Wallace BC, Trikalinos TA, Lau J, Brodley C, Schmid CH. Semi-automated screening of biomedical citations for systematic reviews. *BMC Bioinformatics*. 2010;11(1):55. <https://doi.org/10.1186/1471-2105-11-55> PMID:20102628.
 26. Ng L, Pitt V, Huckvale K, Clavisi O, Turner T, Gruen R, Elliot JH. Title and Abstract Screening and Evaluation in Systematic Reviews (TASER): a pilot randomised controlled trial of title and abstract screening by medical students. *Syst Rev*. 2014;3:121. <https://doi.org/10.1186/2046-4053-3-121> PMID:25335439.
 27. Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, Montori V, Akl EA, Djulbegovic B, Falck-Yitter Y, et al. GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias). *J Clin Epidemiol*. 2011;64(4):407–15. <https://doi.org/10.1016/j.jclinepi.2010.07.017> PMID:21247734.
 28. Higgins JPT, Green S (eds). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from <http://handbook.cochrane.org>.
 29. 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. <https://doi.org/10.1186/1471-2288-3-25> PMID:14606960.
 30. Bolster L, Rourke L. The effect of restricting residents' duty hours on patient safety, resident well-being, and resident education: an updated systematic review. *J Grad Med Educ*. 2015;7(3):349–63. <https://doi.org/10.4300/JGME-D-14-00612.1> PMID:26457139.
 31. Faraone SV. Interpreting estimates of treatment effects: implications for managed care. *P T*. 2008;33(12):700–11. PMID:19750051.
 32. Guyatt GH, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, Norris S, Falck-Yitter Y, Glasziou P, DeBeer H, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol*. 2011;64(4):383–94. <https://doi.org/10.1016/j.jclinepi.2010.04.026> PMID:21195583.
 33. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA-Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol*. 2009;62(10):1006–10. <https://doi.org/10.1016/j.jclinepi.2009.06.005> PMID:19631508.
 34. Yuan Y, Hunt RH. Systematic reviews: the good, the bad, and the ugly. *Am J Gastroenterol*. 2009;104(5):1086–92. <https://doi.org/10.1038/ajg.2009.118> PMID:19417748.
 35. Reed DA, Fletcher KE, Arora VM. Systematic review: association of shift length, protected sleep time, and night float with patient care, residents' health, and education. *Ann Intern Med*. 2010;153(12):829–42. <https://doi.org/10.7326/0003-4819-153-12-201012210-00010> PMID:21173417.
 36. McNeely ML, Campbell KL, Rowe BH, Klassen TP, Mackey JR, Courneya KS. Effects of exercise on breast cancer patients and survivors: a systematic review and meta-analysis. *CMAJ*. 2006;175(1):34–41. <https://doi.org/10.1503/cmaj.051073> PMID:16818906.
 37. Jarczok MN, Jarczok M, Mauss D, Koenig J, Li J, Herr RM, Thayer JF. Autonomic nervous system activity and workplace stressors—a systematic review. *Neurosci Biobehav Rev*. 2013;37(8):1810–1823. <https://doi.org/10.1016/j.neubiorev.2013.07.004>.
 38. Garipey G, Nitka D, Schmitz N. The association between obesity and anxiety disorders in the population: a systematic review and meta-analysis. *Int J Obes (Lond)*. 2010;34(3):407–19. <https://doi.org/10.1038/ijo.2009.252> PMID:19997072.
 39. Buscemi N, Harling L, Vandermeer B, Tjosvold L, Klassen TP. Single data extraction generated more errors than double data extraction in systematic reviews. *J Clin Epidemiol*. 2006;59(7):697–703. <https://doi.org/10.1016/j.jclinepi.2005.11.010> PMID:16765272.