International Journal of Industrial Ergonomics 108 (2025) 103777

Contents lists available at ScienceDirect

International Journal of Industrial Ergonomics

journal homepage: www.elsevier.com/locate/ergon

Developing workload-informed crew configuration recommendations for emergency medical services

Setareh Darvishi [a,*], Paul Misasi [b,c], Laila Cure [a]

[a] Industrial, Systems and Manufacturing Engineering, Wichita State University, 1845 Fairmount St, Wichita, KS, 67220, USA
[b] Sedgwick County EMS, 1015 Stillwell, Wichita, KS, 67213, USA
[c] Kansas College of Osteopathic Medicine, USA

ARTICLE INFO

Keywords:
Emergency medical services (EMS)
Workload analysis
Crew configuration

ABSTRACT

While crew configuration in primary care settings has been studied in terms of its impact on patient outcomes, less is known about how it influences the members' workload experience. This study investigates the workload implications of crew configuration based on members' certification in emergency medical services (EMS). Advanced life support (ALS) ambulance crews are commonly comprised of two paramedics (homogeneous crew) or an emergency medical technician (EMT) and a paramedic (heterogeneous crew). The goals of this study were the following: (1) to investigate differences in workload among members of the same crew, and (2) to use workload assessments to inform crew configuration strategies. We mapped one year of an EMS system's dispatch data to members' workload estimates using the visual, auditory, cognitive, and psychomotor (VACP) approach. We found that lead members (lead paramedics) experience higher workload levels compared to support members (support paramedics or EMTs) in both types of crews. Neither configuration had a consistently lower workload than the other, but differences varied for different shifts and stations. These results informed crew configuration recommendations for stations and shifts in the collaborating system, and in terms of more generalizable variables. A minimum number of staffed crews, half-half shift type (covering both day and night hours), and 30-day frequency of calls with priority P7 most significantly impacted the recommended crew configurations.

1. Introduction

Emergency medical services (EMS) clinicians work in uncontrolled settings that are dynamic, unpredictable, physically demanding, exposed to environmental elements, volatile, and fraught with irreducible uncertainty (Misasi, 2024). They often work long hours during which they may also be exposed to traumatic events and personal safety is not guaranteed (Dropkin et al., 2019). In Israel, one study found that of 533 paramedics responding to an online questionnaire, 73 % left after five years and 93 % after ten years of their training. Work conditions, including extensive physically demanding tasks, inadequate compensation, and long working hours were the main factors in their decisions to quit (Dopelt et al., 2019). In the U.S., 7.9 % of paramedics reported that they were likely to leave EMS within 12 months (Gage et al., 2024). Stress or burnout was reported as the most significant reason for leaving (Gage et al., 2024). High workload has been identified as a contributing factor to fatigue among paramedics (Paterson et al., 2014), and fatigue is recognized as a core dimension of burnout (Melamed et al., 1992; Hammarström et al., 2023).

Several studies have examined fatigue-mitigating strategies for EMS clinicians, often involving specific training programs. These programs covered topics such as sleep education (Patterson et al., 2023), the strategic use of naps to mitigate sleep loss (Patterson et al., 2020), the use of stimulants like caffeine (Barger et al., 2018), and mindfulness practices to reduce fatigue (Ducar et al., 2020; Othman et al., 2023). While these strategies aim at helping individuals better cope with high levels of workload and potentially avoid fatigue, they place the responsibility on individual members to maintain their own well-being. In contrast, operational strategies for balancing workloads focus on system changes that reduce the burden on individuals. Modifying work conditions, such as moderating job demands, has the potential to address the root causes of stress-related health issues (Lovejoy et al., 2021). In an EMS context, operational strategies may include station location, dispatching, redeployment, and crew configuration, among others.

This research focused on the operational strategy of workload-informed crew configurations. We distinguished between the terms

* Corresponding author.
E-mail addresses: Sxdarvishi@shockers.wichita.edu (S. Darvishi), Paul.Misasi@sedgwick.gov (P. Misasi), laila.cure@wichita.edu (L. Cure).

https://doi.org/10.1016/j.ergon.2025.103777
Received 11 November 2024; Received in revised form 28 April 2025; Accepted 29 June 2025
Available online 3 July 2025

crew and member. The term crew refers to a team formed by two members, while a member refers to an individual within that team. Advanced life support (ALS) crews may include two paramedics (or paramedic-paramedic [PP]), or one paramedic with one emergency medical technician (EMT). Crew configurations that involve clinicians with different professional levels within the same crew, such as paramedic-EMT, have become increasingly common due to clinician shortages (Zavadasky). EMTs require no prior medical background and undergo shorter training periods than paramedics (The Difference between a paramedic). Additionally, in 2023, the median annual wage of EMTs was significantly lower than that of paramedics (U.S. Bureau of Labor Statistics, 2024). These factors have contributed to the increasing use of EMTs alongside paramedics in emergency crews. As a result, heterogenous crew configurations present an opportunity to study their potential impact on members' workloads. The remainder of this section explores the existing literature on the impacts of including EMTs in crews, as well as how these heterogenous team configurations influence members' workloads.

The impact of EMS crew configurations on performance has been studied in terms of various clinical metrics. For example, Cortez et al. (2017) compared PP with paramedic-EMT crews in terms of on-scene times, protocol violations, and key clinical outcomes such as time-to-electrocardiogram, time-to-intravenous (IV) insertion, and IV success. Bayley et al. (2008) studied PP and paramedic-EMT crews in terms of cardiac arrest resuscitation performance. Fang et al. (2020) evaluated crew configurations over a three-year period on out-of-hospital cardiac arrest (OHCA), finding that a higher ratio of paramedics to total EMTs on the scene was associated with better patient outcomes. This was consistent with a study by Sun et al. (2018). The findings of these studies (Cortez et al., 2017; Bayley et al., 2008; Fang et al., 2020; Sun et al., 2018) support the fact that paramedic-EMT crews do not compromise performance in terms of care quality and can function effectively as an alternative to PP crews. Interestingly, a study by Sun et al. (2018) highlighted the fact that the optimal survival of OHCA cases was achieved with a balanced combination of EMTs and paramedics. Yet little is known about the extent to which this configuration contributes directly or indirectly to the workload experienced by members.

Team configuration has been examined in other sectors based on characteristics such as members' sex, skills, and professional certification (Abrams et al., 2024). For instance, a study in car manufacturing found that sex-diverse teams could positively influence absenteeism and performance. Teams with at least one female member experienced a 15 % reduction in absenteeism and a 30 % decrease in error rates (Fritzsche et al., 2014). In the context of primary healthcare, team configuration studies mostly focused on its impact on patient outcomes (Vleminckx et al., 2024). For instance, patients in practices with a predominance of female physicians reported better responsiveness of care (Pineault et al., 2017). Burnout was the outcome most closely related to members' workload that has been studied in team configuration research. Studies have shown that female physicians report higher burnout rates than their male counterparts (Dai et al., 2020; Bruhl et al., 2020). One study investigated the relationship between the emotional exhaustion component of burnout and team configuration in a Midwestern primary care practice, analyzing data from 420 family medicine clinicians across 59 communities. Unlike other studies that often group clinicians together, this study distinguished between physicians and nurse practitioners/physician assistants (NPs/PAs) and assessed their experiences separately to better understand how team configuration influenced their burnout experiences. The findings indicated that in heterogeneous teams consisting of both physicians and NPs/PAs, a higher proportion of physicians was associated with lower levels of burnout (Bruhl et al., 2020). As evidenced by the studies reviewed, research on this topic remains scarce, especially regarding the impact of education or professional certification within teams. Our study investigated this gap in the context of EMS.

This research used emergency dispatch data commonly available to EMS systems to identify significant differences in members' workloads considering team configurations based on their professional certifications. We define a "homogeneous" crew as an ambulance team consisting of two paramedics, and a "heterogeneous" crew as consisting of one paramedic and one EMT. We hypothesized the following:

The workload experienced throughout a shift by members in heterogenous crews is significantly higher than the workload of members of homogeneous crews.

To study this hypothesis, we formulated the following research questions.

(RQ1) How do workload estimates differ between members in heterogeneous and homogeneous teams during the shift?
(RQ2) How do workload estimates differ between members with different roles in heterogeneous and homogeneous teams?
(RQ3) What factors influence the recommendation of homogeneous crews instead of heterogenous crews?

We investigated these research questions at the overall level, and then stratified by station assignment, shift, and leadership role assumed by the members. The results of the analysis were used to recommend team configurations for the collaborating EMS system and to explore generalizable variables that can be used by other EMS systems in guiding their team-configuration decisions.

2. Methods

2.1. Study setting

Sedgwick County EMS (SCEMS) is the exclusive provider of ambulance services (including emergency and interfacility transfer calls) for the residents of Sedgwick County, Kansas, which includes the city of Wichita and its surrounding communities. Sedgwick County comprises 1008 square miles with a population of approximately 523,828, 76 % of whom reside in the city of Wichita, which is 166.5 square miles (U.S. Census Bureau, 2020). As a result, SCEMS responds within areas classified as urban, suburban, rural, and frontier. Call volume is historically concentrated geographically on the more densely populated and lower socioeconomic areas, which is primarily centralized in downtown Wichita. Call volume also varies by time of day, with the fewest calls dispatched in the early morning hours, between 2:00 and 4:00 a.m., and the most dispatched between 12:00 and 5:00 p.m. Because the call volume follows predictable patterns by hour-of-day, SCEMS deploys a minimum of 13 ambulances and then reaches peak staffing at around 3:00 p.m. with 21 ambulances. SCEMS maintains 20 geographically distributed stations.

SCEMS employs 155 field staff personnel, including paramedics (87 %) and EMTs (13 %). SCEMS maintains a full ALS service by assuring a minimum of one paramedic in all ambulances and response vehicles. The agency also collaborates with other first response services in the area, including fire departments and law enforcement agencies and, in some rural communities, provides its own first response by way of a single paramedic in a utility vehicle. Additionally, there are two or three free-roaming field supervisors, depending on time of day, who monitor all calls, assist with resource deployment, self-dispatch to assist on critical calls, and perform some additional personnel management duties, including scheduling. In this system, each ambulance team is referred to as a "crew" and is staffed with two members. The crew is assigned to the same ambulance and the same geographical location (i.e., station) for the full shift; however, all calls are dispatched to the closest ambulance by driving time.

While there are several starting times for shifts (6:00 a.m., 9:00 a.m., 12:00 p.m., and 6:00 p.m.), all shifts have a duration of 12 h. When crews are homogeneous, the most senior and ranking paramedic will assume the crew leader role for all activities during the shift; however, the crew will alternate the primary patient care lead for calls during the shift to balance workload between members. If the crew is heterogeneous, then the primary care lead depends on the acuity of the patient conditions they encounter; the lower-level acuity patients are led by the EMT, whereas higher levels of assessment and treatment are led by the paramedic. The paramedic may elect to lead lower-level acuity calls to balance out the workload of the EMT, if the calls to which they are assigned during their shift all happen to be lower acuity. The converse is not true, however, as an EMT cannot lead higher-acuity calls. Calls with priorities P1, P2, P3, P4, P5, and P7 are considered emergency calls, with lower numbers corresponding to emergency medical dispatch determinants designated as more urgent and requiring more resources. Table 1 summarizes priority levels and their corresponding responses at SCEMS [Table 1 near here.].

2.2. Workload estimates

Different methods have been used to assess workload in the EMS literature, such as clinicians' physiologic parameters including heart rate (HR) and metabolic debt (Asselin et al., 2018), Borg scale (Borg, 1990), and the National Aeronautics and Space Administration (NASA) Task Load Index (NASA-TLX) tool (Hart and Staveland, 1988), as well as call volume (He et al., 2019), cumulative worktime (Huynh et al., 2020), and the visual, auditory, cognitive, psychomotor (VACP) score (Zhang et al., 2023), among others (He et al., 2019; Thielmann et al., 2022). While call volume, the number of calls that crews respond to during a period, provides a simple measure of workload, it may not account for variations in call complexity or the time required to complete each call (He et al., 2019). Alternatively, cumulative work time represents the total amount of time that crews work during a given period. This method accounts for the overall time spent on different activities; however, it does not reflect variations in workload intensity imposed by those tasks during a shift (Huynh et al., 2020). The VACP scores are used to assess the demands placed on an individual's resources across four functional domains: visual, auditory, cognitive, and psychomotor (Zhang et al., 2023; McCracken and Aldrich, 1984). It uses a standardized 7-point scale with verbal anchors to promote consistency and minimize inter-rater variability. For example, the cognitive demands for a driving task for a beginner driver may be as high as 6.8 out of 7.0, corresponding to the verbal anchor "evaluation/judgment: consider several actions" (McCracken and Aldrich, 1984). We assessed the workload using the VACP scores approach because these allow for the use of large amounts of dispatch data commonly available to EMS systems and incorporate both the time spent performing work and a measure of the mental effort required in performing these tasks (Rusnock et al., 2015). Furthermore, VACP workload assessments have been shown to have a moderate correlation with NASA-TLX estimates in an EMS context (Ercolani et al., 2024). The workload associated with past call response patterns was assessed by labeling timestamps in 9-1-1 dispatch data with VACP workload scores following the process illustrated in Fig. 1 [Fig. 1 near here].

Table 1
Description of priority levels at Sedgwick County EMS.

Response Priority	Dispatch Urgency	Fire Response	EMS Response
P0	Waiting to Dispatch (EMS in Progress)	–	–
P1	Immediate	Fast	Fast
P2	Immediate	Slow	Fast
P3	Immediate	Fast	Slow
P4	Immediate	Slow	Slow (No Transfers)
P5	Immediate (Request for EMS, Transfer)	No Response	Slow
P6	Immediate	Slow	No response
P7	Immediate (Emergency Transfer)	No Response	Crew Discretion
P8	Can Hold (Transfer)	No Response	Slow
P9	Can Hold (Specialty Care Transfer)	No Response	Slow

Slow: no lights or sirens; fast: lights and sirens.

2.2.1. 9-1-1 dispatch data

One calendar year's worth of 9-1-1 dispatch data from January to December consisting of call records from both emergency and non-emergency calls were provided by the collaborating EMS system. The data contained time stamps of direct call response tasks for all crews and shifts, including dispatch time, driving to the scene (enroute), being on scene, leaving scene, arriving at hospital, transferring care, and back to service time (available). Durations of these activities during each shift were obtained directly from these data. The dispatch data also contained details like patients' acuity level, call priority, leading member, and professional certification of members. This study focused on shifts involving ambulances with two members and included all calls (emergency and non-emergency).

2.2.2. Task analysis

An independent task analysis was done to identify tasks commonly performed by EMS members during their shifts including those that were not recorded in the dispatch data. The task analysis aimed at measuring the durations of these tasks as well as assigning VACP scores. The research team included three researchers, in collaboration with a faculty member and a team of EMS experts. They performed direct observations of EMS crews at randomly selected locations, times of day, and durations ranging from 4 to 13 h. In total, 220 h of observation were conducted, covering 80 EMS calls. The team identified that the overall work content of EMS crewmembers included the following tasks: starting a shift, navigating to the scene, driving to the scene, moving to the patient, providing a preliminary checkup, moving to the ambulance, providing an ambulance checkup, driving to the hospital, checking the patient enroute to the hospital, unloading the patient, relaying information to the doctor, cleaning up the stretcher and vehicle, driving back to the station, documenting the incident, and fueling with gas. These tasks were distributed between the two members. Some tasks were performed by member 1 (leading member), some by member 2 (support member). For example, it was observed that while member 1 was navigating to the scene, member 2 was tasked with driving. Similarly, while member 1 was driving to the hospital, member 2 was checking the patient enroute to the hospital. Some other tasks were performed collaboratively by both members, such as moving to the patient, moving to the ambulance, and unloading the patient. The tasks of each member during a call were identified using dispatch records, which noted information such as who drove to the scene, who drove to the hospital, and who completed the documentation. Both durations and VACP scores of each task were used to estimate the corresponding member's overall workload.

Timestamps associated with these tasks, which are recorded in the dispatch data, were used to estimate their actual durations in a shift. The durations of shift start, documentation, gas fueling, and driving back to the station, which are not commonly recorded in the dispatch data, were estimated from distributions assessed through time studies. The duration of driving back to the station was estimated, assuming normally distributed travel times with means obtained from the Google Maps Application Programming Interface (API) and an arbitrary standard deviation of 2 min (Google Maps API driving time estimates were consistent for different times of day and days of the week for commonly visited locations in Sedgwick County).

The VACP values were assigned by observers to individual tasks according to the scale adapted from Rusnock and Borghetti (2018), shown in Table 2. The research team selected the final workload scores for the tasks by consensus, considering variations in the nature of the entire task and its subtasks. For example, the task of navigating to the scene was scored as a whole with a consistent workload, as shown in the

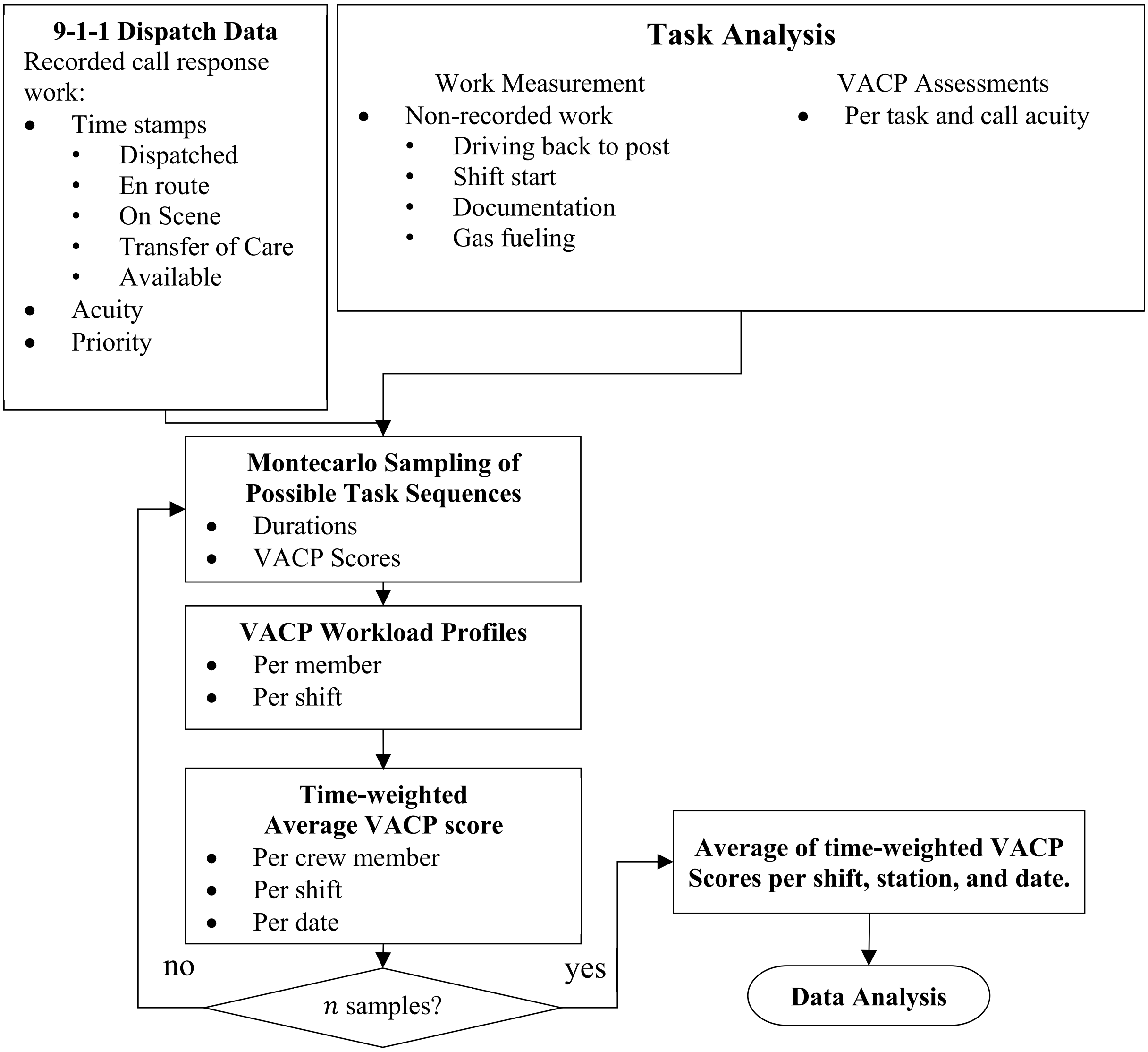

Fig. 1. Illustration of the process of labeling 9-1-1 dispatch data with VACP scores and obtaining workload measure used in analysis, i.e., average of time-weighted VACP scores from n Montecarlo samples of dispatch patterns.

Appendix. Meanwhile, the team identified three workload subcategories for the task of moving to the patient as a whole: walking with light equipment, walking with a stretcher, or involving stairs or heavy equipment. VACP scores were assigned to these workload subcategories, and the subcategories were mapped to dispatch data attributes to facilitate the Montecarlo analysis. For instance, if a call recorded in the dispatch dataset had acuity “green” or “orange,” then crews most likely walked to the patient with light equipment (which was assigned a probability of 95 %, leaving a 5 % probability of having walked to the patient with a stretcher). The online Appendix contains a list of tasks with workload subcategories, along with their VACP scores and the mapping to dispatch data attributes [Table 2 near here.].

Some of the non-recorded tasks also involved subtasks with different workload levels, such as shift start and documentation. VACP scores were initially assigned to the subtasks following the same observation and consensus approach described above. An overall VACP score for the task was then calculated by using a time-weighted average of the VACP scores of all subtasks. Table 3 presents the non-recorded tasks along with their corresponding subtasks and assigned VACP scores [Table 3 near here.].

2.2.3. *Montecarlo sampling*

Recorded events and timestamps were extracted for each shift available in the 9-1-1 dispatch data. Pseudorandom durations for the non-recorded tasks (shift start, documentation, gas fueling, and driving back to the station) were sampled from the duration distributions established in the task analysis. Then, the sampled timestamps were added to the observed dispatch patterns to approximate the time that the member spent performing the work. The assessed VACP scores were added to the timestamps of the corresponding tasks. This data pre-processing task was coded in Python software. Each Montecarlo sample of a station on a given shift and date represents one possible realization of the VACP score over time (workload profile) for that shift. Fig. 2 illustrates how the overall work content, and resulting workload profile, is estimated from task sequences derived from 9-1-1 dispatch data for two Montecarlo samples. In this example, an EMS crew starts their shift at 6:00 a.m., when they start performing the shift start task (time-weighted VACP = 16.30). They are dispatched to a call with priority P5 and green acuity at 6:08 a.m. The enroute task includes driving to the scene with the workload subcategory “Without Lights and Sirens,” because emergency driving is not required (time-weighted VACP = 23.60). On scene tasks (time-weighted VACP = 22.64) include moving to the patient with subcategory “Walking, Using Light Equipment,” the

Table 2
VACP scales (adapted from (Rusnock and Borghetti, 2018)).

Value	Descriptors
Visual	
0.0	No Visual Activity
1.0	Visually Register/Detect (detect occurrence of image)
3.0	Visually Inspect/Check (discrete inspection/static condition)
4.0	Visually Locate/Align (selective orientation)
4.4	Visually Track/Follow (maintain orientation)
5.0	Visually Discriminate (detect visual difference)
5.1	Visually Read (symbol)
6.0	Visually Scan/Search/Monitor (continuous/serial inspection, multiple conditions)
Auditory	
0.0	No Auditory Activity
1.0	Detect/Register Sound (detect occurrence of sound).
2.0	Orient to Sound (general orientation/attention)
3.0	Interpret Semantic Content (speech, simple, 1e2 words)
4.2	Orient to Sound (selective orientation/attention)
4.3	Verify Auditory Feedback (detect occurrence of anticipated sound)
6.0	Interpret Semantic Content (speech, complex, sentence)
6.6	Discriminate Sound Characteristics (detect auditory differences)
7.0	Interpret Sound Patterns (pulse rates, etc.)
Cognitive	
0.0	No Cognitive Activity
1.0	Automatic (simple association)
1.2	Alternative Selection
4.6	Evaluation/Judgment (consider single aspect)
5.0	Sign/Signal Recognition
5.3	Encoding/Decoding, Recall
6.8	Evaluation/Judgment (consider several aspects)
7.0	Estimation, Calculation, Conversion
Fine Motor	
0.0	No Fine Motor Activity
2.2	Discrete Actuation (button, toggle, trigger)
2.6	Continuous Adjustive (flight controls, sensor control)
4.6	Manipulative (tracking)
5.5	Discrete Adjustment (rotary, vertical thumbwheel, lever position)
6.5	Symbolic Production (writing)
7.0	Serial Discrete Manipulation (keyboard entries)
Gross Motor	
0.0	No Gross Motor Activity
1.0	Walking on level terrain
2.0	Walking on uneven terrain
3.0	Jogging on level terrain
3.5	Heavy lifting
5.0	Jogging on uneven terrain
6.0	Complex climbing
Speech	
0.0	No speech activity
2.0	Simple (1–2 words)
4.0	Complex (sentence)
Tactile	
0.0	No tactile activity
1.0	Alerting
2.0	Simple discrimination
4.0	Complex symbolic information

preliminary checkup, ambulance checkup, and checkup enroute all classified under the "Light" workload subcategory, moving to the ambulance and unloading the patient categorized as "Patient Walking," and stretcher and vehicle cleanup with subcategory "Reloading the Stretcher." While the recorded durations of enroute, on scene, and transfer of care are the same for both Montecarlo samples, the estimated (sampled) durations of non-recorded tasks (marked with an asterisk *) are different for the two samples shown, resulting in slightly different workload profiles and time-weighted VACPs [Fig. 2 near here.].

This study used 30 Montecarlo samples for each station-shift-date combination to estimate the mean of the time-weighted VACP scores, which is the main workload measure used in subsequent analyses. The sample size was chosen based on an acceptable absolute error margin of ±1.15 at a 95 % confidence level (Law, 2015). In the remainder of this article, we use the acronym VACP to refer to the average, over the 30 Montecarlo samples, of the time-weighted VACP scores by the end of the corresponding shift, unless specified otherwise.

2.3. Data analysis

The resulting research dataset had records representing unique date, station, and shift combinations, along with variables including type of "shift" (nominal variable with values 6:00 a.m., 9:00 a.m., 12:00 p.m., and 6:00 p.m.); role of each member (lead or support) as well as VACP scores for them (quantitative variables representing the average time-weighted VACP score by the end of the shift); and "Homogenous Crew" (binary variable with a value of 1 denoting a crew with two paramedics, and 0 otherwise). We used a two-sample *t*-test (Leventhal, 2016) to study the differences between members' mean VACPs in homogenous and heterogenous teams overall (RQ1). We inspected the workload difference in the different types of crews in more detail by comparing the VACP of members with different professional certifications within either team (RQ2). We also employed the two-tailed form of the same test to investigate differences in the mean VACPs between heterogeneous and homogeneous crews based on stations and shifts.

Final crew configuration recommendations were developed according to the results of those t-tests as follows. Stations and shifts for which heterogeneous crews had statistically larger VACP scores than homogenous crews were recommended to be staffed with homogeneous crews. Similarly, heterogeneous crews were recommended for stations and shifts where homogeneous crews had significantly larger VACP scores based on the dispatch data. To complement statistically insufficient recommendations, we performed a "what-if" analysis of the dispatch patterns of past heterogeneous crews assuming an "alternating lead" dynamic consistent with a homogeneous crew configuration. We compared the VACP scores between actual (heterogeneous) and simulated (theoretically homogeneous) work patterns. Stations and shifts for which the workload of heterogenous crews were statistically larger than the corresponding workload under the simulated alternating lead dynamic (with 95 % confidence) were recommended as requiring homogenous crews, while others were marked as candidates for heterogeneous configuration.

To investigate the operational factors associated with the recommendation of homogeneous crews (RQ3), we employed logistic regression analysis (Peng et al., 2002). We added the recommendations obtained in the previous analyses as a new binary variable of the research dataset (i.e., "homogenous crew recommendation" with a value of 1 if a homogenous crew was recommended to that station and shift, and 0 otherwise). We also added generalizable potential predictors of these recommendations, such as the type of shift (shifts at 6:00 a.m. and 9:00 a.m. were classified as day shifts, the 6:00 p.m. shift was designated as a night shift, and the 12:00 p.m. shift fell into the half-half category), the minimum number of crews staffed regardless of the configuration over the course of the shift, where each crew referred to a full response team (not an individual member), and the moving frequency of calls per priority level and per initial acuity level. The rolling time window considered for the moving frequencies was 30 days for each shift-date combination. The newly labeled datasets were divided into training (80 %) and testing (20 %).

This study was reviewed and approved by the Wichita State University Institutional Review Board (IRB #4505).

3. Results

3.1. Data overview

The 9-1-1 dispatch data spanned one calendar year between January and December of 2022. During that period, ambulances were dispatched to a total of 67,681 calls. A total of 60,833 (89.9 %) were emergency calls, and 6848 (10.1 %) were non-emergency calls, such as interfacility transfers.

Table 3
Description of non-recorded tasks, Fitted distribution along with parameters, and VACP scores.

		VACP						
Task Sequence	Fitted Distribution (Parameters)	Visual	Auditory	Cognitive	Fine Motor	Gross Motor	Speech	Tactile
Shift Start	*Normal* (18.64, 9.18)							
Check in with Previous Crew		1	6	5.3	0	0	4	0
Log onto Truck		5.1	1	4.6	7	1	0	0
Narcotics Check		5	1	4.6	2.2	2	0	0
Supply Check		5	1	4.6	2.2	2	0	0
Physical Truck Check		5	1	4.6	2.2	2.5	0	0
Ninth Brain Logs (Control substance and mechanical check)		5.1	1	5.3	7	0	0	0
Documentation	*Gamma* (1.54, 13.11)							
On Scene		6	6	5	7	0	2	0
On Ambulance		6	6	5	7	0	2	0
At Station		6	1	5.3	6	0	0	0
Gas Fueling	*Gamma* (11.70, 2.93)	4.4	1	1.2	2.2	1	0	0
Driving Back to the Station	*Gamma* (5.02, 0.45)	6	4.2	5	2.6	1	0	1

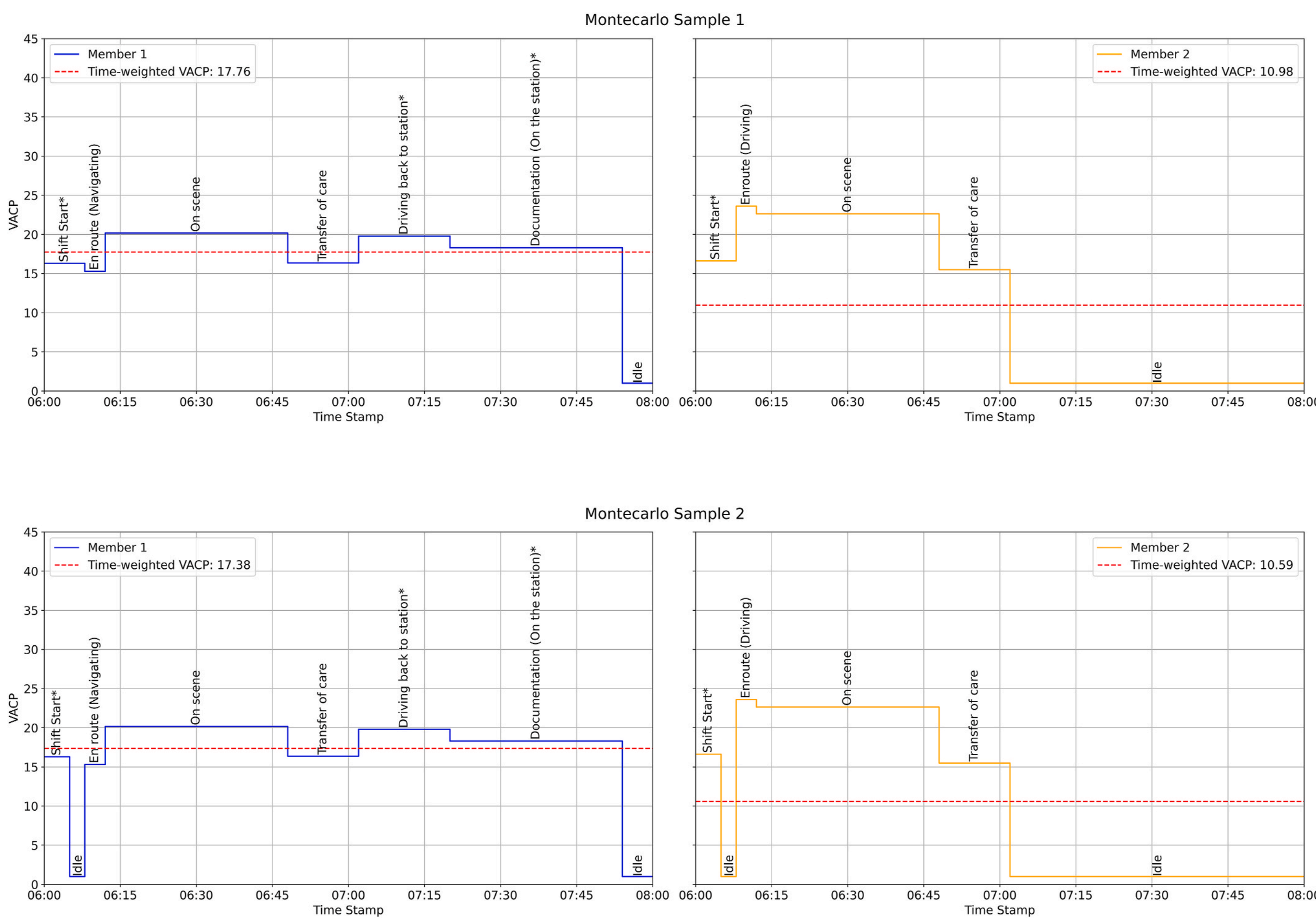

Fig. 2. Illustration of breakdown of task sequences, and VACP workload profiles for two Montecarlo samples.

Of the total calls dispatched, 43,547 (64.3 %) required transport, and 31,819 (47 %) required the use of lights and sirens during the response (calls with priorities P1 and P2). Homogeneous crews responded to 27,608 of the calls (40.8 %) and heterogeneous crews to 29,714 of the calls (43.9 %). The remaining calls (15.3 %) included calls with missing values in one or more of the member licenses, so the configurations of the dispatched crew were not identifiable. Each crew responded to an average of seven calls per shift.

3.2. *Workload*

Fig. 3 illustrates the progression of the VACP of eight different crews on one run of an arbitrary shift on a random day. As shown, the VACP can accumulate at different rates for different crews in the same shift. For example, the lead member of crew J had an estimated VACP of 17.44 at the end of the shift on that day, which was almost twice as much as the workload of the lead member of crew I (8.98). This figure also shows that there may be high variability in the workload incurred by members assigned to the same crew. For example, while crew L had both members incur similar VACPs by the end of the shift, the members in

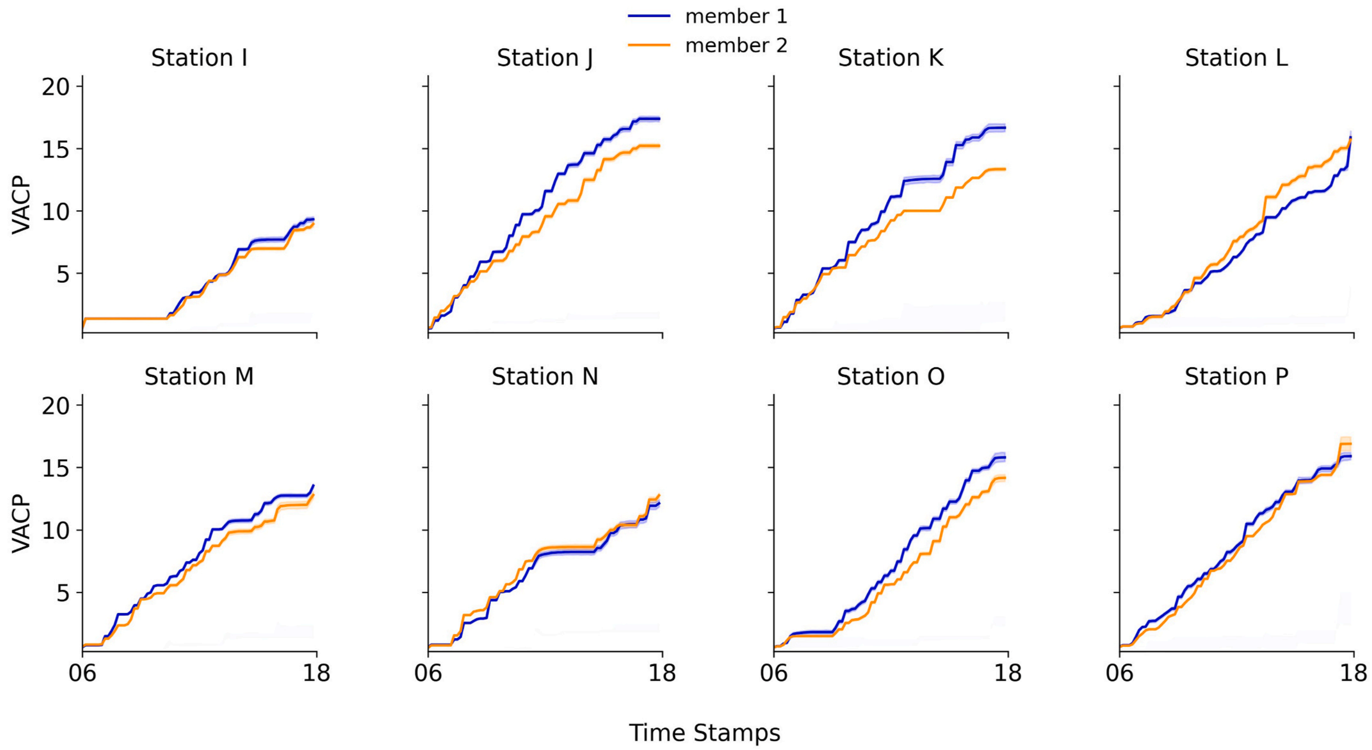

Fig. 3. Sample of time-weighted workload profiles of members in arbitrary 6:00 a.m. shift.

crew K had a 2.50 difference in the VACP they incurred (lead workload = 15.58, support workload = 13.08). A 2.50 difference represents a discrepancy of about 25 % of the observed workload range for that shift and day among all members (between 8.98 and 18.90) [Fig. 3 near here.].

3.3. Workload in homogeneous vs. heterogeneous crews

Fig. 4 shows that, overall, members in homogeneous crews had statistically higher VACP than those in heterogeneous crews ($t = 12.01$, $p < 0.05$). Fig. 5 compares the workload of paramedics and EMTs in the different types of crews. Diamonds show the means and are connected through a dashed line. Lead members in both types of crews had significantly higher workload compared to their support team members

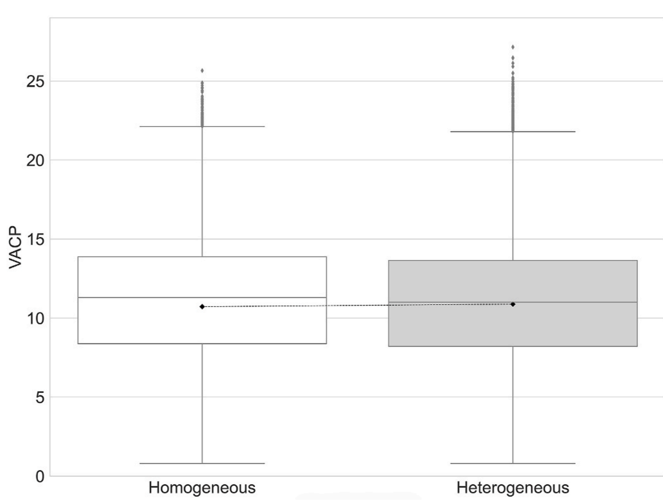

Fig. 4. Comparison of members' VACPs in homogeneous and heterogeneous crews.

($HM : t = 5.65, p < 0.05$ $HT : t = 140.53, p < 0.05$). These differences appear less pronounced in homogenous teams than in heterogeneous teams.

Table 4 shows the results of the t-tests evaluating if the VACPs of members in homogeneous crews were higher than those in heterogenous crews given historical assignments. Members' workloads showed significant differences in 18 of the 32 station-shift combinations considered (about 56 %). Furthermore, some of these statistical differences (8 of the significant 18) had negative t-values suggesting that homogeneous teams may experience higher workloads in those stations and shifts, which contradicts our main hypothesis. The results of the what-if analysis estimating the VACP of members assuming an alternating-lead task pattern are also shown in Table 4. Most station-shift combinations with insignificant differences based on historical task patterns showed a statistical difference at the 5 % level based on the assumed alternating-lead task pattern (e.g., stations B, I, and J [at 6:00 a.m.]; B, I, J, K, and O [at 6:00 p.m.]; and D and S [at 9:00 a.m.], as shown in Table 4). The resulting recommendations are shown in the last column of Table 4. These recommendations suggest heterogenous crews for stations L, M, N, O, and Q at the 6:00 a.m. shift, station F at 6:00 p.m., stations C and S at 9:00 a.m., and station R at 12:00 p.m., and homogeneous crews for others. However, there was no sufficient statistical evidence to prefer one crew configuration over another for four station-shift combinations, i.e., station E at 6:00 a.m.; stations L and Q at 6:00 p.m.; and station H at 12:00 p.m [Table 4 near here.].

3.4. Operational factors influencing crew configuration

Table 5 shows the results of the logistic regression investigating the impact of the type of shift, minimum number of crews staffed, and moving frequency of calls per priority level and per initial acuity level on recommending a homogeneous crew configuration. The model is statistically significant, as the p-value of the likelihood ratio test is less than 1 %. The model suggests that for every additional crew staffed, regardless of the configuration, the odds of recommending a homogenous crew decreased by almost 25 % (multiplied by $e^{-28} \approx 0.75$),

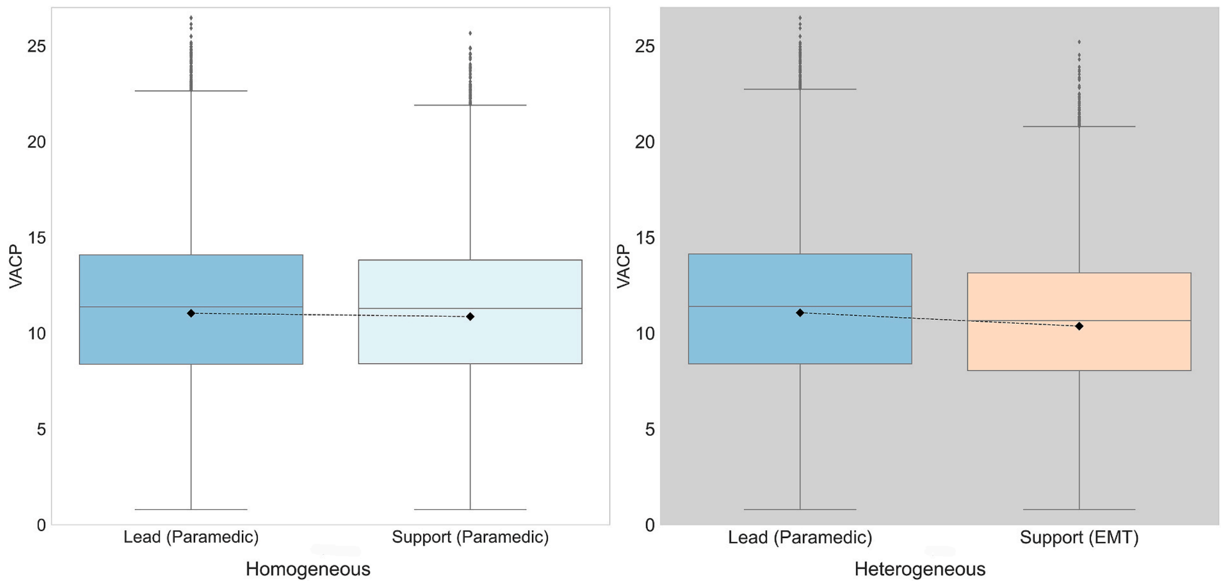

Fig. 5. Comparison of lead and support members' VACPs in homogeneous and heterogeneous crews.

Table 4
T-Test Results Comparing VACPs of Members in Homogenous and Heterogenous Crews per Station and Shift, Based on Historical Task Pattens vs. Alternating Lead Task Patterns and Crew Configuration Recommendations.

Station	Shift	Historical Assignments				HM* Work Pattern (What-If Analysis)			Resulting Recommendation
		Mean VACP		T-Statistic	P-Value	Mean VACP HM*	T-Statistic	P-Value	
		HT	HM						
A	6:00 a.m.	9.13	9.02	2.40	<0.05				HM
B		12.2	11.66	0.13	0.89	11.76	4.73	<0.05	HM
E		9.12	8.96	0.19	0.84	9.17	1.48	0.13	Not Significant
F		10.25	9.57	10.70	<0.05				HM
I		12.94	12.97	−1.73	0.08	12.53	4.49	<0.05	HM
J		12.42	12.29	1.86	0.06	12.30	3.78	<0.05	HM
K		13.05	12.5	9.11	<0.05				HM
L		13.06	13.28	−4.74	<0.05				HT
M		11.61	12.23	−11.22	<0.05				HT
N		10.48	10.66	−2.61	<0.05				HT
O		7.66	8.82	−14.00	<0.05				HT
P		10.05	10.10	6.46	<0.05				HM
Q		11.47	11.69	−4.01	<0.05				HT
A	6:00 p.m.	7.90	7.51	4.41	<0.05				HM
B		10.59	10.72	−1.34	0.18	10.46	3.85	<0.05	HM
E		8.71	8.38	7.40	<0.05				HM
F		8.84	9.44	−4.12	<0.05				HT
I		12.03	11.99	−0.06	0.95	10.12	6.72	<0.05	HM
J		11.42	11.22	1.79	0.07	9.95	6.61	<0.05	HM
K		11.99	11.96	0.38	0.70	10.81	4.73	<0.05	HM
L		12.16	11.98	1.16	0.24	11.83	1.48	0.13	Not Significant
M		11.24	10.70	8.63	<0.05				HM
N		8.95	7.95	11.38	<0.05				HM
O		7.38	7.52	1.53	0.12	7.01	3.78	<0.05	HM
P		8.23	8.27	2.59	<0.05				HM
Q		9.38	9.58	−1.01	0.31	9.32	0.63	0.53	Not Significant
C	9:00 a.m.	13.08	13.75	−11.73	<0.05				HT
D		11.23	11.27	−1.34	0.18	11.20	2.16	<0.05	HM
S		10.91	10.92	−0.92	0.35	9.98	−8.64	<0.05	HT
G	12:00 p.m.	13.24	13.00	3.56	<0.05				HM
H		10.84	11.15	1.28	0.20	11.01	0.17	0.86	Not significant
R		10.89	11.37	−7.31	<0.05				HT

"HM" denotes homogeneous crew; "HT" signifies heterogeneous crew; * indicates result from simulated distribution of workload based on equal distribution of calls among crewmembers.

assuming all other variables remain constant. It suggests that having more crews during a shift allows for more flexibility in crew configuration. All shift types are statistically significant, along with the 30-day moving frequency of calls with priority levels P1, P2, P5, P6, P7, and P9. Similarly, most 30-day moving frequencies of initial acuity levels (except for "Ongoing Resuscitation, Blue") are significant [Table 5 near here.].

A decision tree summarizing the recommendations based on the significant variables of the logistic regression model can be found in Fig. 6. The tree was built using the scikit-learn package of Python and

Table 5
Logistic regression model to identify significant variables influencing recommendation of homogeneous crew.

Variable	Coefficient	P-Value	Odds Ratio
Constant	4.92	0.00	137.80
Day of Week	−0.03	0.07	0.97
Month	0.00	0.69	1.00
Minimum Number of Staffed Crews	−0.28	0.00	0.75
30-Day Frequency of Calls with Priority			
P0	0.15	0.33	1.16
P1	0.07	0.00	1.07
P2	0.13	0.00	1.14
P3	0.01	0.09	1.01
P4	0.00	0.83	1.00
P5	0.04	0.00	1.04
P6	−0.56	0.04	0.57
P7	0.05	0.00	1.05
P8	0.00	0.66	1.00
P9	0.13	0.00	1.14
30-Day Frequency of Calls with Acuity			
Critical (Red)	−0.03	0.01	0.97
Dead without Resuscitation Efforts (Black)	0.43	0.00	1.53
Emergent (Yellow)	−0.11	0.00	0.90
Lower Acuity (Green)	−0.05	0.00	0.95
Ongoing Resuscitation (Blue)	−0.09	0.10	0.90
Psychiatric (Orange)	0.07	0.00	1.08
Shift Type: Half-Half	−1.24	0.00	0.29
Shift Type: Night Shift	−0.39	0.00	0.68
Likelihood Ratio Test	<0.01		

pruned with a cost complexity parameter of 0.008 to allow for readability while maintaining a reasonable testing accuracy (see Fig. 7). To evaluate the predictive performance, the model was trained and tested using an 80 %–20 % split. The model suggests a heterogenous crew for a station during a shift with ten staffed crews working at the same time, at a shift starting during the day and finishing at night, and a 30-day frequency of calls with P7 priority of 20. The testing accuracy of the decision tree was 0.87, while testing accuracy of the logistic regression model above (with a threshold of 0.84 obtained through the receiver operating characteristic (ROC) analysis shown in Fig. 8) was 0.68 [Figs. 6–8 near here.].

4. Discussion

This study serves as a starting point for addressing the gap in research on the impact of operational strategies on workload balance among ambulance crews. Specifically, our investigation focused on the influence of crew configurations on EMS workload. Research in various emergency medical contexts has found that crews involving members with different professional certifications do not compromise safety (Fang et al., 2020; Sun et al., 2018). This has motivated EMS systems to use heterogeneous teams to increase their capacity and decrease cost. Nevertheless, there is no consensus on one optimal, generalizable, team configuration for EMS systems (Shotwell et al., 2018). Such a standard would be difficult to establish given the complexity, variability, and uncertainty inherent to medical emergencies. We acknowledge that achieving a perfectly balanced workload across different certifications and in different team configurations may not always be feasible. The varied levels of responsibility associated with each member's certificate can impact the practical implementation of workload distribution. Still, shortages of paramedics, EMTs, and other medical professionals require EMS organizations to strategize the configuration of their crews and avoid systematic imbalances.

The main contribution of this paper is to provide insight into the impact of crew configuration on workload distribution among members in ALS crews in ambulances with at least two members. Unlike previous

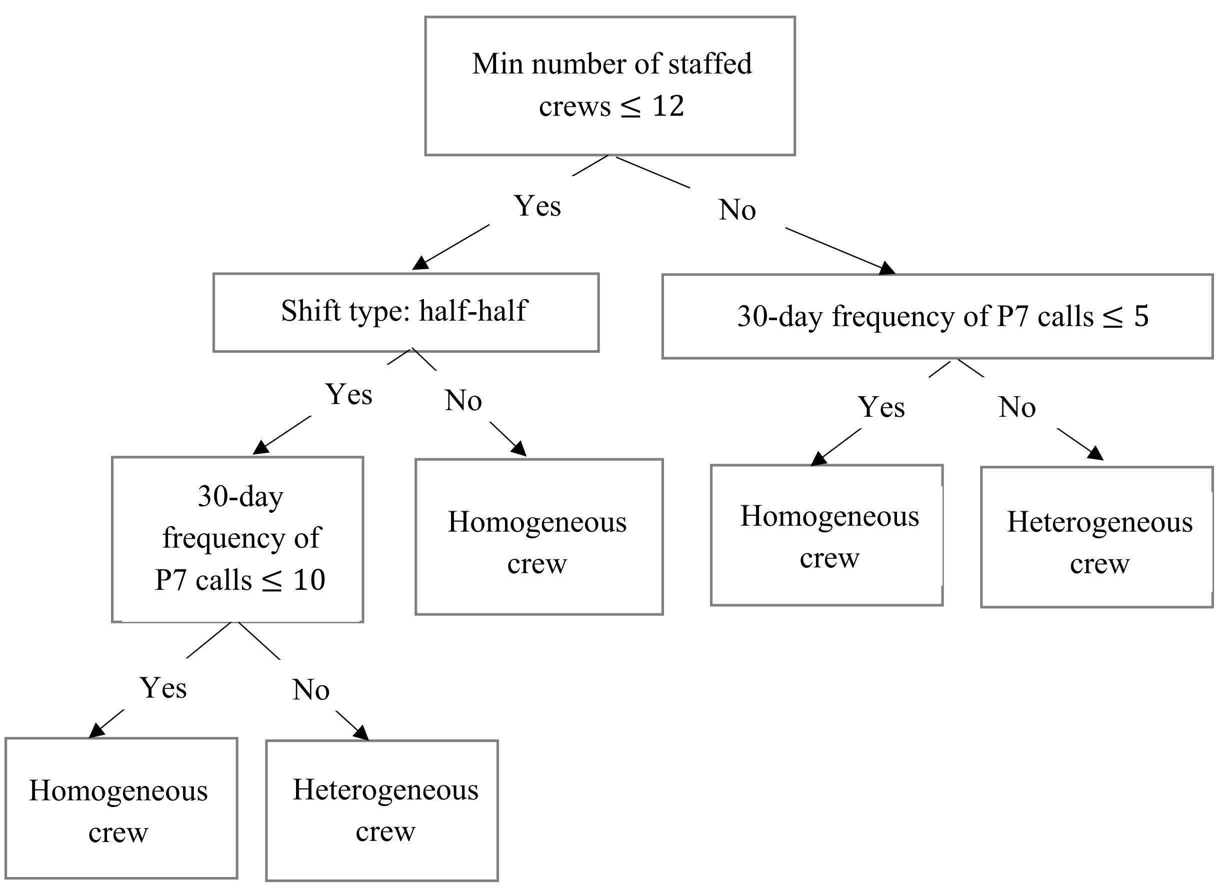

Fig. 6. Decision tree to visualize configuration recommendations procedure.

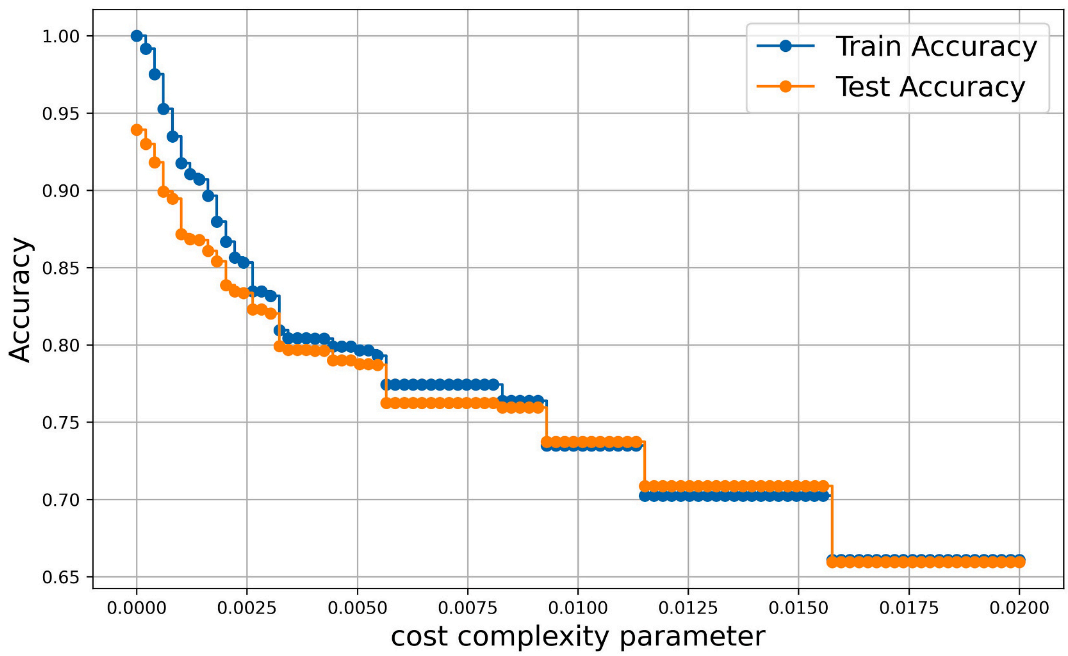

Fig. 7. Accuracy changes via different cost complexity parameters to obtain best tradeoff between tree size and accuracy.

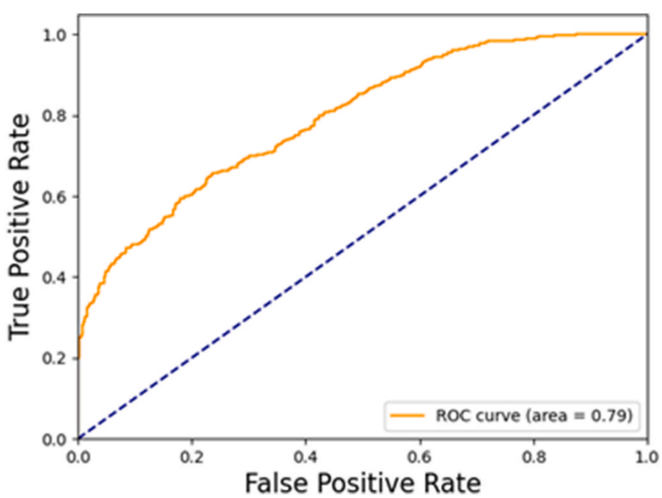

Fig. 8. ROC curve to obtain the best threshold for logistic regression model.

studies that focused on the impact of team configuration on clinical outcomes (Cortez et al., 2017; Bayley et al., 2008; Fang et al., 2020; Sun et al., 2018), our research specifically studied its impact on workloads. Our results have shown that, overall, members in homogeneous crews had statistically higher average time-weighted VACP scores than those in heterogenous crews. Thus, we rejected the hypothesis.

Our analysis showed that the aggregate results did not consistently hold for all stations or shifts in the collaborating organization. We found that both configurations could result in lower workloads for their members, depending on the locations and during specific time periods (RQ1). Furthermore, we disaggregated the data by members' roles and found that crew leaders, who are always paramedics, tended to have significantly higher time-weighed VACP scores than their support counterparts, who could be EMTs or paramedics, depending on the configuration (RQ2). This finding aligns with the work of Tofil et al. (2017), which reported that, on average, team leaders experienced significantly higher workloads than team members in a pediatric emergency context.

Similar to the work of Bruhl et al. (2020), we distinguished between paramedics, EMTs, and the leadership role they assume (lead vs. support). We evaluated their workloads individually to better understand the impact of team configuration on individual workload experiences. However, they did not explore whether homogeneous teams (e.g., all physicians) might have impacted the workload experience. In contrast, we conducted a what-if analysis to examine the potential effects of assigning homogeneous rather than heterogeneous teams. We showed how these nuances could be used to develop collaborator-specific recommendations and found that the minimum number of staffed crews, half-half shift type, and a 30-day moving frequency of calls with priority P7 significantly influence recommending a homogeneous team and result in accurate recommendations based on a classification tree analysis (RQ3). Given that these variables can be calculated directly from commonly available 9-1-1 dispatch data, we hope that other EMS systems can benefit from these results in developing their crew configuration strategies.

The research team had no access to test the model outside of SCEMS, which can be considered a limitation in the generalizability of these results. There were also limitations in crew attributes considered in this study. For example, our investigation did not include familiarity, i.e., how long members had worked together, interpersonal and team dynamics, or members' experience levels. We understand that these factors could influence workload; however, the increased rates of turnover make it less likely that strategies that incorporate these considerations into the analysis can be sustained practically. Additionally, the VACP method itself has limitations, such as subjectivity of the evaluation scale and scoring. However, this subjectivity differs from that of self-reported workload assessment tools, such as the NASA-TLX. While NASA-TLX relies on members' personal perceptions of workload (Wang et al., 2024), VACP involves structured task-level evaluations by external observers (Lu et al., 2023).

Future research may investigate whether crew configuration recommendations seeking to balance workload should vary with time or be based on updated demand forecasts, which would require several years of data. Future research might also investigate other station characteristics such as geographical location, driving distances, and types of patient population, among others requiring additional datasets. In general, future research in EMS systems should incorporate real-time workload assessments into any decision. This article focused on team configuration, but other decisions, such as dispatching and redeployment, may influence the workload experienced by members. Most of the data that would be needed for these types of studies is already embedded in most EMS systems. We showed one way in which the data can be labeled with VACP workload assessments. Thus, future research may also investigate how to label dispatch data with other assessment techniques, such as using physiological measures. The goal, regardless of the workload assessment method, should be to support strategic and real-time operational decisions that allow members to experience a fair and balanced workday.

5. Conclusions

This study provides new insights into the relationship between crew configurations and workload in EMS work, with a specific focus on the impact of homogeneity and heterogeneity in crew compositions. We tested our main hypothesis that the workload of members in heterogenous crews was significantly higher than the workload of members in homogenous crews during the same shift. We found that this does not hold consistently for all shifts and stations and that there are stations and shifts in which heterogeneous teams have a significantly lower workload. Therefore, we rejected the main hypothesis.

We found that lead members experience higher VACP scores compared to support members in both homogeneous and heterogeneous crews. The difference was more noticeable in heterogeneous crews. We were also able to identify stations and shifts specific to the collaborating EMS where the workload of heterogenous crews was statistically lower than homogeneous crews. This information allowed for specific recommendations on crew configurations that could improve workload distribution. Data for some specific shifts and stations was insufficient to recommend any configurations. Hence, we performed a what-if analysis and compared the VACP scores between actual and simulated work patterns. This analysis helped us in recommending the best configurations to as many shifts and stations as possible. We used logistic regression to find all generalizable variables impacting the members' workload and a decision tree to visualize the most significant ones and how to use them to recommend configurations to other EMS systems.

Our findings have practical implications for workload management in EMS systems. They suggest that the choice between homogeneous and heterogeneous crew configurations could be guided by factors such as the shift and the station. Moreover, there are more generalizable variables that impact the appropriate crew configurations significantly such as type of shift (day or night), minimum number of staffed crews, and 30-day frequency of call priorities and acuity levels. Some of these variables, including the minimum number of crews staffed over the shift and moving frequency of tasks, can be utilized in other EMS systems. The analysis approach presented in this article could be applied to any system involving shift teamwork.

CRediT authorship contribution statement

Setareh Darvishi: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Paul Misasi:** Writing – review & editing, Conceptualization. **Laila Cure:** Writing – review & editing, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors report there are no competing interests to declare.

Acknowledgements

This work was supported by the National Science Foundation, award number 2138995.

Appendix

Table
Detailed Description of Call Response Tasks, Workload Subcategories, and Corresponding VACP Scores.

Task Sequence	Workload Subcategory	Dispatch Data Mapping	VACP						
			Visual	Auditory	Cognitive	Fine Motor	Gross Motor	Speech	Tactile
Navigating to the Scene			5.1	1	5	2.2	0	2	0
Driving to the Scene	Without Lights and Sirens	If Priority = P3, P4, P5, P6, P7, P8, P9, then P = 100 %	6	4.2	6.8	2.6	1	2	1
	With Lights and Sirens	If Priority = P1, P2, then P = 100 %	6	6.6	7	4.6	2	2	1
Moving to the Patient	Walking, Using Light Equipment	If green or orange, P = 95 %	4.4	6	1	0	1	2	0
	Walking, Using Stretcher	If green or orange, P = 5 % If yellow, P = 50 % If red, P = 50 %	4.4	6	2	2.2	3	2	0
	Using Stairs or Heavy Equipment	If yellow, P = 50 % If red, P = 100 %	4.4	6	2	2.2	5	2	0
Preliminary Checkup	Light	If green or orange, P = 95 %	4.4	6	6.8	2.2	1	4	1
	Medium	If green or orange, P = 5 % If yellow, P = 50 %	6	6.6	7	5.5	2	4	1
	Heavy	If yellow, P = 50 % If red, P = 100 %	6	7	7	7	3	4	2
Moving to the Ambulance	Patient Walking	If green or orange, P = 95 %	4.4	3	1	0	1	2	0
	Moving Patient to Stretcher	If green or orange, P = 5 % If yellow, P = 50 %	6	4.2	2	2.2	2	2	0
	Lifting Patient to Stretcher	If yellow, P = 50 % If red, P = 100 %	6	6	2	2.2	3.5	2	0
Ambulance Checkup	Light	If green or orange, P = 95 %	6	6	6.8	2.6	0	4	1

(continued on next page)

Table (*continued*)

Task Sequence	Workload Subcategory	Dispatch Data Mapping	VACP						
			Visual	Auditory	Cognitive	Fine Motor	Gross Motor	Speech	Tactile
	Medium	If green or orange, P = 5 % If yellow, P = 50 %	6	6.6	7	5.5	1	4	1
	Heavy	If yellow, P = 50 % If red, P = 100 %	6	7	7	7	2	4	2
Driving to the Hospital	Without Lights and Sirens	If Priority = P3, P4, P5, P6, P7, P8, P9, then P = 100 %	6	4.2	5	2.6	1	0	1
	With Lights and Sirens	If Priority = P1, P2, then P = 100 %	6	4.2	5	4.6	2	0	1
Checking Patient (Enroute to the Hospital)	Light	If green or orange, P = 95 %	6	6	6.8	2.6	0	4	1
	Medium	If green or orange, P = 5 % If yellow, P = 50 %	6	6.6	7	5.5	1	4	1
	Heavy	If yellow, P = 50 % If red, P = 100 %	6	7	7	7	2	4	2
Unloading the Patient/ Transporting to the Hospital Bed	Light	If green or orange, P = 95 %	4.4	6	1	0	1	2	0
	Medium	If green or orange, P = 5 % If yellow, P = 50 %	4.4	6	2	2.2	2	2	0
	Heavy	If yellow P = 50 % If red P = 100 %	4.4	6	2	2.2	3.5	2	0
Relaying Information to the Doctor			5.1	6	5.3	0	0	4	0
Cleaning up Stretcher and Vehicle	Reloading Stretcher	If green or orange, P = 100 %	4	1	1	2.2	2	0	0
	Restocking Supplies from Hospital	If yellow or red, P = 50 %	5	1	4.6	2.2	2	0	0
	Deep Cleaning Vehicle (COVID, etc.)	If yellow or red, P = 50 %	5	1	4.6	2.2	3.5	0	0

Data availability

Research data includes sensitive or confidential information such as patient data, which cannot be shared due to ethical and legal restrictions. However, all codes and deidentified data supporting this study are available upon request from the corresponding author.

References

Abrams, R., Jones, B., Campbell, J., de Lusignan, S., Peckham, S., Gage, H., 2024. The effect of general practice team composition and climate on staff and patient experiences: a systematic review. BJGP Open 8 (1). https://doi.org/10.3399/BJGPO.2023.0111. BJGPO.2023.0111.

Asselin, N., Choi, B., Pettit, C.C., Dannecker, M., Machan, J.T., Merck, D.L., et al., 2018. Comparative analysis of emergency medical service provider workload during simulated out-of-hospital cardiac arrest resuscitation using standard versus experimental protocols and equipment. Simulat. Healthc. J. Soc. Med. Simulat. 13 (6), 376–386.

Barger, L.K., Runyon, M.S., Renn, M.L., Moore, C.G., Weiss, P.M., Condle, J.P., et al., 2018. 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 22 (Suppl. 1), 58–68.

Bayley, R., Weinger, M., Meador, S., Slovis, C., 2008. Impact of ambulance crew configuration on simulated cardiac arrest resuscitation. Prehosp. Emerg. Care 12 (1), 62–68.

Borg, G., 1990. Psychophysical scaling with applications in physical work and the perception of exertion. Scand. J. Work. Environ. Health 16 (Suppl. 1), 55–58.

Bruhl, E.J., MacLaughlin, K.L., Allen, S.V., et al., 2020. Association of primary care team composition and clinician burnout in a primary care practice network. Mayo Clin Proc Innov Qual Outcomes. 4 (2), 135–142. https://doi.org/10.1016/j.mayocpiqo.2019.12.008.

Cortez, E.J., Panchal, A.R., Davis, J.E., Keseg, D.P., 2017. The effect of ambulance staffing models in a metropolitan, fire-based EMS system. Prehospital Disaster Med. 32 (2), 175–179.

Dai, M., Willard-Grace, R., Knox, M., et al., 2020. Team configurations, efficiency, and family physician burnout. J. Am. Board Fam. Med. 33 (3), 368–377. https://doi.org/10.3122/jabfm.2020.03.190336.

Dopelt, K., Wacht, O., Strugo, R., Miller, R., Kushnir, T., 2019. Factors that affect Israeli paramedics' decision to quit the profession: a mixed methods study. Isr. J. Health Pol. Res. 8 (1), 78.

Dropkin, J., Power, P., Rasul, R., Moline, J., Kim, H., 2019. Effect of resiliency and age on musculoskeletal injuries and lost workdays in emergency medical service personnel. Int. J. Ind. Ergon. 69, 184–193. https://doi.org/10.1016/j.ergon.2018.11.008.

Ducar, D.M., Penberthy, J.K., Schorling, J.B., Leavell, V.A., Calland, J.F., 2020. Mindfulness for healthcare providers fosters professional quality of life and mindful attention among emergency medical technicians. Explore 16 (1), 61–68.

Ercolani, J., Cure, L., Misasi, P., 2024. Relationship between conventional workload surrogates and VACP assessments in emergency medical services. In: *Proc IISE Annu Conf. Institute of Industrial and Systems Engineers* (IISE).

Fang, P.H., Lin, Y.Y., Lu, C.H., Lee, C.C., Lin, C.H., 2020. Impacts of emergency medical technician configurations on outcomes of patients with out-of-hospital cardiac arrest. Int. J. Environ. Res. Publ. Health 17 (6).

Fritzsche, L., Wegge, J., Schmauder, M., Kliegel, M., Schmidt, K.H., 2014. Good ergonomics and team diversity reduce absenteeism and errors in car manufacturing. Ergonomics 57 (2), 148–161. https://doi.org/10.1080/00140139.2013.875597.

Gage, C.B., Cooke, C.B., Powell, J.R., Kamholz, J.C., Kurth, J.D., van den Bergh, S., Panchal, A.R., 2024. Factors associated with emergency medical clinicians leaving EMS. Prehosp. Emerg. Care 1, 1–7. https://doi.org/10.1080/10903127.2024.2436047.

Hammarström, P., Rosendahl, S., Gruber, M., Nordin, S., 2023. Somatic symptoms in burnout in a general adult population. J. Psychosom. Res. 168, 111217. https://doi.org/10.1016/j.jpsychores.2023.111217.

Hart, S.G., Staveland, L.E., 1988. Development of NASA-TLX (task load index): results of empirical and theoretical research. Adv. Psychol. 52, 139–183.

He, Z., Qin, X., Renger, R., Souvannasacd, E., 2019. Using spatial regression methods to evaluate rural emergency medical services (EMS). Am. J. Emerg. Med. 37 (9), 1633–1642.

Huynh, N.H., Bui, H., Cure, L., 2020. Characterizing the workday of ambulance crew members. Proc IISE Annu Conf.

Law, A.M., 2015. Simulation Modeling and Analysis, fifth ed. McGraw-Hill Education, New York.

Leventhal, L., 2016. Updating the debate on one- versus two-tailed tests with the directional two-tailed test. Psychol. Rep. 84 (3), 707–718.

Lovejoy, M., Kelly, E.L., Kubzansky, L.D., Berkman, L.F., 2021. Work redesign for the 21st century: promising strategies for enhancing worker well-being. Am. J. Publ. Health 111 (10), 1787–1795. https://doi.org/10.2105/AJPH.2021.306283.

Lu, X., Guo, J., Zeng, S., Che, H., 2023. Mental overload assessment method considering the effects of performance shaping factors. IEEE Access 11, 48375–48391. https://doi.org/10.1109/ACCESS.2023.3277254.

McCracken, J., Aldrich, T.B., 1984. Analyses of selected LHX mission functions: implications for operator workload and system automation goals. Technical Note AD-A232-330.

Melamed, S., Kushnir, T., Shirom, A., 1992. Burnout and risk factors for cardiovascular diseases. Behav. Med. 18 (2), 53–60. https://doi.org/10.1080/08964289.1992.9935172.

Misasi, P., 2024. Recalibrating the design of paramedics' clinical decision support: A cognitive system engineering approach [PhD dissertation]. Wichita State University.

Othman, S.Y., Hassan, N.I., Mohamed, A.M., 2023. Effectiveness of mindfulness-based interventions on burnout and self-compassion among critical care nurses caring for patients with COVID-19: a quasi-experimental study. BMC Nurs. 22, 305.

Paterson, J.L., Sophianopoulos, S., Williams, B., 2014. What paramedics think about when they think about fatigue: contributing factors. Emerg. Med. Australasia (EMA) 26 (2), 139–144.

Patterson, P.D., Weaver, M.D., Guyette, F.X., Martin-Gill, C., 2020. Should public safety shift workers be allowed to nap while on duty? Am. J. Ind. Med. 63 (10), 843–850. https://doi.org/10.1002/ajim.23164.

Patterson, P.D., Martin, S.E., Brassil, B.N., Hsiao, W.H., Weaver, M.D., Okerman, T.S., et al., 2023. The emergency medical services sleep health study: a cluster-randomized trial. Sleep Health 9 (1), 64–76.

Peng, C.Y.J., Lee, K.L., Ingersoll, G.M., 2002. An introduction to logistic regression analysis and reporting. J. Educ. Res. 96 (1), 3–14.

Pineault, R., Borgès Da Silva, R., Provost, S., et al., 2017. Do gender-predominant primary health care organizations have an impact on patient experience of care, use of services, and unmet needs? Inquiry 54, 46958017709688. https://doi.org/10.1177/0046958017709688.

Rusnock, C.F., Borghetti, B.J., 2018. Workload profiles: a continuous measure of mental workload. Int. J. Ind. Ergon. 63, 49–64.

Rusnock, C.F., Borghetti, B.J., McQuaid, I., 2015. Objective-analytical measures of workload—The third pillar of workload triangulation? Lect. Notes Comput. Sci. 9183, 124–135.

Shotwell, D., Merlin, M., Robbins, V., 2018. Ambulance crew configuration: are two paramedics better than one? J. Emerg. Med. Serv.

Sun, J.T., Chiang, W.C., Hsieh, M.J., Huang, E.P., Yang, W.S., Chien, Y.C., et al., 2018. The effect of the number and level of emergency medical technicians on patient outcomes following out-of-hospital cardiac arrest in Taipei. Resuscitation 122, 48–53.

The difference between a paramedic and EMT [Webpage]. [cited 2024 Jul 25]. Available from: https://trueemergency.com/difference-between-a-paramedic-and-emt.

Thielmann, B., Schnell, J., Böckelmann, I., Schumann, H., 2022. Analysis of work-related factors, behavior, well-being outcome, and job satisfaction of workers of emergency medical service: a systematic review. Int. J. Environ. Res. Publ. Health 19 (23), 15887.

Tofil, N.M., Lin, Y., Zhong, J., Peterson, D.T., White, M.L., Grant, V., et al., 2017. Workload of team leaders and team members during a simulated sepsis scenario. Pediatr. Crit. Care Med. 18 (9), e423–e427. https://doi.org/10.1097/PCC.0000000000001258.

U.S. Bureau of Labor Statistics, 2024. In: Emergency medical technicians (EMTs) and paramedics [government document]. https://www.bls.gov/ooh/healthcare/emts-and-paramedics.htm.

U.S. Census Bureau, 2020. Sedgwick County, Kansas demographics [government document]. Sedgwick County KD [cited 2024 Jul 25]. Available from: https://www.census.gov/quickfacts/sedgwickcountykansas.

Vleminckx, S., Van Bogaert, P., De Meulenaere, K., et al., 2024. Factors influencing the formation of balanced care teams: the organisation, performance, and perception of nursing care teams and the link with patient outcomes: a systematic scoping review. BMC Health Serv. Res. 24, 1129. https://doi.org/10.1186/s12913-024-11625-5.

Wang, Z., Liu, S., Wanyan, X., Dang, Y., Chen, X., Zhang, X., 2024. Pilot workload measurement model based on task complexity analysis. Int. J. Ind. Ergon. 104, 103637. https://doi.org/10.1016/j.ergon.2024.103637.

Zavadasky, M.. Exploring new EMS workforce challenges: Q&A with CPSM's Matt Zavadasky [Internet]. [cited 2024 Jul 25]. Available from: https://www.cpsm.us/exploring-new-ems-workforce-challenges-matt-zavadasky.

Zhang, N., Harris, D., Li, W.C., 2023. A mental workload control method based on human performance or safety risk. Lect. Notes Comput. Sci. 14017, 168–177.