Natural Resources for Human Health
eISSN: 2583-1194

Original Research

Received 16 July 2026
Revised 17 Aug 2026
Accepted 03 Sep 2026

Natural Resources for Human Health 2026, Vol. 6 (20s): 864–878

KEYWORDS:
emergency medical services; emergency department; ambulance offload delay; clinical handover; prehospital notification; ambulance diversion; care coordination; healthcare quality; patient safety

The Impact of Coordination Between Ambulance Services and Emergency Departments on Healthcare Quality

Naif Turki Marzook[1], Saeed Ahmed Humran Alzahrani[2], Naif Amir Alfahmi[3], Farida Ahmed Abdali[4], Ahmed Alhussain Alragdi[5], Majed Thayib Aljuaid[6], Jehad Ibrahim Alqarni[7]

[1] Consultant Emergency Medicine
[2] Emergency Medical Specialist
[3] Emergency Medicine Consultant
[4] Clinical Nursing Director of Critical Care
[5] Emergency medical specialist
[6] Emergency Medical Specialist
[7] Emergency Medical Specialist
Workplaces\ Ministry of Defense King Fahad Armed Forces hospital,Jeddah

ABSTRACT:

Background. The interface between emergency medical services (EMS) and hospital emergency departments (EDs) is one of the most consequential, yet least governed, transitions in healthcare. Delays in ambulance offload, incomplete handover, absent prehospital notification and uncoordinated diversion compromise timeliness, safety and effectiveness of care, and remove ambulances from community availability.

Objective. To synthesize evidence on how EMS–ED coordination affects healthcare quality across the Institute of Medicine domains; to identify the mechanisms and system determinants of effective coordination; and to propose an implementable framework for practice, policy and research.

Methods. Narrative review informed by the SANRA criteria. PubMed/MEDLINE, CINAHL, Embase, Scopus and grey literature from ambulance and health-system bodies were searched for studies addressing ambulance offload, clinical handover, prehospital notification and activation pathways, ambulance diversion, information integration and joint governance.

Results. Prehospital notification is consistently associated with faster diagnosis and reperfusion in stroke and myocardial infarction. Structured handover tools improve completeness and efficiency of information transfer. Prolonged offload is associated with harm in several large cohorts and removes ambulance capacity from the community, although associations with mortality are modest and not uniform. Ambulance diversion offers little relief of crowding and has been linked to higher mortality in time-sensitive conditions. Root causes lie predominantly in hospital-wide access block rather than in the ED itself.

Conclusions. Coordination between ambulance services and EDs is a modifiable determinant of quality whose benefits depend on whole-system design. The proposed RELAY framework (Real-time information and prenotification, Efficient offload, Language of handover, Aligned system capacity, and Yardsticks with joint governance)

Natural Resources for Human Health
eISSN: 2583-1194

offers a structured pathway for health systems, including those undergoing transformation under Saudi Vision 2030.

Key Points

- The ambulance–ED interface is a high-risk care transition in which delays and information loss directly affect timeliness, safety and effectiveness of care.
- Prehospital notification and direct activation pathways shorten time to definitive treatment for stroke, STEMI and major trauma.
- Ambulance offload delay is driven mainly by hospital-wide access block; it reduces community ambulance availability and is associated with patient harm.
- The RELAY framework integrates information sharing, offload processes, standardized handover, system-wide capacity alignment and shared measurement.

1. INTRODUCTION

Emergency care is a continuum that begins with a call for help and ends with definitive treatment, yet it is typically delivered by two organizations with different governance, funding, information systems and professional cultures: the ambulance service and the hospital emergency department (ED). The point at which a patient passes from one to the other is a transition of care, and transitions are recognized as periods of heightened risk in which information is lost, responsibility is blurred and delays accumulate (The Joint Commission, 2017).

The quality of this transition matters well beyond the individual patient. When an ED cannot accept patients from arriving ambulances, paramedic crews wait at the hospital, sometimes for hours, and the ambulance is unavailable to respond to the next emergency in the community. In England, ambulance services reported that, on average, around half of all hospital handovers missed the 15-minute contractual standard, and a structured review of patients who waited more than an hour found that more than eight in ten were judged to have potentially experienced some degree of harm (Association of Ambulance Chief Executives [AACE], 2021). Similar pressures have been documented in Australia, Canada and the United States, where terms such as “ramping” and “offload delay” have entered public debate.

Conversely, effective coordination produces some of the clearest quality gains in modern emergency medicine. Prehospital electrocardiography and notification allow catheterization laboratories to be activated before a patient with ST-elevation myocardial infarction (STEMI) arrives; stroke prenotification shortens door-to-imaging and door-to-needle times (Lin et al., 2012); and regionalized trauma systems that direct patients to designated centers reduce mortality (MacKenzie et al., 2006). These benefits arise precisely because the ambulance service and the ED act as a single system rather than as sequential, independent providers.

The World Health Assembly has recognized emergency care systems as an integral component of universal health coverage, emphasizing integration across prehospital and facility-based care (World Health Organization [WHO], 2019). In the Kingdom of Saudi Arabia, where prehospital care is delivered principally by the Saudi Red Crescent Authority (SRCA) and hospital emergency care by Ministry of Health, military, university and private institutions, coordination across organizational boundaries is a particularly salient challenge and a stated priority of the Health Sector Transformation Program under Vision 2030 (Saudi Vision 2030, 2021).

This review addresses four questions: (1) How does coordination between ambulance services and EDs affect healthcare quality across established quality domains? (2) Which mechanisms and processes mediate these effects? (3) What system-level factors enable or impede coordination? (4) What integrated framework can guide improvement and evaluation? The analysis is anchored in Donabedian’s structure–process–outcome model (Donabedian, 1988), the Institute of Medicine (IOM) domains of quality (IOM, 2001) and Reason’s systems model of error (Reason, 2000).

2. REVIEW METHODOLOGY

A structured narrative review was undertaken in accordance with the Scale for the Assessment of Narrative Review Articles (SANRA). A narrative design was selected because the relevant evidence spans heterogeneous designs, including observational cohorts, quality-improvement evaluations, qualitative studies, operations-research models and policy reports, which are unsuitable for meta-analytic pooling.

2.1 Search Strategy and Sources

PubMed/MEDLINE, CINAHL, Embase and Scopus were searched from January 2000 to September 2026 using combinations of terms for the prehospital setting ("emergency medical services", "ambulance", "paramedic"), the hospital setting ("emergency department", "emergency service, hospital") and coordination processes ("handover", "handoff", "offload", "ramping", "turnaround", "prenotification", "pre-alert", "diversion", "interoperability", "integration"). Grey literature from national ambulance associations, health departments and the WHO was included where it provided system-level data unavailable in peer-reviewed sources. Reference lists of key reviews were hand-searched.

2.2 Eligibility and Synthesis

Studies were included if they examined an interaction between prehospital and ED care and reported an outcome relevant to at least one IOM quality domain: safety, effectiveness, patient-centeredness, timeliness, efficiency or equity. Priority was given to large multicenter cohorts, systematic reviews and controlled evaluations. Findings were synthesized thematically by coordination process and mapped to Donabedian's framework. The proposed framework was derived inductively from this synthesis.

3. CONCEPTUAL FOUNDATIONS

3.1 The EMS–ED Interface as a System Boundary

Asplin et al. (2003) described ED crowding through an input–throughput–output model in which ambulance arrivals form a key input, ED processes determine throughput, and the ability to move admitted patients to inpatient beds determines output. Within this model, the ambulance–ED interface sits at the junction of input and throughput: it is where the demands of one system meet the capacity constraints of another.

Coordination failures at this boundary are therefore frequently symptoms of downstream constraints rather than of the interface itself, a point that recurs throughout the evidence.

3.2 Quality Domains

The IOM (2001) defined quality care as safe, effective, patient-centered, timely, efficient and equitable. Each domain is engaged at the EMS–ED interface. Timeliness is affected by offload and treatment delays; safety by information loss and unmonitored waiting; effectiveness by the reliability of time-critical pathways; patient-centeredness by the experience of waiting on stretchers in corridors; efficiency by the idle time of ambulance crews; and equity by the uneven geographic impact of reduced ambulance availability and diversion.

3.3 Coordination Mechanisms

Coordination can be conceptualized through three complementary mechanisms. Information coordination ensures that the right data reach the right clinicians at the right time, through prenotification, structured handover and shared records. Process coordination aligns workflows across organizations, through activation protocols, offload procedures and destination policies. Relational and governance coordination establishes shared goals, mutual respect, joint accountability and common performance measures. Effective systems use all three; improvement efforts that address only one are rarely sustained.

Table 1. Mapping EMS–ED coordination to Donabedian's model and IOM quality domains

Donabedian level	Coordination element	Example	Quality domain primarily affected
Structure	Shared communication infrastructure	Electronic patient care record transmitted to ED	Safety, effectiveness
Structure	Designated arrival and offload areas	Ambulance offload zone staffed by ED nurses	Timeliness, efficiency
Structure	Regionalized specialist centers	Stroke, STEMI and trauma center designation	Effectiveness, equity
Process	Prehospital notification and activation	Pre-arrival cath lab or stroke team activation	Timeliness, effectiveness
Process	Structured clinical handover	IMIST-AMBO, MIST or ISBAR tools	Safety, patient-centeredness
Process	Destination and diversion protocols	Load-balancing in place of unilateral diversion	Equity, effectiveness
Outcome	Offload interval and turnaround time	Proportion offloaded within 15–30 min	Timeliness, efficiency
Outcome	Condition-specific outcomes	Door-to-needle, mortality, harm events	Effectiveness, safety

4. COORDINATION PROCESSES AND THEIR IMPACT ON QUALITY

4.1 Prehospital Notification and Direct Activation Pathways

Prehospital notification is the most extensively evidenced form of EMS–ED coordination. In an analysis of 371,988 patients with acute ischemic stroke transported by EMS in the Get With The Guidelines–Stroke program, approximately two-thirds were prenotified, and prenotification was independently associated with door-to-imaging within 25 minutes, door-to-needle within 60 minutes and onset-to-needle within 120 minutes; among patients arriving within two hours, a higher proportion of prenotified patients received thrombolysis (Lin et al., 2012). Because roughly one-third of patients were not prenotified, the study also revealed a substantial gap between recommended and actual practice.

For acute coronary syndromes, prehospital 12-lead electrocardiography combined with transmission or notification allows early diagnosis and, where systems permit, direct transfer to the catheterization laboratory, bypassing the ED. In a linked national cohort from the Myocardial Ischaemia National Audit Project in England and Wales, prehospital electrocardiography was associated with shorter times to reperfusion and lower mortality (Quinn et al., 2014). These pathways illustrate the principle that the ED need not be a mandatory waypoint when prehospital diagnostic information is sufficiently reliable and organizational agreements are in place.

Mobile stroke units extend this logic by bringing computed tomography and thrombolysis to the patient. Two large controlled studies showed that dispatch of mobile stroke units was associated with faster treatment and better functional outcomes than conventional ambulance transport (Ebinger et al., 2021; Grotta et al., 2021). Such models depend entirely on tight integration between dispatch, prehospital teams and receiving stroke centers.

For major trauma, coordination operates through field triage and destination protocols. In a national evaluation in the United States, care at a level I trauma center was associated with a significantly lower risk of death than care at non-trauma centers (MacKenzie et al., 2006). The benefit of such regionalization depends on accurate prehospital triage, reliable communication with the receiving trauma team, and the readiness of that team at the time of arrival.

> ◆ **KEY MESSAGE**
>
> When prehospital teams can identify time-critical conditions and hospitals can act on that information before arrival, the EMS–ED interface becomes a source of time saved rather than time lost. The gains are largest for stroke, STEMI and major trauma.

4.2 Clinical Handover

Handover is the transfer of professional responsibility and accountability for some or all aspects of a patient's care. Inadequate hand-off communication has been identified as a contributing factor in a large proportion of sentinel events, prompting the Joint Commission to issue specific guidance on standardized handover (The Joint Commission, 2017). The ambulance-to-ED handover is especially vulnerable: it often occurs in noisy, crowded environments, under time pressure, to clinicians who may be simultaneously occupied with other patients, and frequently more than once as the patient passes from triage nurse to treating nurse to physician.

A literature review of ambulance-to-ED handover identified recurrent problems: information omitted or not retained, lack of structure, poor listening by receiving staff, and inconsistent documentation (Bost et al., 2010). Paramedics often report that verbal information is not heard in full, while ED staff report that handovers are unstructured or that crucial details such as medications and allergies are missing.

Structured tools address these weaknesses. The IMIST-AMBO protocol (Identification, Mechanism or medical complaint, Injuries or information relating to the complaint, Signs and vital signs, Treatment and trends; Allergies, Medications, Background history, Other information) was co-designed with paramedics and ED clinicians using video-reflexive ethnography. After its introduction, handovers contained more information, were delivered in a more consistent order, were shorter, and generated fewer questions and less repetition (Iedema et al., 2012). Other mnemonics, including MIST for trauma and ISBAR for general use, share the same principles: predictable structure, protected time for the verbal handover, and confirmation of receipt.

Electronic patient care records transmitted directly into hospital systems complement the verbal handover by providing a durable, retrievable account of prehospital findings and interventions. Their value depends on interoperability: records that arrive as printouts or separate portals, or that are unavailable until after the patient has been assessed, do little to reduce information loss.

4.3 Ambulance Offload Delay

Ambulance offload delay (AOD) occurs when the transfer of a patient from ambulance stretcher to ED care cannot be completed promptly. In a review of 137 articles, Li et al. (2019) found that relatively few studies addressed AOD directly; most examined ED crowding or ambulance diversion, and the authors concluded that mitigation requires system-wide rather than ED-centered interventions. The National Association of EMS Physicians characterized AOD as a symptom of hospital crowding that endangers both the waiting patient and the community awaiting an ambulance (Cooney et al., 2011).

The consequences of AOD fall on two populations. For the waiting patient, risks include delayed assessment and treatment, deterioration while monitored by crews unfamiliar with ED resources, pressure injuries, and loss of dignity. For the community, each hour of offload delay is an hour in which an ambulance is unavailable, prolonging response times for other emergencies, including cardiac arrest.

Outcome evidence is consistent in direction but modest in magnitude and not uniform. In a population-based cohort of 213,544 patients with chest pain in Victoria, Australia, longer offload times were associated with higher adjusted 30-day mortality, with

the highest tertile (more than 28 minutes) carrying an adjusted risk difference of 0.28 percentage points compared with the lowest (Dawson et al., 2022). In contrast, a propensity-matched study of 70,290 high-acuity patients in Calgary, Canada, found no increase in 7- or 30-day mortality among patients delayed for 60 minutes or more, although ED length of stay was longer; the authors suggested that early hallway physician assessment may have mitigated harm (Stewart et al., 2019). The English ambulance services' structured harm review further suggests that prolonged waits expose frail and seriously ill patients to clinically meaningful risk (AACE, 2021).

These mixed findings carry an important lesson: the harm of AOD depends on what happens during the delay. Where waiting patients are triaged, monitored and treated by ED clinicians, outcomes may be protected; where they wait unassessed, harm is more likely. Offload delay is thus both a measure of system strain and a marker of whether shared responsibility for waiting patients has been clearly defined.

Table 2. Selected evidence on ambulance offload delay, diversion and outcomes

Study	Setting / sample	Exposure	Principal finding
Dawson et al., 2022	Victoria, Australia; 213,544 chest-pain patients	Offload time tertiles	Longer offload associated with higher adjusted 30-day mortality
Stewart et al., 2019	Calgary, Canada; 70,290 matched	Offload ≥60 vs ≤15 min	No mortality difference; longer ED stay
AACE, 2021	England; ten ambulance services	Handover >60 min	Over 80% potentially harmed; ~9% severe harm
Shen & Hsia, 2011	California; 13,860 AMI patients	Nearest ED on diversion ≥12 h	Higher 30-day and 1-year mortality
Burke et al., 2013	Boston, USA; nine hospitals	Statewide diversion ban	No increase in ED length of stay; shorter turnaround
Li et al., 2019	Review of 137 articles	Offload delay literature	Root causes systemic; whole-system solutions needed

4.4 Ambulance Diversion and Destination Coordination

Ambulance diversion, in which an ED temporarily requests that ambulances bypass it, is the most commonly studied response to crowding, yet its effects are controversial (Li et al., 2019). Diversion can create a cascade in which neighboring hospitals become overwhelmed and divert in turn, prolonging transport times without increasing system capacity. In a study of Medicare patients with acute myocardial infarction in four California counties, patients whose nearest ED was on diversion for 12 hours or more had higher mortality at 30 days, 90 days, 9 months and 1 year (Shen & Hsia, 2011).

Massachusetts banned diversion statewide in 2009. An evaluation of nine Boston-area hospitals found that ED length of stay for admitted patients did not increase, and ambulance turnaround time decreased modestly, despite rising ED volume (Burke et al., 2013). The authors attributed this to concurrent hospital-wide operational changes, reinforcing the view that crowding is a hospital-level rather than an ED-level problem. Coordinated, real-time load balancing, in which a regional coordination center distributes ambulances across hospitals according to capacity and clinical need, is increasingly favored over unilateral diversion.

4.5 Information Integration and Digital Coordination

Digital technologies offer the potential to transform coordination from episodic, voice-based communication into continuous shared situational awareness. Key capabilities include real-time transmission of vital signs and electrocardiograms; electronic

patient care records accessible within the hospital record; hospital capacity dashboards visible to dispatch centers; automated arrival notifications linked to vehicle location; and telemedicine consultation between paramedics and emergency physicians. Evidence for individual technologies is growing, but the principal barriers are organizational: incompatible systems, unclear data governance, and the absence of shared performance targets that would make integration a joint priority.

4.6 Continuity of Resuscitation and Critical Care

For patients in cardiac arrest, shock or respiratory failure, coordination determines whether resuscitation continues seamlessly or is interrupted at the hospital door. Prehospital notification allows the ED to assemble a resuscitation team, prepare equipment and, where available, alert cardiac catheterization, extracorporeal support or trauma surgical teams. The physical transfer from ambulance stretcher to ED bed is itself a high-risk moment, during which chest compressions, ventilation and infusions are vulnerable to interruption. Structured, rehearsed transfer procedures, with explicit role allocation and a brief "hands-off" handover once the patient is stabilized, protect against these risks.

Interventions initiated in the prehospital phase, including airway management, vasopressor infusions, blood products, tranexamic acid, analgesia and anticoagulation, must be accurately communicated so that they are neither duplicated nor omitted. Discrepancies in drug administration records at the interface are a recognized source of medication error, and electronic patient care records shared in real time are an important safeguard. In systems where physicians or critical care paramedics deliver advanced prehospital care, protocols should be co-authored with receiving EDs so that treatment begun in the field is continued, rather than reconsidered, on arrival.

4.7 Patient-Centeredness and Experience

The patient-centeredness domain is often neglected in discussions of the interface, which tend to focus on time intervals. Patients who wait on ambulance stretchers in corridors experience discomfort, lack of privacy, uncertainty about who is responsible for their care, delayed analgesia and limited access to toilets, food and family. Older patients and those with cognitive impairment are particularly vulnerable to distress, delirium and pressure injury during prolonged waits. A patient-centered interface therefore requires not only speed but also clear communication with patients and families about what is happening, who is responsible, and what to expect.

Handover itself can be designed to include the patient. Conducting the verbal handover at the bedside, where appropriate, allows patients to correct or supplement information, confirms identity and allergies, and signals continuity of care. Paramedics also hold unique contextual information, such as the condition of the home environment, social circumstances and the wishes of family members, that is relevant to discharge planning and is often lost if not explicitly transmitted.

5. ROOT CAUSES: WHY COORDINATION FAILS

The coordination failures described above share common root causes, which are summarized in Table 3.

5.1 Access Block and Hospital-Wide Flow

Systematic reviews of ED crowding identify inpatient boarding, the inability to transfer admitted patients to ward beds, as the dominant cause (Hoot & Aronsky, 2008; Morley et al., 2018). Crowding in turn is associated with delays in time-sensitive treatment, poorer patient experience and higher mortality (Bernstein et al., 2009). In a cohort of admitted patients in California, those admitted on days of high ED crowding had higher inpatient mortality and longer, costlier stays (Sun et al., 2013). When ED cubicles are occupied by boarded patients, arriving ambulance patients cannot be offloaded, regardless of the efficiency of the ED itself.

5.2 Fragmented Governance and Misaligned Incentives

Ambulance services are typically measured on response times, while hospitals are measured on ED length of stay or admission targets. When the clock for one organization stops at a point that starts the clock for the other, each may optimize its own metric at the expense of the system. For example, a hospital may delay formally "accepting" a patient to protect its ED time target, leaving the patient in the care of an ambulance crew. Joint metrics and shared accountability are necessary to remove such perverse incentives.

5.3 Professional Culture and Communication

Paramedics and ED clinicians differ in training, professional identity and situational pressures. Qualitative studies describe paramedics feeling unheard or undervalued at handover and ED staff perceiving incoming patients as additional load. Without structured interaction and mutual understanding of each other's constraints, these tensions undermine information transfer and collaborative problem-solving. Team-training approaches such as TeamSTEPPS, which emphasize structured communication, mutual support and shared mental models, provide a vehicle for addressing these cultural barriers across organizational boundaries (Agency for Healthcare Research and Quality [AHRQ], 2023).

5.4 Geographic and Workforce Constraints

In geographically dispersed systems, long transport distances amplify the cost of any delay. A qualitative study of EMS in Saudi Arabia found that rural crews faced long distances to hospitals, aging vehicles, variable dispatch knowledge of local geography, and a workforce with a lower proportion of paramedics relative to emergency medical technicians; participants also highlighted the need for better coordination between the SRCA and hospital-based Ministry of Health ambulance services (Alanazy et al., 2021). Such constraints make destination accuracy and prenotification even more valuable, since secondary transfers are costly and slow.

Table 3. Root causes of EMS–ED coordination failure and corresponding levers

Root cause	Manifestation at the interface	Primary improvement lever
Hospital access block	No cubicle available; prolonged offload; ramping	Whole-hospital flow, discharge planning, escalation plans
Misaligned metrics	Delayed formal acceptance; clock-stopping behavior	Shared offload and handover metrics; joint accountability
Unstructured communication	Omitted medications, allergies, trends	Standardized handover tool with protected time
Information-system fragmentation	Prehospital record unavailable to ED team	Interoperable electronic records; data standards
Uncoordinated diversion	Cascading diversion; longer transport	Regional load-balancing and capacity dashboards
Cultural and professional distance	Mistrust; handover not heard	Joint training, rotations, liaison roles
Geographic and workforce limits	Long transports; limited skill mix	Accurate destination protocols; telemedicine support

6. INTERVENTIONS TO IMPROVE COORDINATION

6.1 Offload Zones and Dedicated Arrival Staffing

Dedicated offload or arrival zones, staffed by ED nurses who assume responsibility for multiple ambulance patients simultaneously, allow paramedic crews to return to service while patients await a cubicle. Several Canadian and Australian evaluations have reported shorter offload intervals with such models. Their safety depends on appropriate patient selection, adequate nurse-to-patient ratios and clear criteria for escalation. They are a mitigation rather than a cure: by relocating rather than resolving waiting, they may mask the underlying access block if used in isolation.

6.2 Liaison Roles

Hospital ambulance liaison officers and EMS–hospital liaison programs place an experienced clinician or manager at the interface to coordinate arrivals, expedite handover, and escalate delays. These roles improve situational awareness and relationships between organizations, although robust outcome evaluations remain limited.

6.3 Standardized Handover and Joint Training

Implementing a single, co-designed handover tool across the ambulance service and all receiving hospitals in a region reduces variation and strengthens mutual expectations (Iedema et al., 2012). Joint simulation training, inter-professional education and clinical rotations of paramedics into EDs, and of ED staff into ambulance services, build shared mental models and trust.

6.4 Alternative Pathways and Demand Management

Not every patient who calls an ambulance requires ED care. Clinical hub triage, "hear and treat" telephone assessment, "see and treat" at scene, and direct referral to urgent-care centers, mental-health services or community teams reduce unnecessary conveyance and thus ED demand. These pathways require strong clinical governance and feedback loops between ambulance services and hospitals to ensure safety.

6.5 Whole-System Capacity Management

Because access block is the dominant cause of offload delay, the most effective interventions are those that improve hospital-wide flow: early-in-the-day discharges, discharge lounges, seven-day consultant review, full-capacity protocols that distribute boarded patients to inpatient corridors rather than the ED, and investment in community and social care capacity. The Massachusetts experience illustrates that removing the option of diversion, when accompanied by hospital-wide operational change, can be achieved without worsening ED performance (Burke et al., 2013).

Table 4. Interventions to strengthen EMS–ED coordination

Intervention	Mechanism	Evidence	Main caveat
Prehospital notification / activation	Pre-arrival mobilization of specialist teams	Strong	Requires reliable prehospital diagnosis
Structured handover (e.g., IMIST-AMBO)	Complete, ordered information transfer	Moderate	Needs protected time and training
Offload zones	Releases ambulance crews; pooled ED nursing	Moderate	May mask access block
Liaison officers	Real-time coordination and escalation	Limited–moderate	Few outcome evaluations
Regional load balancing	Distributes demand by capacity	Moderate	Needs real-time capacity data
Alternative pathways	Reduces avoidable conveyance	Moderate	Requires clinical governance
Whole-hospital flow measures	Addresses root cause of access block	Strong (crowding literature)	Organizationally demanding

7. THE RELAY FRAMEWORK FOR INTEGRATED PREHOSPITAL–HOSPITAL EMERGENCY CARE

Drawing on the evidence synthesized above, we propose the RELAY framework, which conceptualizes the ambulance–ED interface as a relay in which performance depends on the quality of the baton exchange as much as on the speed of either runner. The framework integrates Donabedian's structure–process–outcome logic with Reason's concept of layered defenses: each component provides an independent barrier against delay and information loss, so that a failure in one is less likely to cause harm.

Table 5. The RELAY framework: components, core actions and indicators

Component	Core actions	Suggested indicators
R – Real-time information and prenotification	Standardized prenotification criteria for stroke, STEMI, trauma, sepsis and cardiac arrest; transmission of ECGs and vital signs; direct activation pathways	% eligible patients prenotified; door-to-needle and door-to-balloon times
E – Efficient offload	Shared definition of offload time; arrival zones; clear transfer of responsibility on arrival; escalation thresholds at 30 and 60 min	Median offload time; % offloaded within 15 and 30 min; ambulance hours lost
L – Language of handover	Single regional handover tool; protected verbal handover; electronic record available at the bedside	Handover completeness audits; information-loss incidents
A – Aligned system capacity	Hospital-wide flow and full-capacity protocols; regional load balancing instead of diversion; alternative care pathways	Boarding time; diversion hours; non-conveyance rate with safety follow-up
Y – Yardsticks and joint governance	Joint EMS–hospital committee; shared dashboards; joint incident review; inter-professional training	Joint review cycle completed; staff experience scores; harm events at the interface

RELAY is designed to be scalable. At the level of a single hospital and its ambulance provider, it supports a local agreement on offload, handover and escalation. At regional level, it supports coordination centers that manage destination decisions and capacity. At national level, it informs standards, accreditation criteria and data infrastructure. Its emphasis on joint governance responds directly to the evidence that coordination failures arise from the boundary between organizations rather than within either one.

"The ambulance–ED interface is not a hand-off between two systems; it is the point at which they must become one."

7.1 Measuring Coordination Quality

Measurement is the foundation of the Yardsticks component. A balanced indicator set should combine time intervals, which are readily captured but incomplete, with measures of information quality, safety and experience. Definitions must be agreed jointly: for example, offload time should run from ambulance arrival at the hospital, recorded automatically where possible, to the documented transfer of clinical responsibility, not merely to physical transfer to a bed. Indicators should be reported jointly to ambulance and hospital leadership, stratified by acuity and time of day, and interpreted alongside hospital occupancy and boarding data so that interface performance is not attributed solely to the ED.

Table 6. Proposed indicator set for EMS–ED coordination quality

Domain	Indicator	Type	Data source
Timeliness	Median and 90th-percentile offload time; % within 15 and 30 min	Process	Ambulance and ED timestamps

Domain	Indicator	Type	Data source
Timeliness	Door-to-needle, door-to-balloon and trauma-team activation times	Process	Disease registries
Effectiveness	% eligible patients prenotified; % direct-to-lab activations	Process	Dispatch and ED records
Safety	Deterioration or harm events during offload wait	Outcome	Incident reporting; joint review
Safety	Handover completeness (medications, allergies, treatments)	Process	Structured audits
Efficiency	Ambulance hours lost to delays beyond 30 min	Outcome	Ambulance service data
Equity	Diversion hours; response times by region	Outcome	Coordination center
Patient-centeredness	Patient-reported experience of arrival and handover	Outcome	Patient surveys

7.2 Phased Implementation

Implementation of RELAY is best approached in phases that build relationships and data before structural change. In the first phase, foundation, a joint governance committee is established, offload and handover definitions are agreed, baseline data are collected and a common handover tool is selected. In the second phase, process redesign, prenotification criteria, offload procedures, escalation thresholds and arrival-area arrangements are implemented and tested through plan–do–study–act cycles. In the third phase, system integration, interoperable records, capacity dashboards and regional load balancing are introduced, and hospital-wide flow initiatives are aligned with interface targets. In the fourth phase, sustainment, indicators are embedded in routine reporting, accreditation and contracts, and joint training becomes part of continuing professional development.

Table 7. Phased roadmap for implementing the RELAY framework

Phase	Indicative timing	Key activities	Milestone
1. Foundation	0–6 months	Joint committee; shared definitions; baseline audit; handover tool selection	Signed interface agreement
2. Process redesign	6–18 months	Prenotification criteria; offload and escalation protocols; joint training	Offload and handover targets met in pilot sites
3. System integration	12–36 months	Interoperable records; capacity dashboards; regional load balancing; hospital-wide flow	Region-wide data sharing operational
4. Sustainment	Ongoing	Indicators in accreditation and contracts; continuous improvement; research	Sustained performance with public reporting

8. IMPLICATIONS FOR SAUDI ARABIA AND VISION 2030

Saudi Arabia offers an instructive context for applying these principles. Prehospital emergency care is delivered principally by the SRCA, while emergency departments operate under multiple governmental and private providers. The Health Sector Transformation Program, a component of Vision 2030, aims to improve access, quality and efficiency, and is reorganizing Ministry of Health services into regional health clusters with integrated accountability for population health (Saudi Vision 2030, 2021). This restructuring creates an opportunity to embed EMS–ED coordination into cluster-level governance rather than leaving it to bilateral arrangements between individual stations and hospitals.

Several practical priorities follow from the RELAY framework. First, national definitions and targets for offload and handover times, jointly owned by the SRCA and receiving hospitals, would provide a shared yardstick. Second, a nationally endorsed handover standard, incorporated into SRCA protocols and hospital policies, would reduce variation. Third, interoperability between SRCA electronic patient records and hospital information systems, consistent with the Kingdom's national digital health agenda, would allow prehospital data to follow the patient. Fourth, accreditation standards of the Saudi Central Board for Accreditation of Healthcare Institutions (CBAHI) could incorporate interface metrics into emergency department performance indicators. Fifth, education and professional development regulated by the Saudi Commission for Health Specialties (SCFHS) could include inter-professional training in handover and coordination for paramedics, emergency nurses and emergency physicians.

Mass-gathering events, notably the annual Hajj and Umrah seasons, add a distinctive dimension. They require surge coordination between prehospital and hospital services at scale, and the experience gained in these settings, including regional command structures and pre-planned destination protocols, could inform routine coordination nationally. Assessments of SRCA preparedness for mass-casualty incidents have identified variation across regions and highlighted the importance of personnel, training and evaluation (Alotaibi & Khan, 2019).

Finally, national research capacity is needed. Peer-reviewed data on offload times, handover quality and their association with outcomes in Saudi hospitals remain sparse. Linked prehospital–hospital datasets, supported by academic centers and national research institutes, would allow the impact of coordination initiatives to be measured and benchmarked internationally.

9. DISCUSSION

The evidence reviewed supports three overarching conclusions. First, coordination between ambulance services and EDs has measurable effects on quality, most clearly for timeliness and effectiveness in time-critical conditions, where prenotification and activation pathways are associated with faster treatment and, in some studies, lower mortality. Second, failures of coordination, manifested as offload delay, information loss and uncoordinated diversion, are associated with harm, although the magnitude of measured mortality effects is modest and context-dependent. Third, the root causes of these failures lie predominantly outside the interface, in hospital-wide capacity, fragmented governance and misaligned incentives.

The apparently inconsistent findings regarding offload delay deserve emphasis. Mortality is a blunt outcome measure for a phenomenon whose harms include delayed analgesia, pressure injury, loss of dignity, missed deterioration and, critically, the unmeasured harm to community patients who wait longer for an ambulance. Studies that report no mortality association may also reflect local mitigation, such as early physician assessment. Future research should therefore adopt broader outcome sets, including patient-reported experience, harm events and community response-time effects, and should evaluate the processes that occur during delay rather than delay alone.

From a patient-safety perspective, the interface exemplifies Reason's systems model: harm rarely results from a single individual failure but from the alignment of latent conditions, such as access block, ambiguous responsibility and incompatible information systems, with active failures such as an incomplete handover. Interventions that rely solely on individual effort, such as exhorting staff to offload faster, are unlikely to succeed. Durable improvement requires redesigning the structures and processes that create these latent conditions.

From a management perspective, the interface is a case study in inter-organizational coordination. Shared metrics, joint governance and relational coordination, characterized by shared goals, shared knowledge and mutual respect, are prerequisites for sustained performance. Health systems that measure ambulance services and hospitals separately will continue to generate behaviors that optimize each organization at the expense of the patient.

10. RESEARCH GAPS AND FUTURE DIRECTIONS

- **Causal evidence:** controlled or quasi-experimental evaluations of offload and handover interventions with patient-level outcomes are scarce; stepped-wedge designs are well suited to regional implementation.
- **Community-level impact:** studies should quantify how offload delay affects response times and outcomes for other patients awaiting ambulances.
- **Patient experience:** the experience of patients waiting on ambulance stretchers is rarely measured and should be included in quality frameworks.
- **Digital integration:** the effect of interoperable prehospital–hospital records and capacity dashboards on information loss and destination decisions requires rigorous evaluation.
- **Equity:** rural, elderly and linguistically diverse populations may be disproportionately affected by coordination failures and need targeted study.
- **Regional evidence:** linked prehospital–hospital data from Saudi Arabia and the wider Gulf region are needed to inform context-specific standards.

11. LIMITATIONS

As a narrative review, this work is subject to selection and interpretive bias, mitigated but not eliminated by a structured search and explicit criteria. The evidence base is dominated by observational studies from high-income countries with established EMS systems, limiting causal inference and generalizability. Definitions of offload time and handover quality vary across studies, and some system-level data derive from grey literature rather than peer-reviewed sources. The RELAY framework is a synthesis intended to guide implementation; it has not yet been prospectively validated.

12. CONCLUSIONS

The interface between ambulance services and emergency departments is a critical determinant of healthcare quality. Effective coordination, through prehospital notification, structured handover, timely offload, coordinated destination decisions and integrated information, improves timeliness and effectiveness of care and protects patient safety. Failures at the interface harm both the waiting patient and the wider community. Because the causes of these failures are largely systemic, solutions must be system-wide, combining hospital-wide flow improvement with joint governance and shared measurement. The RELAY framework provides a structured approach to this task, and health systems undergoing major transformation, such as that of Saudi Arabia under Vision 2030, are well placed to embed it from the outset.

DECLARATIONS

Funding. This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflicts of interest. The author(s) declare no conflicts of interest.

Ethical approval. Not applicable; this review analyzed previously published data only.

Data availability. No new data were generated or analyzed in this study.

Author contributions. [To be completed according to the CRediT taxonomy before submission.]

Use of generative AI. Generative AI tools were used to assist with drafting and structuring this manuscript. The author(s) reviewed, verified and edited all content and take full responsibility for the final text, in accordance with ICMJE recommendations and the target journal's policy.

REFERENCES

[1] Agency for Healthcare Research and Quality. (2023). TeamSTEPPS 3.0: Team strategies and tools to enhance performance and patient safety. AHRQ.

[2] Alanazy, A. R. M., Fraser, J., & Wark, S. (2021). Organisational factors affecting emergency medical services' performance in rural and urban areas of Saudi Arabia. BMC Health Services Research, 21. https://doi.org/10.1186/s12913-021-06565-3

[3] Alotaibi, M. S., & Khan, A. A. (2019). Assessing the pre-hospital care preparedness to face mass casualty incident in Saudi Arabia in 2017–2018. Saudi Medical Journal, 40(10), 1032–1039. https://doi.org/10.15537/smj.2019.10.24292

[4] Asplin, B. R., Magid, D. J., Rhodes, K. V., Solberg, L. I., Lurie, N., & Camargo, C. A., Jr. (2003). A conceptual model of emergency department crowding. Annals of Emergency Medicine, 42(2), 173–180.

[5] Association of Ambulance Chief Executives. (2021). Delayed hospital handovers: Impact assessment of patient harm. AACE.

[6] Bernstein, S. L., Aronsky, D., Duseja, R., Epstein, S., Handel, D., Hwang, U., et al. (2009). The effect of emergency department crowding on clinically oriented outcomes. Academic Emergency Medicine, 16(1), 1–10.

[7] Bost, N., Crilly, J., Wallis, M., Patterson, E., & Chaboyer, W. (2010). Clinical handover of patients arriving by ambulance to the emergency department: A literature review. International Emergency Nursing, 18(4), 210–220.

[8] Burke, L. G., Joyce, N., Baker, W. E., Biddinger, P. D., Dyer, K. S., Friedman, F. D., et al. (2013). The effect of an ambulance diversion ban on emergency department length of stay and ambulance turnaround time. Annals of Emergency Medicine, 61(3), 303–311.

[9] Cooney, D. R., Millin, M. G., Carter, A., Lawner, B. J., Nable, J. V., & Wallus, H. J. (2011). Ambulance diversion and emergency department offload delay: Resource document for the National Association of EMS Physicians position statement. Prehospital Emergency Care, 15(4), 555–561.

[10] Dawson, L. P., Andrew, E., Stephenson, M., Nehme, Z., Bloom, J., Cox, S., Anderson, D., Lefkovits, J., Taylor, A. J., Kaye, D., Smith, K., & Stub, D. (2022). The influence of ambulance offload time on 30-day risks of death and re-presentation for patients with chest pain. Medical Journal of Australia, 217(5), 253–259. https://doi.org/10.5694/mja2.51613

[11] Donabedian, A. (1988). The quality of care: How can it be assessed? JAMA, 260(12), 1743–1748.

[12] Ebinger, M., Siegerink, B., Kunz, A., Wendt, M., Weber, J. E., Schwabauer, E., et al. (2021). Association between dispatch of mobile stroke units and functional outcomes among patients with acute ischemic stroke in Berlin. JAMA, 325(5), 454–466.

[13] Grotta, J. C., Yamal, J. M., Parker, S. A., Rajan, S. S., Gonzales, N. R., Jones, W. J., et al. (2021). Prospective, multicenter, controlled trial of mobile stroke units. New England Journal of Medicine, 385(11), 971–981.

[14] Hoot, N. R., & Aronsky, D. (2008). Systematic review of emergency department crowding: Causes, effects, and solutions. Annals of Emergency Medicine, 52(2), 126–136.

[15] Iedema, R., Ball, C., Daly, B., Young, J., Green, T., Middleton, P. M., Foster-Curry, C., Jones, M., Hoy, S., & Comerford, D. (2012). Design and trial of a new ambulance-to-emergency department handover protocol: 'IMIST-AMBO'. BMJ Quality & Safety, 21(8), 627–633.

[16] Institute of Medicine. (2001). Crossing the quality chasm: A new health system for the 21st century. National Academies Press.

[17] Li, M., Vanberkel, P., & Carter, A. J. E. (2019). A review on ambulance offload delay literature. Health Care Management Science, 22(4), 658–675. https://doi.org/10.1007/s10729-018-9450-x

[18] Lin, C. B., Peterson, E. D., Smith, E. E., Saver, J. L., Liang, L., Xian, Y., et al. (2012). Emergency medical service hospital prenotification is associated with improved evaluation and treatment of acute ischemic stroke. Circulation: Cardiovascular Quality and Outcomes, 5(4), 514–522.

[19] MacKenzie, E. J., Rivara, F. P., Jurkovich, G. J., Nathens, A. B., Frey, K. P., Egleston, B. L., Salkever, D. S., & Scharfstein, D. O. (2006). A national evaluation of the effect of trauma-center care on mortality. New England Journal of Medicine, 354(4), 366–378.

[20] Morley, C., Unwin, M., Peterson, G. M., Stankovich, J., & Kinsman, L. (2018). Emergency department crowding: A systematic review of causes, consequences and solutions. PLoS ONE, 13(8), e0203316.

[21] Quinn, T., Johnsen, S., Gale, C. P., Snooks, H., McLean, S., Woollard, M., & Weston, C. (2014). Effects of prehospital 12-lead ECG on processes of care and mortality in acute coronary syndrome: A linked cohort study from the Myocardial Ischaemia National Audit Project. Heart, 100(12), 944–950.

[22] Reason, J. (2000). Human error: Models and management. BMJ, 320(7237), 768–770.

[23] Saudi Vision 2030. (2021). Health Sector Transformation Program: Delivery plan. Government of the Kingdom of Saudi Arabia.

[24] Shen, Y. C., & Hsia, R. Y. (2011). Association between ambulance diversion and survival among patients with acute myocardial infarction. JAMA, 305(23), 2440–2447.

[25] Stewart, D., Lang, E., Wang, D., & Innes, G. (2019). Are emergency medical services offload delay patients at increased risk of adverse outcomes? CJEM, 21(4), 505–512. https://doi.org/10.1017/cem.2018.478

[26] Sun, B. C., Hsia, R. Y., Weiss, R. E., Zingmond, D., Liang, L. J., Han, W., McCreath, H., & Asch, S. M. (2013). Effect of emergency department crowding on outcomes of admitted patients. Annals of Emergency Medicine, 61(6), 605–611.

[27] The Joint Commission. (2017). Inadequate hand-off communication (Sentinel Event Alert, Issue 58). The Joint Commission.

[28] World Health Organization. (2019). Emergency care systems for universal health coverage: Ensuring timely care for the acutely ill and injured (Resolution WHA72.16). Seventy-second World Health Assembly.