Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Implementation and real-world evaluation of an integrated pre- hospital emergency service system: a retrospective before-after implementation study

Zichun Tang
Sichuan University
Lin Tang
Sichuan University
Cong Li
Sichuan University
Yi Chen
Sichuan University
Xingying Ren
Sichuan University
Bin Zhou
Sichuan University
Le Tong
Sichuan University
Shiyuan Tang
Sichuan University
Wei Zhang
Sichuan University
Jun Guo
guojun@wchscu.cn
Sichuan University

Research Article

Keywords: emergency medical services, pre-hospital emergency care, health information system, dispatch, triage, interrupted time-series, usability, before-after study

Posted Date: November 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11703940/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Emergency medical services (EMS) operate under tight time constraints, but dispatch, ambulance, and hospital systems often remain poorly interoperable, causing repeated data entry and verbal handover. Evidence on end-to-end integrated pre-hospital platforms, especially segment-specific operational effects, triage consistency, and staff usability, remains limited.

Methods

We conducted a single-centre retrospective before-after implementation study at West China Hospital, Sichuan University, comparing missions that ended in transport to the study emergency department with completed handover before (September 2024 to August 2025) and after (October 2025 to September 2026) roll-out of the Pre-hospital Emergency Service System (PHESS). Of 4917 records screened, 968 were excluded, corresponding to 21.3% of records before and 17.7% after implementation. Outcomes were dispatch response time, scene arrival time, on-scene management time, total completion time, pre-hospital-in-hospital triage agreement, and mHealth App Usability Questionnaire (MAUQ) scores. Adjusted log-linear regression, demand- and disease-stratified comparisons, and monthly and weekly interrupted time-series analyses were used.

Results

Among 3949 missions (2138 before and 1811 after implementation), median recorded scene arrival time decreased from 594 to 567 seconds (P<.001). The adjusted common period time ratio was 0.949 (95% CI 0.928–0.969; P<.001), and the interaction with the green-channel-eligible group was nonsignificant (P=.100). Dispatch response, on-scene management and total completion times showed no clear between-period differences. Exact triage agreement was 98.78% before and 96.20% after implementation; quadratic-weighted kappa was 0.981 and 0.938. Among 32 surveyed providers, mean MAUQ score was 6.27 (SD 0.82) of 7 (87.8/100).

Conclusions

Among missions that ended in transport to the study emergency department with completed handover, PHESS was associated with shorter recorded scene arrival time and with lower documented triage agreement, and hands-on on-scene duration was unchanged. The composite interval cannot separate information transfer from en-route movement, and the change in timestamp source between periods may have contributed to the timing difference. Staff acceptance was favourable. These single-centre implementation findings require multicentre validation, clinical outcome linkage and broader confounder adjustment.

Background

Emergency medical services (EMS) face rising demand under time constraints that leave little room for delay. Ischemic heart disease, stroke, and injuries together cause more than 19 million deaths each year [1, 2], and most of them begin in the community, where EMS is the first and most time-critical point of contact. Delayed cardiopulmonary resuscitation lowers cardiac arrest survival by 7%–10% for each minute of delay [3]. Stroke guidelines require door-to-needle times within 60 minutes [4], and trauma outcomes depend on the "golden hour" [5]. Because call volumes are high, even a small per-case efficiency gain produces a large population-level benefit, which makes pre-hospital efficiency an attractive target for improvement. EMS information systems, however, remain fragmented. Dispatch, ambulance, and hospital run systems that cannot exchange data. Patient details are collected again at every interface, and handover depends on verbal communication that is prone to error [6, 7]. The resulting workflow is serial and duplicative, and it consumes exactly the minutes on which survival depends.

Integrated health information systems have shown substantial efficiency gains across several health care domains. A systematic review of 44 studies found that such systems reduced length of stay and waiting times, streamlined care processes, and strengthened interdisciplinary coordination [8]. Clinical information systems with embedded decision support improved protocol adherence, reduced medication errors, and increased documentation completeness [9]. In emergency care, however, adoption remains limited to component-level solutions rather than fully integrated systems. Optimised dispatch protocols have improved pre-hospital recognition of acute coronary syndrome [10]. Direct admission pathways have shortened reperfusion time in STEMI [11, 12]. Mobile stroke units have reduced time-to-treatment [13]. None of these constitutes an end-to-end integrated EMS information platform [14].

Three evidence gaps remain. Most evaluations have looked at component-level tools such as dispatch aids, routing modules, or disease-specific pathways, not at a platform that spans dispatch, ambulance, and hospital. How much end-to-end integration changes system performance is therefore still unclear [15, 16]. Reported benefits are also usually limited to a single interval, most often scene arrival or transport time, so it is hard to tell which segment of the response chain actually changed. Without segment-specific measurement, an observed gain cannot be attributed to faster information transfer any more than to concurrent operational changes [17]. Finally, unintended consequences for process quality are seldom assessed. Agreement between pre-hospital and in-hospital triage shows whether digitised data entry supports or disrupts clinical judgment, and before-after comparisons that ignore underlying temporal trends cannot separate a system effect from secular change [18, 19].

To address these gaps, we developed and implemented an integrated Pre-hospital Emergency Service System (PHESS) at West China Hospital, Sichuan University. The system has a three-tier architecture that connects the municipal 120 dispatch platform, ambulance crews, and hospital information systems in both directions.

This study aimed to describe the development and implementation of PHESS and evaluate its real-world associations with operational efficiency, triage process quality, and user satisfaction using a before-after implementation design. We expected deployment to be associated with shorter scene arrival time because it improved dispatch, location, routing, and information sharing. We expected on-scene management time not to change materially, because it reflects hands-on clinical work rather than information transfer. Triage agreement was treated as a process-quality indicator; no clinical outcome or patient-care benefit was assumed.

Methods

Study design and population

This single-centre retrospective before-after implementation study was conducted in the pre-hospital emergency team and emergency department of West China Hospital, Sichuan University. The pre-implementation period served as a within-site comparison because no concurrent control group was available. Two 12-month periods were compared: September 2024 to August 2025 under the traditional manual workflow and October 2025 to September 2026 under routine PHESS use; September 2025 was excluded as the transition month. Interrupted time-series analysis was used to assess underlying temporal trends, and the report follows the STROBE guideline (Additional file 1).

System development and data interfaces

The PHESS is a dual-terminal platform comprising a PC workstation client and an Android mobile client. A shared backend database and data coordination layer support real-time bidirectional synchronisation between the two terminals.

Relative to the previous workflow, PHESS converts serial dispatch notification into parallel task assignment and connects pre-hospital and hospital information systems. The electronic workflow is intended to reduce duplicate post-mission entry and reliance on retrospective recall; direct documentation-time or cost savings were not measured in this study. The mobile client supports one-tap calling, intelligent navigation, structured documentation, triage-guided ordering, electronic informed

consent, and digital handover signatures. Dedicated green-channel routing, automated statistics, scheduling reminders, and embedded quality-control checks are described in Figs. 1 to 3; screenshots use demonstration data.

PC workstation client for in-hospital dispatch and management staff

1. A real-time mission dashboard.
2. Structured export of mission records with automated descriptive statistics.
3. Monthly pre-hospital quality control against preset documentation, timeliness and triage indicators.
4. Shift scheduling with automatic conflict detection and one-click crew notification.

Figure 1 PC workstation interfaces of the West China integrated PHESS.

Android mobile client for on-site pre-hospital crews

1. One-tap calling of patients or witnesses with call logs attached to the mission.
2. Geographic information system based route planning to the scene and the target hospital.
3. Structured pre-hospital record documentation.
4. Nurse confirmation and documentation of order execution.
5. Electronic capture of informed consent, fee confirmation and handover signatures.

System workflow and implementation

The PHESS workflow is organised in two stages. Stage 1 covers call reception, dispatch, and en-route care. Stage 2 covers on-scene care, completion, and handover (Fig. 3). Screenshots were obtained using demonstration data.

Stage 1: Call Reception, Dispatch and En-Route Care

When the Chengdu Emergency Medical Command Center issues a command sheet, a corresponding mission is created in PHESS, automatically linked to the 120 dispatch record and synchronised bidirectionally. The scheduled physician, nurse, and driver are then notified in parallel according to the duty roster, and cancellations are transmitted to all notified crew members and archived with their recorded reason. During transport, the crew uses the Android terminal for one-tap calls to the patient or bystanders and for intelligent navigation to the scene. Every key time node including task receipt, ambulance departure, arrival at scene, departure from scene, and hospital arrival is captured automatically as a timestamped field rather than recalled retrospectively.

Stage 2: On-Scene Care, Completion and Handover

On arrival at the patient, the crew establishes a hospital visit card using the patient's national identification number. For patients with previous visits, prior records are retrieved automatically and identity information is populated without re-entry. The physician completes the pre-hospital medical record while the nurse enters vital signs, and both independently assign a four-level triage grade (I–IV). When the two ratings differ, the higher grade is adopted. The physician issues pre-hospital medical orders through the system and the nurse executes and documents them. Before leaving the scene or on arrival at the hospital, fee collection and the informed consent signature are completed electronically. On arrival at the emergency department an electronic handover is performed with the receiving team, both parties sign electronically, and the complete record is transmitted to the 120 electronic medical record platform and archived.

Eligibility Criteria

Missions were eligible for inclusion if they met all of the following criteria, and were excluded if they met any of the exclusion criteria (Table 1).

Table 1
Eligibility criteria for pre-hospital emergency mission records

Items	Criteria
Inclusion	
I-1	Pre-hospital emergency missions dispatched by the Chengdu Emergency Medical Command Center, or generated in-hospital for emergency response or inter hospital transfer. Inter hospital transfer missions were eligible only when the patient was subsequently evaluated in our emergency department.
I-2	Missions executed by the West China Hospital prehospital emergency team during the defined pre-implementation or post-implementation period.
I-3	Missions were eligible if they ended in transport to our emergency department and completed handover
I-4	Missions with an analysable record: for the post-implementation period, a PHESS mission record; for the pre-implementation period, a record in the emergency department prehospital visit registry.
Exclusion	
E-1	Patient declined prehospital care or transport after crew arrival (refusal of care)
E-2	Patient transferred directly to another hospital without retrieval by the West China Hospital prehospital team
E-3	Patient found dead on crew arrival (no prehospital intervention possible)
E-4	Patient not found at the reported scene location
E-5	Mission cancelled en route before patient contact was established

Outcome measures

Source and comparability of the time data

Time records came from two sources. After implementation every time node was written by PHESS itself, which stored it as a server-side timestamp at the moment the corresponding action was entered, so each timestamp marks the moment the action was documented and not a later reconstruction of it. Before implementation the same nodes were extracted retrospectively from the routine pre-hospital emergency registry kept under the existing manual workflow. Call receipt, the start of three of the four intervals, was written by the municipal 120 command platform, which was not replaced as part of the intervention, so the same start point was used in both periods. The remaining nodes were recorded by the attending crew, in the registry before implementation and in PHESS afterwards. Both sources stored clock times, and intervals were computed from them without rounding or truncation. Because the two sources captured time in different ways, one set of interval definitions was applied to both periods. Each interval was calculated as the difference between the same pair of documented time nodes, no field or definition specific to either period was used, and no interval was reconstructed from free-text notes. The four metrics therefore describe the same operational segments before and after implementation, and the only intended difference between the periods is the way in which the underlying timestamps were recorded.

Primary outcomes

The primary outcomes were four operational time metrics, each defined by a pair of documented time nodes:

1. Dispatch response time: from call receipt at the municipal 120 command centre to departure of the ambulance from the ambulance station.
2. Scene arrival time: from call receipt to arrival of the ambulance at the scene.
3. On-scene management time: from arrival at the scene to departure of the ambulance from the scene.
4. Total completion time: from call receipt to mission completion, taken as the end of patient handover at the receiving emergency department.

Time records were complete for all 3949 missions in both periods, so no time metric required imputation. Scene arrival time was the metric carried forward to multivariable modelling. It is the interval in which a change in the transfer of information among dispatch, crew and receiving hospital would show most directly, because it combines dispatch processing with the journey to the scene and leaves out the clinical work performed there. The three remaining metrics were analysed to locate any change within the response chain. Dispatch response time reflects processing upstream of the crew, and on-scene management time reflects care delivered at the scene. Total completion time shows whether any gain persisted to the end of the mission.

Secondary outcomes

Secondary outcomes were process quality and user experience. Process quality was measured as concordance between the pre-hospital and the in-hospital triage grade, quantified by exact agreement, quadratic weighted κ, the under-triage rate and the over-triage rate. Under-triage was defined as a pre-hospital grade indicating lower acuity than the in-hospital grade, and over-triage as a pre-hospital grade indicating higher acuity. Exact agreement, under-triage and over-triage together partition the paired records. User experience was measured with the 18-item 7-point MAUQ (stand-alone version), which covers ease of use (items 1–5), interface and satisfaction (items 6–12) and usefulness (items 13–18). The items were rated from 1 (strongly disagree) to 7 (strongly agree), so a higher score indicates greater usability. Item, subscale and total scores were summarised as means and expressed on both the original 1–7 scale and a linear 0–100 scale, converted as (mean − 1) × 100/6.

Classification variables

Three classification variables were used both to describe the cohort and as covariates in the models. Free-text emergency diagnoses were standardised into 13 emergency categories with the keyword mapping rules given in Additional file 2, Table S1. The rules were applied in the order listed, so every record received exactly one category. Calls were classified by hourly call volume into early morning (00:00–06:59), peak (08:00–12:59 and 17:00–22:59) and off-peak (all remaining hours). Green-channel eligibility was derived from the documented diagnosis and identified missions whose diagnosis fell in one of three categories served by dedicated clinical pathways, namely stroke/neurological deficit, cardiovascular emergencies including cardiac arrest, and trauma/injury. Because activation of a green-channel pathway was not logged in either period, this variable indicates eligibility by diagnostic category and not confirmed pathway use. It was used both as a stratum and as an effect modifier.

Missing data

Triage grades were the only variables with missing values. An in-hospital grade was missing for 5 missions before implementation and for 21 after implementation. In 13 of these 26 missions the patient died during or shortly after arrival, before a grade could be assigned. In 5 of the 21 post-implementation missions the pre-hospital grade was also missing, so these 5 records carried no grade of any kind. Handling differed by analysis. (i) All 3949 missions were retained in every analysis of the four time metrics and in the interrupted time series, because no time value was missing. (ii) The 26 missions without an in-hospital grade were excluded from the triage concordance analysis only, which was based on 2133 pre-implementation and 1790 post-implementation paired records. (iii) For the descriptive comparison in Table 2 the 5 missions without any grade were coded as triage level 1, so that the four triage levels summed to the full cohort. That code was used for description only. In the multivariable models triage level entered as two indicator terms (levels II and III, with level I as reference), so no record was left out of the models because of a missing grade. No outcome or exposure variable was imputed.

Statistical analysis

Descriptive and unadjusted comparisons

Categorical variables were summarised as counts and percentages, and continuous variables as medians with 25th and 75th percentiles, because all four time metrics were right-skewed. Time metrics were compared between periods with Mann–

Whitney U tests. Exact triage agreement was compared with χ^2 tests, and disease-stratified agreement with Fisher exact tests where expected cell counts were small. Two-sided P values below .05 were considered statistically significant.

Triage concordance

Triage concordance was quantified with quadratic weighted κ. This statistic weights each disagreement by the square of the number of triage levels separating the two grades and is therefore appropriate for an ordered four-level scale. κ was reported with 95% confidence intervals, and under-triage and over-triage as proportions of paired records with 95% confidence intervals.

Association between implementation period and scene arrival time

The association between implementation period and scene arrival time was estimated by ordinary least squares on the natural logarithm of scene arrival time, with heteroscedasticity-consistent (HC3) standard errors. Model 1 adjusted for pre-hospital triage level, demand period (peak hours as reference) and green-channel eligibility, and returned one adjusted period association for the whole cohort. Triage level entered as two indicator terms for levels II and III; level I formed the reference category together with the three post-implementation level IV records and the five records that carried no grade, so all 3949 missions were retained in both models. Model 2 added a period × green-channel-eligibility product term, which allowed the period association to differ between missions whose diagnosis was green-channel-eligible and all other missions. Exponentiated coefficients are reported as time ratios, that is, the ratio of the adjusted geometric mean after implementation to that before implementation, so a time ratio below 1 indicates a shorter interval. For green-channel eligible missions the time ratio was calculated as $\exp(\beta_{period} + \beta_{interaction})$. Its variance, $\mathrm{Var}(\beta_{period}) + \mathrm{Var}(\beta_{interaction}) + 2\mathrm{Cov}(\beta_{period}, \beta_{interaction})$, was taken from the model covariance matrix and evaluated with the delta method.

Unmeasured confounders

Distance from station to scene, traffic conditions, weather, vehicle type, crew experience and concurrent operational changes were not recorded in either source system and could not be adjusted for. Their likely influence on the period estimate is addressed in the Discussion.

Interrupted time series

A supplementary interrupted time-series analysis was performed on monthly geometric means of each time metric. The series covered the 25 calendar months from September 2024 to September 2026, of which 24 contributed observations, because September 2025, the month in which PHESS went live, was excluded as the transition month. Time was indexed in months from the start of the series. Each model estimated a pre-implementation slope (β_1), a level change at the transition month (β_2) and a change in slope after implementation (β_3); the post-implementation slope is therefore $\beta_1 + \beta_3$, and the level change is expressed relative to the fitted value in August 2025, the last pre-implementation month. No observation was carried across the excluded transition month. Newey–West standard errors were used with two lags in the monthly models and four lags in the weekly model, chosen after residual autocorrelation was assessed, and the Durbin–Watson statistic is reported for every model. Two sensitivity analyses were performed. One used a weekly specification of the same model, with weekly geometric means and the same indexing (Additional file 2, Table S2), and the other refitted the monthly model without the first post-implementation month (October 2025).

Multiplicity

Because the disease-stratified and time-stratified analyses involved multiple related comparisons, P values were additionally adjusted with the Benjamini–Hochberg procedure. Adjustment was applied separately within each family of related tests, that is, within each time metric across the three demand strata for the time-stratified analysis and within each time metric across the 13 diagnostic categories for the disease-stratified analysis. Unadjusted and adjusted P values are reported together. The demand-stratified and disease-stratified estimates are exploratory, and the primary comparisons were the common adjusted period association for scene arrival time and the overall triage concordance analysis.

User experience

MAUQ scores were summarised as means with standard deviations and compared with the neutral midpoint of the 7-point scale (score 4) by one-sample t tests. Standardised mean differences are reported as Cohen d. Internal consistency (Cronbach α) was taken from the original validation of the MAUQ and not re-estimated in this sample, and it is annotated as such in Table 4.

Software

Analyses were performed in R 3.3 (R Foundation for Statistical Computing, Vienna, Austria) and IBM SPSS Statistics version 24 (IBM Corp, Armonk, NY).

Results

Participant Selection and Flow

All records identified in the two source systems were screened against the eligibility criteria. The selection process is summarised in the study flow diagram (Fig. 4). For each study period, the diagram reports the number of records identified, the number excluded at each step with the reason, and the number finally included in the analysis.

A total of 4917 mission records were identified and screened: 2716 in the pre-implementation period and 2201 in the post-implementation period. Of these, 968 missions were excluded (578 pre-implementation and 390 post-implementation). The final analysis set comprised 2138 records in the pre-implementation period and 1811 in the post-implementation period.

Study population

A total of 3949 pre-hospital emergency records were analysed. Baseline characteristics differed between the two periods (Table 2). The distribution of calls across demand periods was similar (P=.597), while the disease mix changed substantially (P<.001), with a lower proportion of trauma/injury (29.3% vs 23.5%) and higher proportions of neurological (20.0% vs 23.7%) and respiratory (7.2% vs 8.6%) emergencies after implementation. Pre-hospital triage level did not differ between periods (P=.109). Level 4 cases were rare and were recorded only after implementation (0 vs 3), and the two periods still did not differ when levels 1–3 were compared alone (P=.286). The most prevalent diagnoses overall were trauma/injury (26.6%), neurological emergencies (consciousness/seizure) (21.7%), and cardiovascular emergencies (incl. cardiac arrest) (10.4%). Pre-hospital triage was predominantly level 3 (urgent, 62.8%), followed by level 2 (emergent, 32.9%) and level 1 (critical, 4.3%).

Table 2
Baseline characteristics of the study population

Characteristic	Before (n = 2138)	After (n = 1811)	Total (N = 3949)	P value
Time period, n (%)				.597
Peak hours (08:00–12:59, 17:00–22:59)	1179 (55.1)	1012 (55.9)	2191 (55.5)	
Off-peak hours (07:00–07:59, 13:00–16:59, 23:00–23:59)	511 (23.9)	443 (24.5)	954 (24.2)	
Early-morning (00:00–06:59)	448 (21.0)	356 (19.7)	804 (20.4)	
Disease category, n (%)				< .001
Trauma/injury	627 (29.3)	425 (23.5)	1052 (26.6)	
Neurological emergencies (consciousness/seizure)	427 (20.0)	429 (23.7)	856 (21.7)	
Digestive system emergencies (incl. GI bleeding)	229 (10.7)	173 (9.6)	402 (10.2)	
Cardiovascular emergencies (incl. cardiac arrest)	196 (9.2)	213 (11.8)	409 (10.4)	
Poisoning (alcohol/drug/other)	99 (4.6)	61 (3.4)	160 (4.1)	
Respiratory emergencies	155 (7.2)	156 (8.6)	311 (7.9)	
Other/special circumstances	126 (5.9)	92 (5.1)	218 (5.5)	
Stroke/neurological deficit	105 (4.9)	116 (6.4)	221 (5.6)	
Pain/musculoskeletal (non-specific)	67 (3.1)	59 (3.3)	126 (3.2)	
Fever/infection	65 (3.0)	54 (3.0)	119 (3.0)	
Hemorrhagic emergencies (non-traumatic)	16 (0.7)	13 (0.7)	29 (0.7)	
Psychiatric & behavioural disorders and self-harm/suicide	16 (0.7)	12 (0.7)	28 (0.7)	
Environmental & accidental injuries (drowning/heatstroke)	10 (0.5)	8 (0.4)	18 (0.5)	
Pre-hospital triage, n (%)				.109
Level 1 (Critical)	86 (4.0)	82 (4.5)	168 (4.3)	
Level 2 (Emergent)	726 (34.0)	574 (31.7)	1300 (32.9)	
Level 3 (Urgent)	1326 (62.0)	1152 (63.6)	2478 (62.8)	

Characteristic	Before (n = 2138)	After (n = 1811)	Total (N = 3949)	P value
Level 4 (Minor)	0 (0.0)	3 (0.2)	3 (0.1)	

Pre-hospital triage grade was not documented for 5 post-implementation missions, and these missions were counted as level 1 in this table.

Operational time outcomes

Table 3 summarises the four time metrics for the overall cohort. Median recorded scene arrival time was shorter after implementation (594 to 567 s; P<.001), and the adjusted ratio of geometric mean times was 0.949 (95% CI 0.928 to 0.969), reported in the multivariable analysis below. Median dispatch response time was 143 s in both periods, although the distributions differed (P=.011). Total completion time (1530 to 1526 s; P=.313) and on-scene management time (911 to 914 s; P=.487) showed no clear between-period differences. Trauma/injury was the only disease category in which the scene-arrival reduction survived Benjamini-Hochberg correction (591 to 523 s; -68 s; -11.5%; BH-adjusted P<.001). Stroke/neurological deficit showed a nominal reduction (670 to 612 s; -58 s; -8.6%) that did not survive correction (BH-adjusted P=.105). Full disease-stratified results are reported in Additional file 2, Table S3.

Table 3
Overall operational time metrics by period

Emergency category	Time metric	Pre M (P25, P75)	Post M (P25, P75)	P
Overall (All cases)	Dispatch response time	143 (122, 165)	143 (113, 166)	.011
Overall (All cases)	Scene arrival time	594 (471, 725)	567 (440, 716)	< .001
Overall (All cases)	On-scene management time	911 (620, 1267)	914 (668, 1220)	.487
Overall (All cases)	Total completion time	1530 (1169, 1965)	1526 (1172, 1923)	.313

Triage agreement and usability evaluation

System usability and user acceptance

All 32 pre-hospital providers on duty during the usability evaluation period were invited and completed the questionnaire (response rate 100.0%), including 21 nurses (65.6%) and 11 physicians (34.4%). Reported system-use duration and weekly use frequency are summarised in Additional file 2, Table S4.

Table 4 presents subscale and overall MAUQ results. Item-level scores are reported in Additional file 2, Table S6. All 18 items scored above the neutral midpoint of 4 (all P<.001), with means ranging from 5.75 to 6.53 on the 7-point scale. The three subscales were well balanced: interface and satisfaction 6.35 +/- 0.78, usefulness 6.22 +/- 0.83, and ease of use 6.21 +/- 0.85. The overall MAUQ total score was 6.27 +/- 0.82 (87.8/100), indicating favourable subjective acceptance.

Table 4
MAUQ subscale and overall scores (n = 32)

Subscale	Items	Mean	SD	t	P	Cohen d	Original MAUQ α
Ease of use	1–5	6.21	0.85	14.66	< .001	2.59	0.847
Interface and satisfaction	6–12	6.35	0.78	17.15	< .001	3.03	0.908
Usefulness	13–18	6.22	0.83	15.20	< .001	2.69	0.717
MAUQ total	1–18	6.27	0.82	15.66	< .001	2.77	0.914

Cronbach α values are from the original validation of the standalone MAUQ [23]. They were not estimated from the 32 respondents in this study.

Pre-hospital–in-hospital triage agreement

Among record pairs with both pre-hospital and in-hospital triage levels documented, exact agreement decreased after implementation (Table 5): 98.78% (2107/2133) before versus 96.20% (1722/1790) after, and quadratic-weighted kappa decreased from 0.981 to 0.938 (P<.001). Under-triage increased from 0.70% to 2.23% and over-triage from 0.52% to 1.56%. The in-hospital triage distribution did not differ between periods (P=.39). Disease-stratified agreement results are reported in Additional file 2, Table S5.

Table 5
Triage agreement before and after implementation

Indicator	Before (n = 2133)	After (n = 1790)	Difference (after − before)	Test statistic	P value
Agreement, % (95% CI)	98.78 (98.22–99.17)	96.20 (95.21–96.99)	-2.58	χ^2=27.70	< .001
Under-triage, % (95% CI)	0.70 (0.43–1.16)	2.23 (1.65–3.03)	+ 1.53	χ^2=16.51	< .001
Over-triage, % (95% CI)	0.52 (0.29–0.92)	1.56 (1.08–2.25)	+ 1.05	χ^2=10.87	< .001
Quadratic-weighted κ (95% CI)	0.981 (0.974–0.988)	0.938 (0.922–0.955)	-0.043	z = 4.72	< .001

Time-stratified analysis

Table 6 reports the four time metrics by demand period. Scene arrival time decreased in all three strata; reductions during peak hours (607 to 575 s; -32 s; BH-adjusted P=.003) and early-morning hours (546 to 517 s; -28 s; BH-adjusted P=.003) remained significant after correction, whereas the off-peak reduction was smaller and nonsignificant (607 to 584 s; BH-adjusted P=.101). On-scene management and total completion times did not change significantly in any stratum after correction.

Table 6
Operational time metrics by demand period

Time period	Metric	Pre-implementation M (P25, P75)	Post-implementation M (P25, P75)	Δ Median (s)	Δ (%)	P	BH-adjusted P
Peak hours	Dispatch response time	142 (119, 166)	144 (116, 166)	+ 2	+ 1.4	.390	.390
	Scene arrival time	607 (475, 741)	575 (446, 726)	-32	-5.3	.002	.003
	On-scene management time	913 (622, 1285)	926 (686, 1250)	+ 13	+ 1.4	.418	.554
	Total completion time	1554 (1180, 2005)	1540 (1192, 1940)	-14	-0.9	.614	.628
Off-peak hours	Dispatch response time	146 (125, 166)	144 (114, 167)	-2	-1.4	.195	.293
	Scene arrival time	607 (486, 740)	584 (460, 739)	-23	-3.8	.101	.101
	On-scene management time	970 (656, 1302)	925 (695, 1210)	-45	-4.6	.554	.554
	Total completion time	1605 (1238, 1993)	1567 (1202, 1928)	-38	-2.4	.250	.628
Early-morning hours	Dispatch response time	143 (122, 165)	138 (107, 161)	-4	-3.1	.004	.011
	Scene arrival time	546 (438, 668)	517 (396, 632)	-28	-5.2	.002	.003
	On-scene management time	852 (566, 1160)	882 (610, 1190)	+ 30	+ 3.6	.534	.554
	Total completion time	1431 (1108, 1803)	1408 (1055, 1811)	-22	-1.6	.628	.628

Differences and percentage changes were calculated from unrounded medians. The Benjamini–Hochberg adjusted P values were calculated within each time metric across the three demand strata (m = 3).

Multivariable analysis

Table 7 presents two log-linear models for scene arrival time, adjusted for pre-hospital triage level, demand period, and green-channel eligibility. Model 1 estimated the common adjusted period association: TR = 0.949 (95% CI 0.928–0.969; P<.001). In Model 2, the green-channel-ineligible period TR was 0.964 (95% CI 0.936–0.992; P=.011), the green-channel-eligible main effect was 1.029 (95% CI 0.998–1.060; P=.065), and the period-by-green-channel-eligibility interaction was 0.964 (95% CI 0.922–1.007; P=.100), indicating no statistically significant heterogeneity. The exploratory green-channel-eligible period TR was 0.929 (95% CI 0.898–0.960; P<.001), estimated from the model covariance matrix by the delta method. This P value refers to the within-group period contrast and is not the interaction test, which is the .100 reported above. Because that test was nonsignificant, the stratum-specific estimate should be interpreted as exploratory.

Table 7
Adjusted scene-arrival regression models

Variable	β (95% CI)	Time ratio (95% CI)	P value
Model 1 · Main-effects model — common adjusted period association (n = 3949)			
Post-implementation vs pre-implementation	−0.052 (−0.074 to −0.032)	0.949 (0.928 to 0.969)	< .001
Model 2 · Period × green-channel-eligibility interaction (n = 3949)			
Post-implementation vs pre-implementation, green-channel-ineligible	−0.037 (−0.066 to −0.008)	0.964 (0.936 to 0.992)	.011
Green-channel-eligible group (vs other diagnostic groups)	+ 0.028 (−0.002 to +0.058)	1.029 (0.998 to 1.060)	.065
Interaction: green-channel-eligible × post (ratio of period TRs)	−0.037 (−0.081 to +0.007)	0.964 (0.922 to 1.007)	.100
Green-channel-eligible group period TR = exp(βpost + βinteraction), delta method	—	0.929 (0.898 to 0.960)	< .001
Adjusted covariates in Model 2 (reference: level 1, peak hours, other diagnostic groups)			
Pre-hospital triage level 2 (vs level 1)	+ 0.029 (−0.026 to +0.085)	1.030 (0.974 to 1.088)	.300
Pre-hospital triage level 3 (vs level 1)	+ 0.011 (−0.043 to +0.065)	1.011 (0.958 to 1.068)	.683
Off-peak hours (vs peak hours)	+ 0.011 (−0.015 to +0.038)	1.011 (0.985 to 1.038)	.406
Early-morning hours (vs peak hours)	−0.103 (−0.131 to −0.075)	0.902 (0.877 to 0.928)	< .001

Interrupted time-series analysis

Figure 5 shows the trend changes before and after the application of the system. Before implementation, scene arrival time had been increasing by + 5.62 s per month (P<.001). Implementation did not produce an immediate step change (-8.9 s; 95% CI -53.3 to + 35.6; P=.695) but reversed the trend: the monthly slope changed by -13.38 s (95% CI -19.36 to -7.41; P<.001), leaving a post-implementation slope of -7.76 s per month. Dispatch response time increased immediately by + 21.7 s (95% CI + 4.3 to + 39.1; P=.014) and then declined by -4.20 s per month (P<.001), consistent with an initial run-in period. On-scene management time showed no significant slope change (P=.635). The weekly specification (Additional file 2, Table S2) reproduced the monthly pattern for scene arrival time, with a level change of -11.7 s (95% CI -51.0 to + 27.7; P=.562) and a slope change of -2.59 s per week (95% CI -4.14 to -1.05; P=.001), leaving a post-implementation slope of -1.60 s per week. For total completion time the monthly slope change was − 14.88 s per month (P=.043) whereas the weekly slope change was − 2.83 s per week (95% CI -5.94 to + 0.27; P=.074), and this inconsistency prevents a robust conclusion that total completion time improved. Excluding the first post-implementation month (October 2025) yielded TR = 0.954 (95% CI 0.925–0.983; P=.002) for scene arrival time (Table 8).

Table 8
Interrupted time-series models based on monthly geometric means

Metric	Pre-implementation slope β_1 (s per month)	Level change β_2 (95% CI)	Slope change β_3 (95% CI)	Post-implementation slope	R^2	Durbin–Watson	Newey–West lags
Dispatch response time	−0.12 (P=.808)	+ 21.7 (+ 4.3 to + 39.1; P=.014)	−4.20 (− 5.45 to − 2.94; P<.001)	−4.32	0.72	1.68	2
Scene arrival time	+ 5.62 (P<.001)	−8.9 (− 53.3 to + 35.6; P=.695)	−13.38 (− 19.36 to − 7.41; P<.001)	−7.76	0.55	1.41	2
On-scene management time	−2.25 (P=.561)	+ 44.4 (− 18.6 to + 107.4; P=.167)	−2.37 (− 12.13 to + 7.40; P=.635)	−4.62	0.08	1.54	2
Total completion time	+ 3.52 (P=.475)	+ 25.8 (− 81.5 to + 133.0; P=.638)	−14.88 (− 29.32 to − 0.44; P=.043)	−11.35	0.22	1.55	2

Discussion

Principal findings

After deployment, recorded scene arrival time decreased, whereas dispatch response, on-scene management and total completion times showed no clear between-period differences. This pattern is compatible with a change concentrated in the part of the response chain that precedes on-scene care rather than in hands-on clinical work, although the composite interval cannot separate information transfer from en-route movement. Unchanged on-scene duration may reflect stable clinical workload, teamwork, and the time required for structured field assessment [20]. Bidirectional information exchange may support in-hospital preparation, but clinical outcomes were not measured in this study.

The unchanged on-scene duration also indicates a boundary for informatics interventions: digitisation can reduce coordination and documentation latency, but hands-on assessment and treatment depend on clinical content, crew composition, and scene conditions rather than information flow alone [20, 21]. Additional system features are therefore unlikely to shorten on-scene intervals without changes to clinical pathways or crew configuration [22]. Segment-specific metrics are preferable to a single aggregate interval because aggregate measures can conceal which part of the response chain is modifiable.

Beyond measured response intervals, PHESS changed the documentation workflow. The paper-based process required post-return records and could rely on retrospective recall; concurrent electronic entry may reduce repeated transcription and documentation burden, although direct time or cost savings were not measured. This potential documentation benefit is reflected in the usefulness subscale but should not be interpreted as a measured cost-effectiveness outcome.

The MAUQ evaluation showed favourable overall usability, with all subscales above the neutral midpoint [23, 24]. Subscale scores indicated acceptable ease of use, interface satisfaction, and perceived usefulness. The lowest item related to use under poor network conditions, suggesting that offline caching and local temporary storage should be prioritised in future iterations [25].

Poor-network performance was the main usability limitation; larger standard deviations for some interface and operation-time items also suggest variability across field network and device conditions. Future development should emphasise fault-

tolerant offline behaviour and consistent performance across devices.

Although absolute agreement remained high, exact agreement and weighted kappa declined. The larger change was in under-triage, which rose by 1.53 percentage points, from 0.70% to 2.23%, and identifies the main calibration target for refresher triage training and future decision-support prompts, particularly where field information is limited [28, 29]. Over-triage rose by 1.05 percentage points, from 0.52% to 1.56%, a smaller change that may reflect a cautious field threshold permitting proactive pre-hospital management and early hospital activation, at some cost in downstream resource use [26, 27]. The two grades are also recorded at different points in care, and the in-hospital grade can be revised once the patient is seen, so part of the discordance may reflect the timing of assessment and the availability of pre-hospital information to the receiving team [19].

Benefits were observed across demand strata, with significant corrected reductions in peak and early-morning scene arrival time. Peak periods contribute the largest call volume, so even modest per-case reductions may have operational relevance, although this study did not measure patient outcomes or resource savings.

Trauma/injury was the only disease category whose scene-arrival reduction survived multiple-comparison correction, potentially because geographically dispersed trauma events benefit from routing and dispatch information [22]. Stroke/neurological deficit showed a nominal reduction but did not survive correction and requires validation in a larger sample [20]. These subgroup estimates were exploratory, and a difference in within-category significance is not evidence that the period association differed between categories.

The interrupted time-series analysis provides two complementary observations. First, scene arrival time had been worsening before roll-out, so a simple before-after comparison would have mixed the pre-existing trend with implementation effects. Second, the lack of an immediate step change followed by a slope reversal is compatible with gradual adoption and workflow learning, rather than a sudden operational transformation. The absence of a slope change in on-scene management time is consistent with a change concentrated in the interval that precedes on-scene care, although this analysis cannot identify which component of that interval changed.

Comparison with prior work

Many prior evaluations of EMS digital interventions focused on partial, component-level tools rather than complete multi-tier integrated dispatch-ambulance-hospital platforms. Previous work has documented time savings from individual functions such as mobile stroke units, computer-aided dispatch decision aids, and optimised routing modules [10, 13]. Consistent with our findings, several of those studies reported reductions in scene arrival or transport time, whereas on-scene clinical durations remained largely unchanged [10]. Unlike disease-specific or single-step solutions, our system integrates municipal dispatch, pre-hospital crews, and hospital-side modules and contributes implementation evidence for an end-to-end workflow [22].

Few published EMS informatics studies have described demand-stratified and disease-stratified patterns alongside aggregate effects [22]. Trauma/injury was the only disease category with a corrected scene-arrival reduction; this pattern may be related to geolocation and intelligent ambulance assignment for geographically dispersed events, consistent with geospatial EMS research [10]. Stroke/neurological deficit showed a nominal but BH-nonsignificant reduction, likely limited by the subgroup sample size [20].

Most evaluations of EMS information-system roll-outs emphasise positive operational gains and report unintended process-quality changes less often. As in other real-world EMS digital projects, our platform captured triage data but had not yet implemented triage-weighted priority dispatch rules [29]; collecting technical data is not enough on its own to change operational behaviour, which also requires protocol revision and staff training. By assessing time-efficiency metrics and triage classification quality together, our study adds real-world evidence for integrated pre-hospital informatics deployments.

Limitations

Several limitations should be weighed against these strengths. This single-centre retrospective before-after study lacked a concurrent control, and secular trends cannot be fully ruled out. The timestamp source also changed between periods: before implementation the time nodes were extracted from the routine registry, and afterwards they were written automatically by PHESS. Although one set of interval definitions was applied to both periods, part of the between-period difference may reflect this change in recording rather than a change in practice. Eligible missions were limited to those ending in transport to the study emergency department with completed handover, and the excluded proportion differed between periods (578/2716, 21.3% before and 390/2201, 17.7% after), so the estimates apply to missions that reached that end point and may not extend to missions that were not transported or not retrieved by the study team. Adjustments were restricted to triage level, demand period and green-channel eligibility, while key contextual variables including distance, traffic and crew experience were unavailable, so residual confounding may bias associations in either direction. Standard errors were heteroscedasticity consistent but did not account for possible clustering of missions within crews or calendar days. Downstream clinical outcomes were not collected, precluding clinical or cost-effectiveness claims. MAUQ results came from 32 post-implementation providers and used reference Cronbach α from external work. Three rare post-implementation level 4 triage cases, all of which fell into the reference category of the model, constrain subgroup interpretation. Multicentre studies incorporating richer covariates and clinical outcomes are needed to improve generalisability.

Conclusions

In this single-centre retrospective before-after implementation study of missions that ended in transport to the study emergency department with completed handover, PHESS was associated with shorter recorded and adjusted scene arrival time, while on-scene management and total completion times did not change materially. Triage agreement declined and under-triage increased; this is a calibration and safety signal that should be addressed through triage training and decision support rather than being read as a benefit of the system. Frontline providers reported favourable usability. The evidence supports PHESS as an operational coordination tool and not as proof of improved clinical outcomes. Implementation should be paired with triage calibration, offline reliability improvements and multicentre studies with broader confounder adjustment and clinical outcome linkage.

Abbreviations

BH Benjamini-Hochberg

CI Confidence interval

ED Emergency department

EMS Emergency medical services

GI Gastrointestinal

HC3 Heteroscedasticity-consistent standard errors

ITS Interrupted time-series

MAUQ mHealth App Usability Questionnaire

PHESS Pre-hospital Emergency Service System

SD Standard deviation

STEMI ST-elevation myocardial infarction

STROBE Strengthening the Reporting of Observational Studies in Epidemiology

TR Time ratio.

Declarations

Ethics approval and consent to participate The Biomedical Research Ethics Committee of West China Hospital of Sichuan University approved the study (No. 2026[2527]). Written informed consent was waived for the retrospective analysis of de-identified operational mission records. Providers gave electronic informed consent before completing the usability survey, and survey responses were collected and analysed anonymously.

Consent for publication Not applicable as no individual person's data are contained in this manuscript in any identifiable form.

Availability of data and materials Mission-level records contain linked clinical identifiers and cannot be deposited publicly under the approved ethics protocol. De-identified aggregated tables and reproducible analysis code may be shared with qualified researchers after institutional review and data-use approval; requests should be addressed to the corresponding author.

Competing interests The authors declare no competing interests.

Funding This work was supported by the National Natural Science Foundation of China (No. 82572539); Hospital Management Research Project, West China Hospital, Sichuan University (ZYYG26004); and the 1.3.5 project for hospital management research, West China Hospital, Sichuan University (No. HX20220935). The funders had no role in study design, data collection, analysis, interpretation, decision to publish, or manuscript preparation.

Author contributions ZT and Lin Tang conceived and designed the study. ZT developed the methodological framework; ZT and Lin Tang built the analytical scripts. ZT and BZ performed data validation. Lin Tang and XR conducted formal data analysis and interpreted results on de-identified mission datasets. ZT and Lin Tang carried out the investigation. JG provided study resources. ST and JG completed data curation. ZT, Lin Tang, CL and YC drafted the original manuscript. ZT, Lin Tang, ST and JG critically revised the manuscript for important intellectual content. ZT and Lin Tang generated study visualisations. Le Tong, WZ and JG supervised the overall project. Le Tong, ST, WZ and JG undertook project administration. ST and JG secured study funding acquisition. All authors read and approved the final manuscript.

Acknowledgements We would like to thank the emergency departments that participated in the system operation throughout the hospital working days, as their feedback is crucial for evaluating system performance.

References

1. Roth GA, Mensah GA, Johnson CO, Addolorato G, Ammirati E, Baddour LM, et al. Global burden of cardiovascular diseases and risk factors, 1990–2019. J Am Coll Cardiol. 2020;76:2982–3021. https://doi.org/10.1016/j.jacc.2020.11.010.
2. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. Lancet. 2020;396:1204–22. https://doi.org/10.1016/S0140-6736(20)30925-9.
3. Harris AW, Kudenchuk PJ. Cardiopulmonary resuscitation: the science behind the hands. Heart Br Cardiac Soc. 2018;104:1056–61. https://doi.org/10.1136/heartjnl-2017-312696.
4. Prabhakaran S, Gonzalez NR, Zachrison KS, Adeoye O, Alexandrov AW, Ansari SA, et al. 2026 guideline for the early management of patients with acute ischemic stroke: a guideline from the American heart association/american stroke association. Stroke. 2026;57:e316–436. https://doi.org/10.1161/STR.0000000000000513.
5. Hsieh S-L, Hsiao C-H, Chiang W-C, Shin SD, Jamaluddin SF, Son DN, et al. Association between the time to definitive care and trauma patient outcomes: every minute in the golden hour matters. Eur J Trauma Emerg Surg. 2022;48:2709–16. https://doi.org/10.1007/s00068-021-01816-8.

6. Braverman A, Frenkel A, Schwarzfuchs D, Jaffe E, Bitan Y. Verbal and visual information exchange in EMS-to-ED patient handovers: an observational and attitudinal study. Int Emerg Nurs. 2026;84:101735. https://doi.org/10.1016/j.ienj.2025.101735.
7. Montano IH, De La Torre Díez I, López-Izquierdo R, Villamor MAC, Martín-Rodríguez F. Mobile triage applications: a systematic review in literature and play store. J Med Syst. 2021;45:86. https://doi.org/10.1007/s10916-021-01763-2.
8. Nguyen Q, Wybrow M, Burstein F, Taylor D, Enticott J. Understanding the impacts of health information systems on patient flow management: a systematic review across several decades of research. Kuo Y-H. editor PLOS One. 2022;17:e0274493. https://doi.org/10.1371/journal.pone.0274493.
9. Chaudhry B, Wang J, Wu S, Maglione M, Mojica W, Roth E, et al. Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. Ann Intern Med. 2006;144:742–52. https://doi.org/10.7326/0003-4819-144-10-200605160-00125.
10. Dong X, Ding F, Zhou S, Ma J, Li N, Maimaitiming M, et al. Optimizing an emergency medical dispatch system to improve prehospital diagnosis and treatment of acute coronary syndrome: nationwide retrospective study in China. J Med Internet Res. 2022;24:e36929. https://doi.org/10.2196/36929.
11. Hagiwara MA, Bremer A, Claesson A, Axelsson C, Norberg G, Herlitz J. The impact of direct admission to a catheterisation lab/CCU in patients with ST-elevation myocardial infarction on the delay to reperfusion and early risk of death: results of a systematic review including meta-analysis. Scand J Trauma Resusc Emerg Med. 2014;22:67. https://doi.org/10.1186/s13049-014-0067-x.
12. Jacobsson A, Kurland L, Höglund E. Direct in-hospital admission via ambulance (DIVA): a retrospective observational study. Int Emerg Nurs. 2020;52:100906. https://doi.org/10.1016/j.ienj.2020.100906.
13. Grotta JC, Yamal J-M, Parker SA, Rajan SS, Gonzales NR, Jones WJ, et al. Prospective, multicenter, controlled trial of mobile stroke units. N Engl J Med. 2021;385:971–81. https://doi.org/10.1056/NEJMoa2103879.
14. Asadiara E, Emami H, Safari M, Rabiei R. Development and usability evaluation of a mHealth application for prehospital triage. Sci Rep. 2025;15:21576. https://doi.org/10.1038/s41598-025-06952-4.
15. Martin T, Ranney M, Dorroh J, Asselin N, Sarkar I. Health information exchange in emergency medical services. Appl Clin Inf. 2018;9:884–91. https://doi.org/10.1055/s-0038-1676041.
16. Zhang Z, Brazil J, Ozkaynak M, Desanto K. Evaluative research of technologies for prehospital communication and coordination: a systematic review. J Med Syst. 2020;44:100. https://doi.org/10.1007/s10916-020-01556-z.
17. Fukaguchi K, Goto T, Yamamoto T, Yamagami H. Experimental implementation of NSER mobile app for efficient real-time sharing of prehospital patient information with emergency departments: interrupted time-series analysis. JMIR Form Res. 2022;6:e37301. https://doi.org/10.2196/37301.
18. Bhattarai HK, Bhusal S, Barone-Adesi F, Hubloue I. Prehospital emergency care in low- and middle-income countries: a systematic review. Prehospital Disaster Med. 2023;38:495–512. https://doi.org/10.1017/S1049023X23006088.
19. Kim M-J, Moon HS, Paek S-H. Agreement between EMS provider-assigned prehospital triage and initial emergency department triage in pediatric and adult EMS-transported encounters: a retrospective observational study. Alamri MS, editor. PLOS One. 2026;21:e0352969. https://doi.org/10.1371/journal.pone.0352969
20. Drenck N, Viereck S, Bækgaard JS, Christensen KB, Lippert F, Folke F. Pre-hospital management of acute stroke patients eligible for thrombolysis – an evaluation of ambulance on-scene time. Scand J Trauma Resusc Emerg Med. 2019;27:3. https://doi.org/10.1186/s13049-018-0580-4.
21. Zhang Z, Joy K, Harris R, Park SY. Characteristics and challenges of clinical documentation in self-organized fast-paced medical work. Proc ACM Hum Comput Interact. 2022;6:1–21. https://doi.org/10.1145/3555111.
22. Medrano NW, Villarreal CL, Mann NC, Price MA, Nolte KB, MacKenzie EJ, et al. Activation and on-scene intervals for severe trauma EMS interventions: an analysis of the NEMSIS database. Prehosp Emerg Care. 2023;27:46–53. https://doi.org/10.1080/10903127.2022.2053615.

23. Zhou L, Bao J, Setiawan IMA, Saptono A, Parmanto B. The mHealth app usability questionnaire (MAUQ): development and validation study. JMIR mHealth uHealth. 2019;7:e11500. https://doi.org/10.2196/11500.
24. Hsiao J-L, Chen R-F. Understanding determinants of health care professionals' perspectives on mobile health continuance and performance. JMIR Med Inf. 2019;7:e12350. https://doi.org/10.2196/12350.
25. Woo M, Mishra P, Lin J, Kar S, Deas N, Linduff C, et al. Complete and resilient documentation for operational medical environments leveraging mobile hands-free technology in a systems approach: experimental study. JMIR mHealth uHealth. 2021;9:e32301. https://doi.org/10.2196/32301.
26. Olaussen A, Abetz JW, Smith K, Bernard S, Gaddam R, Banerjee A, et al. Paramedic streaming upon arrival in emergency department: a prospective study. Emerg Med Australas. 2021;33:286–91. https://doi.org/10.1111/1742-6723.13618.
27. Reay G, Rankin JA, Smith-MacDonald L, Lazarenko GC. Creative adapting in a fluid environment: an explanatory model of paramedic decision making in the pre-hospital setting. BMC Emerg Med. 2018;18:42. https://doi.org/10.1186/s12873-018-0194-1.
28. Yoder A, Bradburn EH, Morgan ME, Vernon TM, Bresz KE, Gross BW, et al. An analysis of overtriage and undertriage by advanced life support transport in a mature trauma system. J Trauma Acute Care Surg. 2020;88:704–9. https://doi.org/10.1097/TA.0000000000002602.
29. Zhao Z, Li L, Zhang J, Shi X, Jiang Z, Jiang F, et al. Evaluating the real-world utility of four physiological scores for prehospital interventions in resource-limited EMS systems: a population-based study of 15,720 cases. BMC Health Serv Res. 2026;26:199. https://doi.org/10.1186/s12913-026-14019-x.

Figures

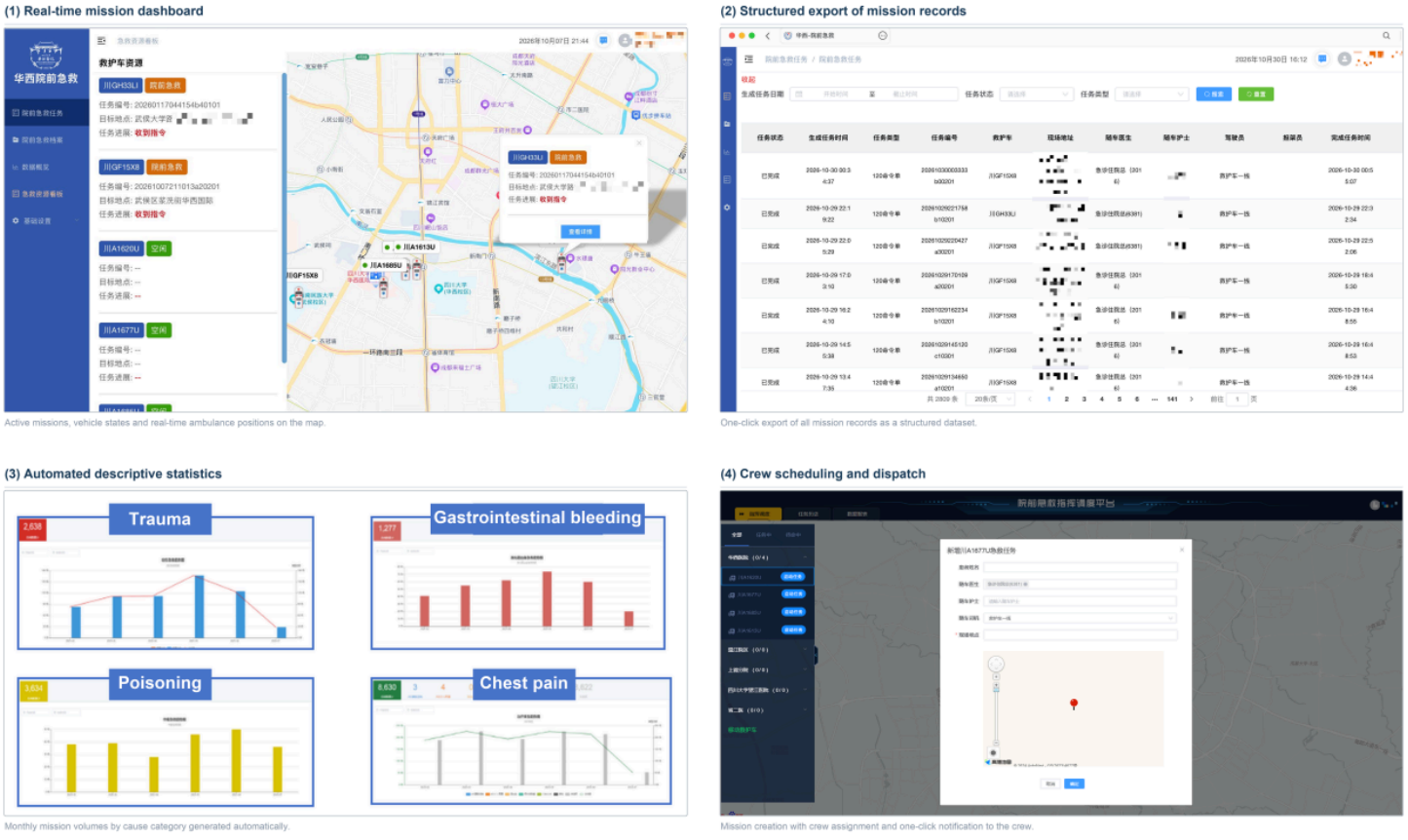

Figure 1

PC workstation interfaces of the West China integrated PHESS.

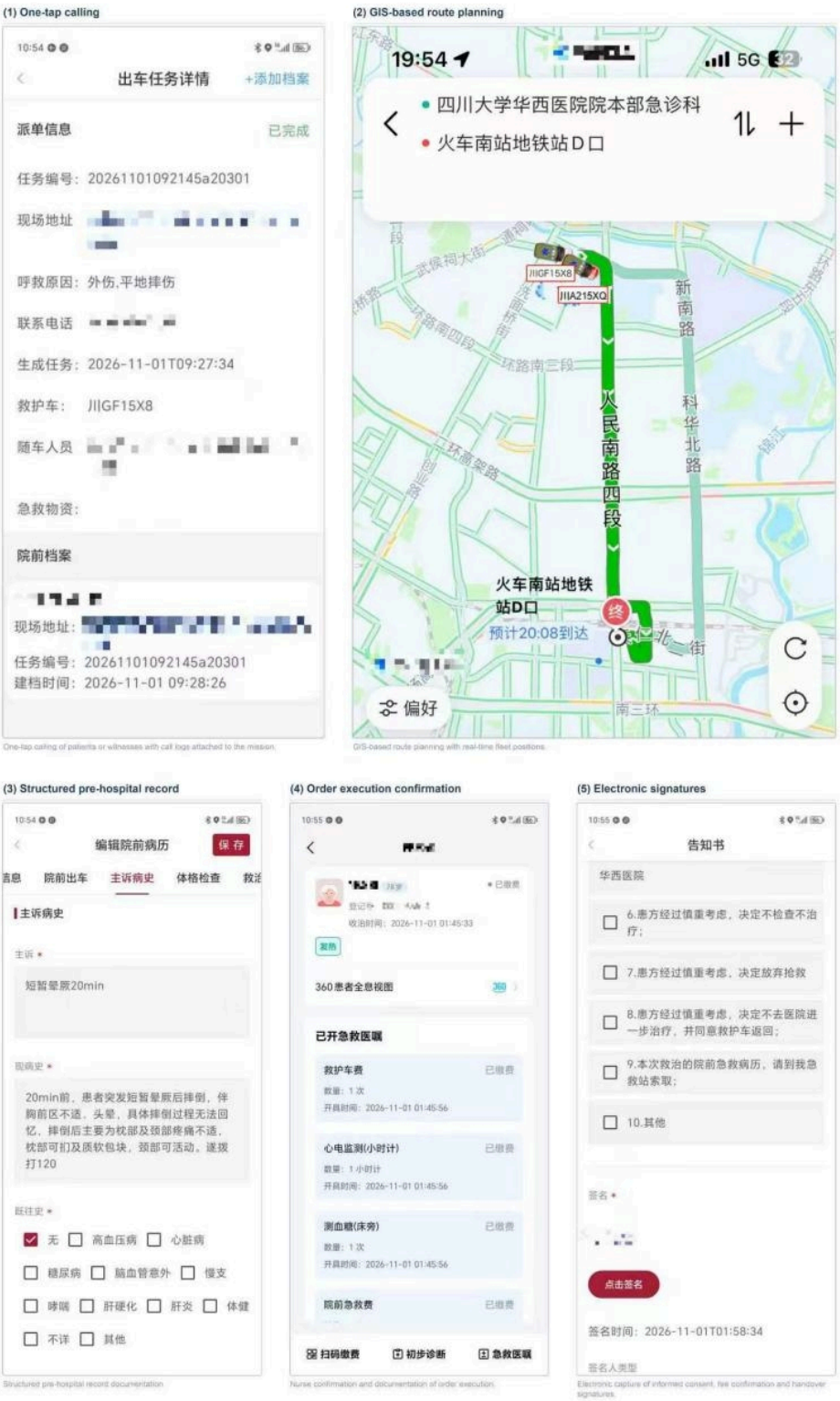

Figure 2

Android mobile client interfaces of the West China integrated PHESS.



Figure 3

Operational workflow of the West China pre-hospital emergency information platform, showing the ten steps from reception of the 120 command sheet to electronic handover. The screenshots show the deployed interface populated with synthetic demonstration records; the displayed records were created for demonstration and were not included in the analysis.

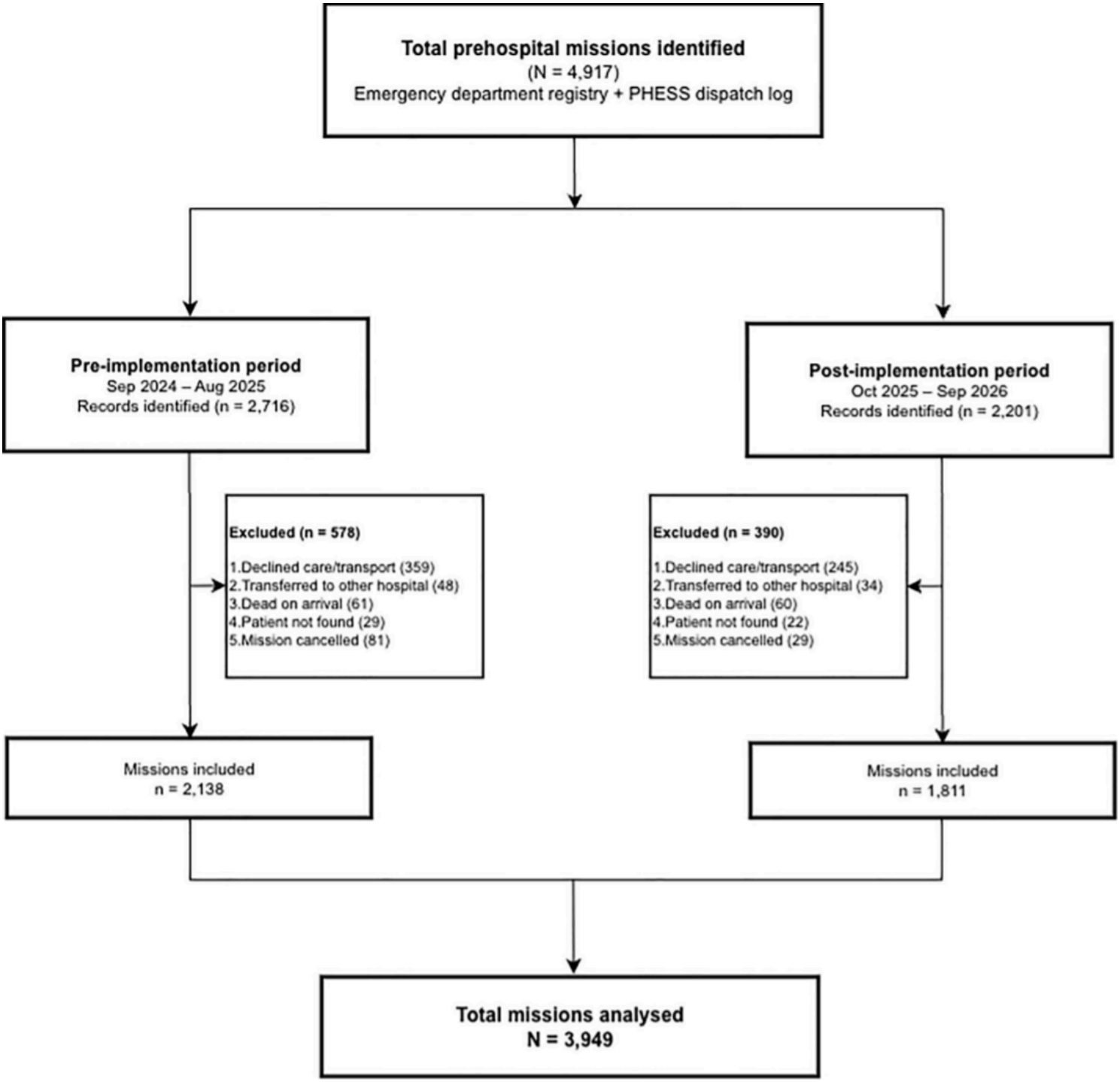

Figure 4

Flow diagram of mission selection and inclusion in the pre- and post-implementation periods.

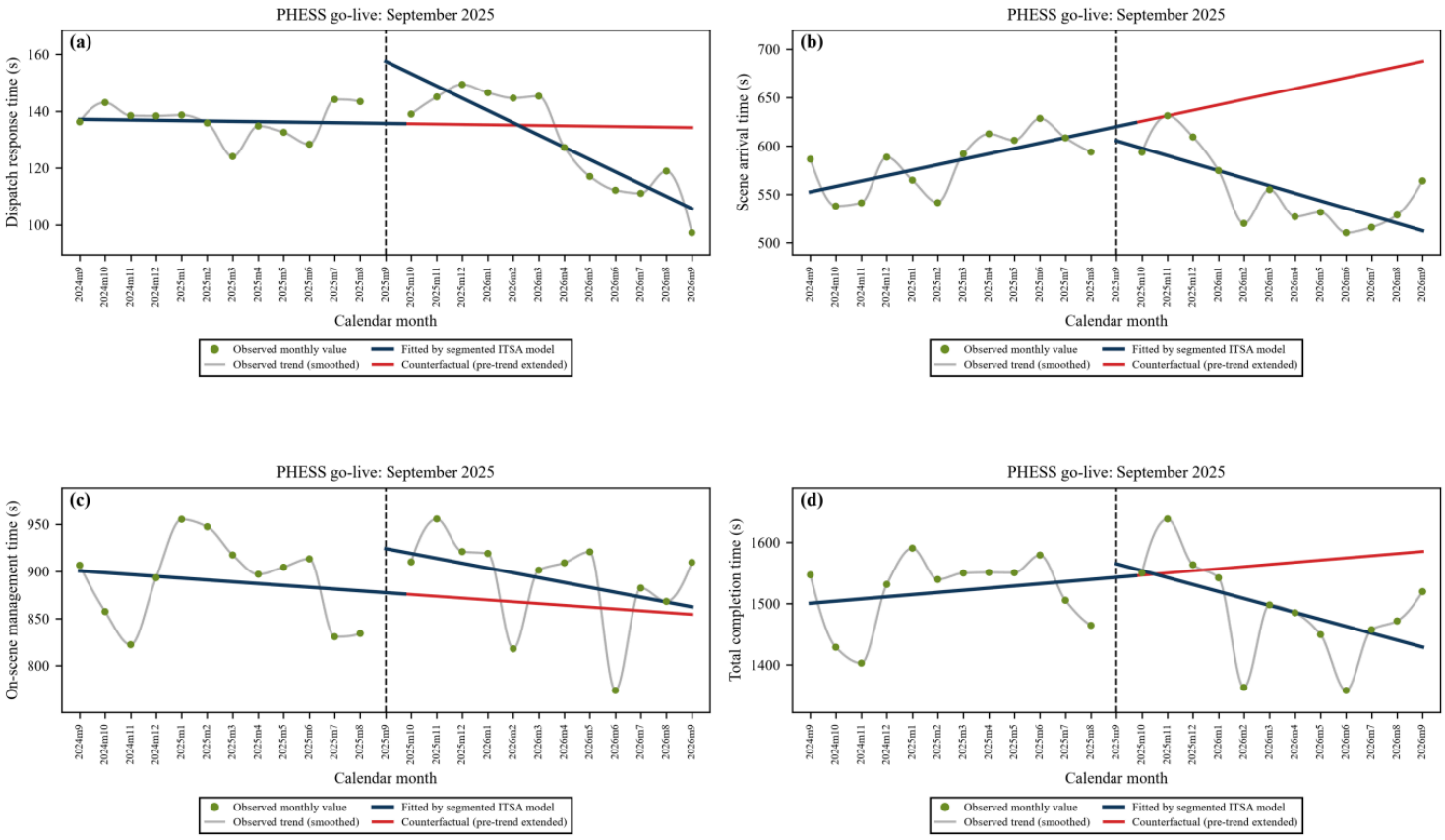

Figure 5

Interrupted time-series analysis of four pre-hospital emergency time metrics around PHESS implementation: (a) dispatch response time, (b) scene arrival time, (c) on-scene management time, and (d) total completion time.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1STROBEchecklist.docx
- Additionalfile2Supplementarytables.docx