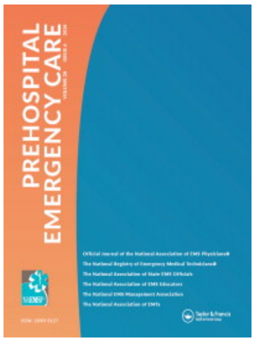

Prehospital Emergency Care

ISSN: 1090-3127 (Print) 1545-0066 (Online) Journal homepage: www.tandfonline.com/journals/ipec20

Survival Benefit of Helicopter Scene Response for Patients with an Injury Severity Score of at Least Nine: A Systematic Review and Meta-Analysis

Christie L. Fritz, Sarah Alice Thomas, Samuel M. Galvagno Jr. & Stephen H. Thomas

To cite this article: Christie L. Fritz, Sarah Alice Thomas, Samuel M. Galvagno Jr. & Stephen H. Thomas (2024) Survival Benefit of Helicopter Scene Response for Patients with an Injury Severity Score of at Least Nine: A Systematic Review and Meta-Analysis, Prehospital Emergency Care, 28:6, 841-850, DOI: 10.1080/10903127.2023.2232453

To link to this article: https://doi.org/10.1080/10903127.2023.2232453

PREHOSPITAL EMERGENCY CARE
2024, VOL. 28, NO. 6, 841–850
https://doi.org/10.1080/10903127.2023.2232453

ORIGINAL RESEARCH

Survival Benefit of Helicopter Scene Response for Patients with an Injury Severity Score of at Least Nine: A Systematic Review and Meta-Analysis

Christie L. Fritz[a], Sarah Alice Thomas[b], Samuel M. Galvagno, Jr.[c,d], and Stephen H. Thomas[a,d]

[a]EMS Division, Department of Emergency Medicine, Beth Israel Deaconess Medical Center & Harvard Medical School, Boston, Massachusetts; [b]Department of Molecular Microbiology & Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland; [c]Department of Anesthesiology, University of Maryland School of Medicine & R. Adams Cowley Shock Trauma Center, Baltimore, Maryland; [d]Barts and The London School of Medicine & Dentistry, Queen Mary University of London, London, UK

ABSTRACT

Background and Aim: Helicopter EMS (HEMS) is a well-established mode of rapid transportation for patients with need for time-sensitive interventions, especially in patients with significant traumatic injuries. Traditionally in the setting of trauma, HEMS is often considered appropriate when used for patients with "severe" injury as defined by Injury Severity Score (ISS) >15. This may be overly conservative, and patients with a lower ISS may benefit from HEMS-associated speed or care quality. Our objective was to perform a meta-analysis of trauma HEMS transports to evaluate for possible mortality benefit in injured cases defined by an ISS score >8, lower than the customary ISS cutoff of >15.

Methods: A broad search of the literature was performed including PubMed, EMBASE, SCOPUS, Cochrane Central Register of Controlled Trials, and Google Scholar from the years 1970 to 2022. The gray literature and reference lists of included publications were also examined. We included studies with the outcome of mortality in HEMS vs control in trauma transports from scene of injury for patients (adult or pediatric) with ISS > 8.

Results: Nine eligible studies were used in the final analysis: six in the primary analysis and three in sensitivity analysis due to patient overlap. All studies reported statistically significant survival benefit in HEMS compared to control group. The minimum survival odds ratio (OR) benefit observed was OR 1.15 (95% CI 1.06–1.25) and maximum was OR 2.04 (95% CI 1.18–3.57). Risk of bias tool (ROBINS-I) application yielded moderate to low risk of bias, mainly due to the observational nature of the studies included.

Conclusions: There was a statistically significant survival benefit in patients with ISS > 8 when HEMS was used over traditional ground ambulance transportation, although novel and more inclusive trauma triage criteria may be more appropriate in the future to guide HEMS utilization decision-making. Restricting HEMS to trauma patients with ISS >15 likely misses survival benefit that could be afforded to the subset of trauma patients with serious injury.

ARTICLE HISTORY
Received 8 March 2023
Revised 8 May 2023
Accepted 26 June 2023

Introduction

Helicopter emergency medical services (HEMS) have been dispatched to civilian trauma scenes since the 1970s. As compared to ground EMS, HEMS may offer patient-centered benefits attributable to a combination of clinical care and logistics.

There is a large body of evidence addressing the question of potential survival benefit associated with air medical scene response (1–6), but evidentiary quality has been questioned. There is a virtual absence of randomized controlled trial (RCT) literature; all outcomes publications but one (7) have used observational non-randomized study designs.

A 2015 Cochrane review (8) assessing adults (16+ years) with severe injury – defined as Injury Severity Score (ISS) at least 15 – concluded there was need for more high-quality evidence to enable definitive evaluation of HEMS benefits. In the years since that review, HEMS has continued to constitute a major part of prehospital care systems, with frequent deployment for patients ultimately identified as having ISS < 15.

An ISS > 15 corresponds to a single-system Abbreviated Injury Scale (AIS) score of at least 4, the lowest AIS representing a threat to life (per Baker's original description of the ISS) (9). An ISS > 8 cutoff corresponds to a decrement from a single-system injury of level AIS 4 ("severe, life-threatening, survival probable") to AIS 3 ("severe, not life-threatening") (9).

Given challenges of trauma triage in the real world (*i.e.*, frequent utilization of HEMS for lower-acuity patients), we chose to focus on studies of patients with an ISS cutoff lower than the ISS > 15 cutoff usually employed for HEMS evaluations. This review's objective is to assess whether HEMS scene response is associated with a survival benefit

CONTACT Christie L. Fritz cfritz@bidmc.harvard.edu.
Supplemental data for this article is available online at https://doi.org/10.1080/10903127.2023.2232453

when assessed over the set of trauma patients with severity of ISS > 8.

Methods

Review Framework

This systematic review with evidentiary quality rating (using Grading of Recommendations Assessment, Development, and Evaluation; GRADE) was conducted in accordance with recommendations of the Cochrane Collaboration (10). The structure follows the Preferred reporting items for systematic review and meta-analysis (PRISMA) version 2020 (11). The PRISMA checklist is provided in Supplement 1.

Bias was assessed using Cochrane's Risk of bias in non-randomized studies of interventions (ROBINS-I) (12) with GRADE co-implementation recommendations (13). ROBINS-I plotting employed McGuinness and Higgins' tool (https://mcguinlu.shinyapps.io/robvis/) (14).

PICO Structure: Participants – Intervention – Comparator – Outcome

Eligible studies were those reporting on mortality (*outcome*) associated with HEMS (*intervention*) *vs.* ground EMS (*comparator*) scene trauma response in adults or children with serious (ISS > 8) injury (*participants*). This review's ISS-based inclusion criterion was simple: ISS > 8.

Studies were excluded if they used a different ISS cutoff (*e.g.*, ISS > 15) or – with the exception of the typical exclusion criterion of scene death – any eligibility criteria other than ISS > 8. Any method of reaching ISS of 9 was accepted; there was no minimum AIS requirement for a given category. Studies were not included if they capped the ISS range due to restriction to single-system trauma (*e.g.*, AIS 3+ in one system with requirement for AIS < 3 in other systems) (15).

Any study design was eligible for inclusion. The only requirement was for HEMS to be compared against a control group. The control group could either be a same-setting ground EMS group or a historical cohort such as used in trauma and injury severity score (TRISS) methodology (16).

For inclusion in the review, and in keeping with methods commonly employed to evaluate HEMS (17–19), HEMS response was required to routinely comprise both helicopter response to the scene and air medical transport of the patient to the receiving hospital. Rare exceptions to this rule (such as for patients with cardiac arrest) did not result in study exclusion. The review excluded reports from systems where helicopters are commonly used to transport clinicians to scenes (and not transport patients to hospitals) (7, 20). The ground EMS comparator group in this review's study set consisted of patients for whom ground ambulances responded to trauma scenes and then transported patients (by ground) to the receiving centers.

Our focus on scene transports did allow for inclusion of studies including occasional interfacility transfers. Sensitivity analysis was planned to assess effects of including studies that focused on scene transports but also encompassed occasional interfacility cases.

There were no study exclusions for this review based on HEMS or ground EMS crew configuration. The judgment to not restrict the review to certain crew configurations inevitably added to study heterogeneity. The decision to not limit study assessment due to crew configuration was based on two factors.

The first crew factor is that staffing compositions for helicopter EMS and/or ground EMS within a given system may change from case to case. Studies often do not report or adjust for these variations.

The second crew consideration is that similarly credentialed crews in two locations (even within the same region) may have markedly different practice scopes for important interventions such as rapid sequence intubation (RSI). A particular region's non-physician crew may have an extended practice scope rendering such a non-physician crew more comparable to a second region's physician (rather than non-physician) model.

To properly incorporate crew configuration into multivariable analysis, it is the crew's practice scope capabilities (*e.g.*, RSI), rather than the nurse, paramedic, physician, or other credentials, that must be assessed. Such information is not routinely available in the available helicopter EMS literature published to date and is a foreseeable limitation.

Survival to hospital discharge was the primary outcome of interest. There were no secondary outcomes assessed. An individual study's time frame for assessment of the mortality endpoint (*e.g.*, in-hospital, 30-day) did not affect review eligibility.

Information Sources & Search Strategy

The review employed multiple sources for constituent studies, in an attempt to identify the maximum number of eligible publications. Study accrual included review of previously published HEMS study collections, complemented by execution of new literature database searches.

The first step was assessment of bibliographic publications summarizing HEMS outcomes studies from 1980 through 2016 (1–6). Annotated bibliographies' data were supplemented with references from other evidentiary summaries (8, 21, 22).

After assembling information from bibliographies, the next step in accruing HEMS research information was execution of database searches. We purposefully employed a recommended broad strategy (23) including search of PubMed, EMBASE, SCOPUS, and gray literature sources for terms including HEMS, air medical transport, and air ambulance. Details of the search strategy, including specific search terms, are provided in Supplement 2.

Data Collection and Management

Study data were entered into an Excel spreadsheet. Two authors (SAT and SHT) entered data; each study's data were cross-checked by both. In the event clarifications were needed regarding a particular study, the study's contact author was emailed.

If an individual study reported results from multiple modeling techniques, and one of those models incorporated propensity scoring (PS), this review prioritized PS findings. When such prioritization was executed, the study *n* was reported as the PS-matched case number as a more conservative estimate of precision.

Preliminary ROBINS-I and GRADE ratings were generated by two of this review's authors (SMG and SHT). These ratings were used as a basis on which consensus was reached among all study authors.

Analysis

Stata 17MP was used for statistical analysis, reporting, and graphing. Significance was defined as $p < 0.05$, and 95% confidence intervals (CIs) were used.

Study planning called for recording of studies' *n* and results for HEMS' mortality association, using the odds ratio (OR) as the main effect measure. For studies in which raw acuity-matched data were available – if HEMS and ground EMS patients were similar before adjustment, or if PS-matched data were available – classic 2×2 tables of exposure (HEMS) and outcome (mortality) were generated to serve as the basis to calculate the secondary effect-measure endpoint of absolute risk reduction (ARR) and its inverse, number needed to treat (NNT).

Meta-analysis was performed using the Sidik-Jonkman random effects model. Sidik-Jonkman modeling was selected based on Cochrane group recommendations for its use as preferred over other random effects techniques such as DerSimonian-Laird (24).

The review's studies encompassed a breadth of focus with regard to patients, crew configurations, and logistics. Clinical heterogeneity was known *a priori* to be high, and overall heterogeneity assessed by I^2 was projected to be considerable (8, 24). Study planning thus included attention to parsing statistical heterogeneity into clinical and methodological subcategories.

Small-study (mainly publication) bias was assessed with both funnel and Galbraith plots; Galbraith plots were also used to evaluate for outliers. Publication bias was also evaluated by comparing fixed-effect and random effects estimates (25). As recommended by the Cochrane group, small-study bias and heterogeneity were to be formally evaluated (using prediction intervals, Egger testing, and meta-regression) if the review accrued the required minimum of 10 studies (24, 25).

Pre-planned sensitivity analyses, which consisted of repeating meta-analysis using techniques such as leave-one-out plotting, revolved around two main issues. One such issue, studies' potential inclusion of interfacility transports, has been previously noted. The other issue was that of data overlap across different studies. The approach to sensitivity analysis regarding data-overlap was to constitute a separate study set, the sensitivity analysis set, comprising studies for which some data was encompassed within other (larger) studies included in the main analysis set.

We excluded any study for which data were fully subsumed within other, larger studies already included in the review. For sets of studies with any degree of incomplete data overlap, we retained only one representative publication for use in the review's main analysis set. The conservative approach was dictated by the lack of precision with regard to knowledge of which cases overlapped; this lack of knowledge precluded employment of preferred data-overlap solutions such as decoupled analysis or generalized-weight meta-estimation (26). Studies removed from the main analysis set due to data overlap were retained to constitute the review's sensitivity analysis set. Statistical calculations and graphing were executed on both the main study set and the combined main and sensitivity analysis sets.

Results

Study Identification and Screening

From an initial set of >24,000 records, screening yielded 12 publications identified as potentially meeting inclusion criteria. From this set of 12, three were dropped due to their cases being fully subsumed within other (included) studies. There were thus nine studies (27–35) that underwent review for data overlap.

Within the review's set of nine studies there were a total of five for which partial data overlap was an issue. Two studies (28, 34) had overlapping data from a German trauma registry, and three studies (32, 33, 35) had overlapping data from a Japanese database. The process of selecting a representative study from each of the two groups is detailed in Supplement 2.

After selection of a study from each of the two sets of partial data overlap publications, this review's main analysis set was constituted with six studies (27–31, 33). The review's sensitivity analysis study set comprised three reports (32, 34, 35) characterized by partial data overlap with studies that were already included in the review's main analysis set.

A PRISMA (11) diagram is shown in Figure 1. Supplement 2 provides details on the search, review, and screening of records that culminated in constitution of the main analysis and sensitivity analysis study sets.

Study Characteristics

Studies comprising this review set (see Table 1) were all published in English. All were retrospective reviews of trauma databases that were regional (27, 30) or national (28, 29, 31–35) in scope. Varying definitions of adult *vs.* pediatric patients were used (thus preventing subgroup analysis).

The main study set did not contain any studies that accrued interfacility transports. The sensitivity analysis study set included one such study (Jitsuiki (32) with 21% interfacility transports).

For the primary analysis set of six studies, overall mortality median was 10.9% with interquartile range (IQR) 6.6% to 18.1%; corresponding mean ± standard deviation (SD) was $12.9 \pm 7.8\%$. Pediatric studies (29, 31) and the French Alpine

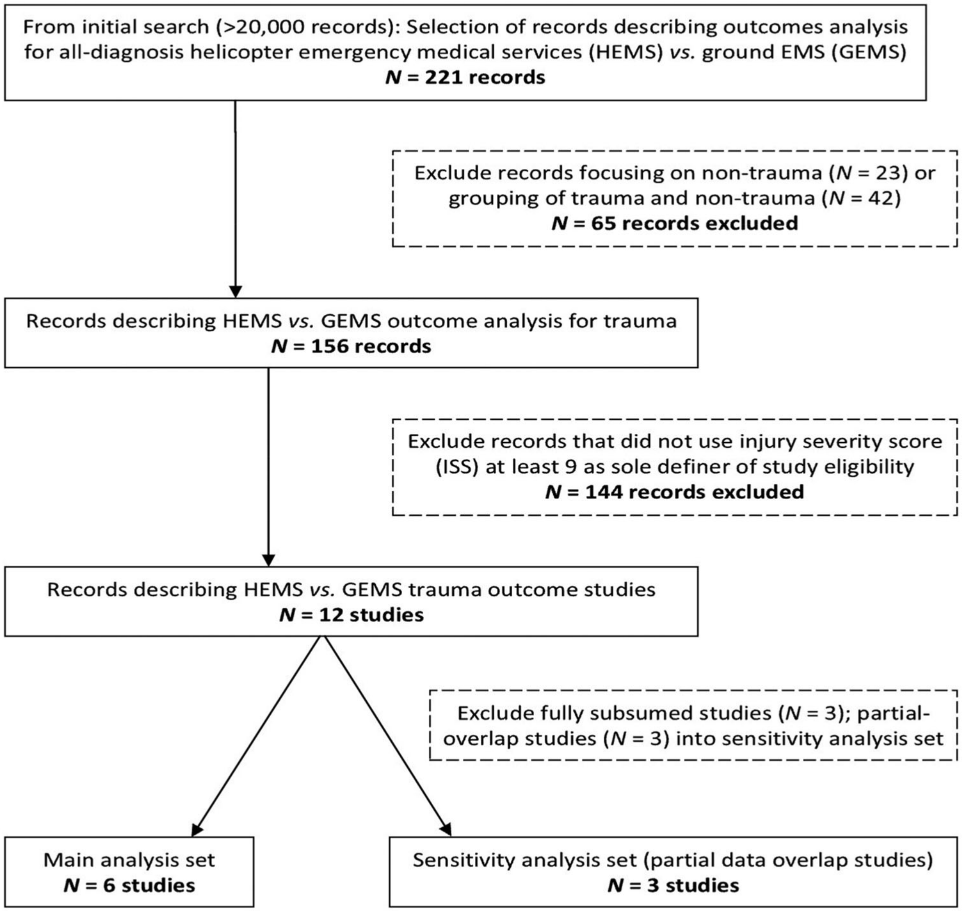

Figure 1. Screening and selection of publications for inclusion in review.

Table 1. Studies comprising this review's main ($N = 6$) and sensitivity ($N = 3$) analysis sets.

First author & publication year	Patient geography, data source, and ages included	Accrual years	Ground EMS *n* (ISS mean)	HEMS *n* (ISS mean)
Studies ($N = 6$) in the main analysis set				
Ageron (2020) (27)	Regional database: French Alps (TRENAU); all ages	2009-2017	5253 (16)	3524 (16)
Andruszkow (2016) (28)	National data base: Germany (TraumaRegister DGU); all ages	2002-2012	35974 (21.7)	16307 (24.8)
Blasius (2021) (29)	National data base: Germany (TraumaRegister DGU); <16 years	2013-2017	1929 (18.1)	826 (21.4)
Davis (2005) (30)	Regional database: San Diego County Trauma Registry; >14 years	1987-2003	7295* (25.5)	3017 (28.4)
Englum (2017) (31)	National database: USA (NTDB); <19 years	2007-2011	26808* (not reported)	10565 (not reported)
Kushida (2021) (33)	National database: Japan (JTDB); all ages	2004-2019	7328^ (27.2)	7328 (27.9)
Studies ($N = 3$) in the sensitivity analysis set (data overlap with main-set studies)				
Hosomi (2022) (35)	National database: Japan (JTDB); all ages	2004-2018	1858^ (26.1)	1855 (26.1)
Jitsuiki (2022) (32)	National database: Japan (JTDB); all ages	2004-2019	1774^ (27.7)	1732 (27.7)
Weinlich (2019) (34)	Frankfurt-area data from TraumaRegister DGU; all ages	2009-2013	915 (not reported)	81 (not reported)

Abbreviations: EMS - emergency medical services; HEMS – helicopter EMS; ISS – Injury Severity Score; TRENAU – Trauma system of the Northern French Alps Emergency Network; DGU – Deutsche Gesellschaft fur Unfallchirurgie (German Trauma Society); NTDB – National Trauma Data Bank of the American College of Surgeons.

*Propensity-score (PS) methods used; PS-matched *n* not explicitly reported but 1:1 matching was employed (thus effective GEMS *n* was 10,565 if all HEMS cases were matched).

^PS methods used, study *n* is number of patients in each PS-matched HEMS and ground EMS group.

study (27) had the lowest mortality (6–8%) whereas the highest death proportion (25%) was seen in the study (30) accruing patients based on AIS_{Head} 3+. Mortality proportions in the three-study sensitivity analysis set had a median (IQR) of 17.8 (12.6 to 20.0) and mean (± SD) of $16.8 \pm 3.8\%$.

All studies assessed HEMS outcome *vs.* same-setting ground EMS controls. No reports in this review used TRISS predictions as the sole comparator against which HEMS outcomes were assessed.

Adjustment for confounding by acuity was executed in all studies using multivariable logistic regression. PS techniques were used in three studies (Davis (30), Englum (31), Kushida (33)) in the main analysis set and two (Hosomi (35), Jitsuiki (32)) in the sensitivity analysis set.

While all studies used logistic regression, only one (Ageron (27)) reported postestimation assessments of both calibration (goodness of fit) and discrimination (*c* statistic). Three other studies reported results for either calibration or discrimination. In the main analysis set, Davis (30) reported calibration and Andruszkow (28) reported discrimination. In the sensitivity analysis set, Hosomi (35) reported discrimination.

With regard to missing data, one study (Ageron (27)) employed multiple imputation and sensitivity analyses. Another study (Englum (31)) performed limited sensitivity analysis focusing on one covariate (transport interval). The remaining studies in the main and sensitivity analysis sets executed complete-case analyses.

Bias Assessment

Since all studies were non-randomized designs, the application of ROBINS-I was known *a priori* to be unlikely to result in an overall assignation of "low risk of bias" status (36). ROBINS-I was therefore executed to determine whether there was moderate, serious, or critical risk of bias.

The detailed multi-stage ROBINS-I table results are provided in Supplement 3. Results are provided both for the main analysis set and the sensitivity analysis set. Similarities in design across studies in this review translated into concomitant similarity in bias judgments, which are reported in detail in Supplement 3. ROBINS-I table highlights are discussed in this section.

The first and most important issue is that, regardless of methodology, observational HEMS scene trauma mortality analyses are at high risk for residual confounding and selection bias. These sources of bias may be mitigated with methodology. Geographic confounding can be addressed with cluster analysis, missing data can be addressed with multiple imputation and sensitivity analysis, and covariate-pattern confounding can be addressed with PS. Nevertheless, at best, mitigation is partial. Even with employment of these advanced methods, non-randomized studies have limited protection against confounding and selection bias (36).

Some forms of bias were minimized by the narrow focus of this review. The objective and fixed nature of both transport mode and mortality translate into low likelihood of misclassification. All studies reported the mortality endpoint as survival status at hospital discharge.

One issue, that of on-scene arrest, deserves mention as it could theoretically cause change of intervention group. Clinical practice protocols tend to favor ground EMS transport of traumatic arrest patients, even if HEMS is on scene. Inclusion of these patients is problematic since a majority die, biasing results against ground EMS. However, all but two studies (30, 31) explicitly excluded patients who died on scene or were dead on hospital arrival.

Since patients who are perceived as "peri-arrest" could imaginably be triaged to ground EMS (in case of concerns over need for intra-transport cardiopulmonary resuscitation) or HEMS (in case of prioritization of rapid transport for moribund patients), neither bias direction nor degree can be reliably predicted in peri-arrest.

Other than for patients in arrest, acuity stratification was accomplished using ISS. ISS is broadly accepted, although imperfect (*e.g.*, in terms of accuracy in penetrating trauma). Adjustment for blunt *vs.* penetrating injury mechanism was executed in all studies covered in this review.

The overall ROBINS-I summary ratings for the main analysis set are shown in Figure 2. Due largely to the simplicity and concrete nature of a dichotomously assessed intervention (HEMS or ground EMS) and outcome (lived or died), there was low bias risk associated with many ROBINS-I domains. It was judged unlikely that results of the studies were affected substantively by participant selection, intervention classification or deviation, outcome classification, or selective outcome reporting. As shown in Supplement 3, the overall ROBINS-I findings for the sensitivity analysis study set mirrored the findings for the main analysis study set.

This review's interpretation of the summary data is that the available studies provide results that, while reasonable for studies using non-randomized designs, are clearly inferior to data that could be generated from RCT designs. This interpretation was not changed with the additional consideration of the partial data overlap studies in the sensitivity analysis set.

HEMS Effect on Survival

All studies in both sets reported statistically significant survival increments in their predefined endpoints, the OR effect estimates. Inter-study variation in OR reporting methods is discussed in detail in Supplement 4. The remainder of this section presents results for PS models (where both PS and non-PS ORs were available).

Of the studies in the main analysis set, the minimum and maximum HEMS survival OR benefits were reported by Kushida (33) (OR 1.15, 95% CI 1.06–1.25) and Blasius (OR 2.04, 95% CI 1.18–3.57). Summary results of HEMS evaluation for effect on survival odds are shown as a modified forest plot depicted in Figure 3. While the null hypothesis of no HEMS effect on survival was rejected ($p < 0.01$), the forest plot as shown in Figure 3 is modified to suppress display of a pooled (summary) effect estimate.

It is important to note that pooled estimation was judged inappropriate given the calculated random effects model's I^2 of 89%. A hypothesis-generating pooled estimate, provided only as a preliminary finding, was calculated as OR 1.45 (95% CI 1.13–1.86). Further details regarding the pooled effect estimate and modeling statistics are reported in Supplement 4.

After forest plotting and identification of heterogeneity, additional analyses were performed in an effort to identify potential explanations of heterogeneity. These analyses (see Supplement 4) suggested neither outlier effects nor small-study (publication) bias. There were insufficient studies within a given subgroup (*e.g.* pediatric) to enable focused analysis on such groups.

The HEMS survival ORs discussed in this review are not to be interpreted as approximations of the risk ratio, since the majority of studies in this meta-analysis violated the rare-disease assumption. In fact, given the difficulties with meaningful interpretation of the OR, experts (24) recommend that OR results be converted to a more meaningful metric. Using standard formulae (37) for conversion of ORs to NNT values at varying levels of overall (baseline) trauma mortalities, an NNT table has been included in Supplement 4. The table's specificity to the current meta-analysis lies

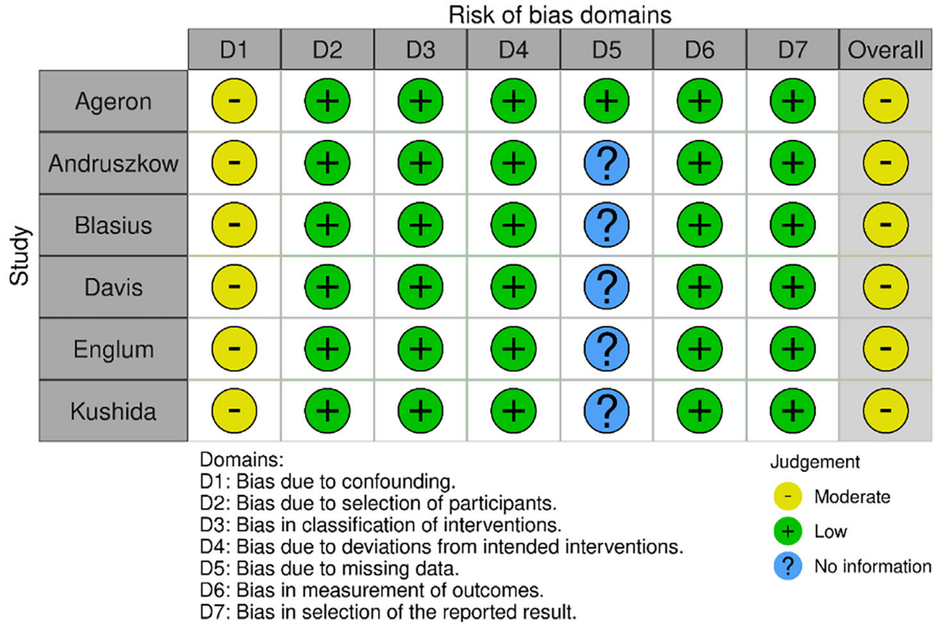

Figure 2. ROBINS-I summary plot (14).

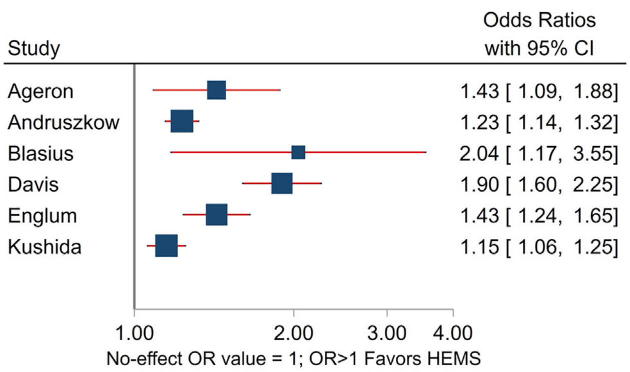

Figure 3. HEMS association with survival odds ratio (OR) with confidence interval (CI).

only in the choice of ranges for baseline mortality risk and range of ORs (both were chosen to match the ranges seen in this meta-analysis).

Secondary Effect Metrics

In addition to the primary endpoint metric of OR, a few studies reported other measures of HEMS effect on mortality. These secondary results were not subject to meta-analysis since they were reported with insufficient frequency and consistency. For example, standardized mortality was calculated in two studies, but the calculations used two different standardization populations and methods (TRISS and Revised Injury Severity Classification) (27, 29).

The review's initial plan called for use of individual study data reporting to generate classic 2×2 tables of exposure (HEMS) and outcome (survival) to facilitate calculation of effect measures such as ARR and NNT. Upon examination of the review set's publications (including supplementary information), there was insufficient information for generation of these data. Limited results, intended to facilitate future studies (*e.g.*, sample size calculations) are presented in Supplement 4).

Discussion

This review and meta-analysis assessed observational (non-RCT) studies of HEMS trauma scene response and transport, for patients with ISS > 8. The selected ISS cutoff, which was lower than the traditionally used ISS > 15 cutoff for "severe" injury, generally corresponds to (but does not actually require) an AIS of at least 3 in at least one system. The selection of the lower-than-usual ISS cutoff was deliberate, given widely acknowledged challenges with trauma triage; HEMS is used for a substantial number of patients whose ISS does not exceed 15.

There are at least two essential findings in this review. First, in every study that was identified as meeting inclusion criteria for this review, HEMS was associated with a statistically significant survival benefit. Second, due to differences in the analyzed studies' case mix, prehospital crews, and logistics, there is sufficient clinical heterogeneity to prevent generation of a pooled effect estimate.

One important ramification of the study heterogeneity is that it appears that the evidence base is sufficiently disparate as to not support definition of a given ISS cutoff for HEMS use. Additionally, there may be no single ISS cutoff applicable to all populations. It is possible, for example, that different ISS cutoffs may be appropriate for patients in different age groups, or with single- *vs.* multi-system trauma.

Whereas we selected a cutoff of ISS > 9 cutoff in an effort to address the realities and challenges of field triage, the literature does not provide sufficient support to precisely define a minimum ISS at which HEMS becomes useful. Logically, there is a certain point at which the ISS is so low that HEMS cannot improve survival (though HEMS may still be necessary for logistical indications). It is not clear that such a cutoff lies at the commonly used level of ISS 15.

Given the difficulties of trauma triage, it is quite common for patients in the ISS range of 9–15 to undergo air transport, since the ISS can only be accurately calculated after all injuries are diagnosed and coded. Uncertainty as to whether HEMS is useful in the ISS 9–15 range was explicitly identified by one of this review's constituent studies (28) as the reasoning for their selection of the acuity cutoff. Nonetheless, as recommended by the current trauma triage guidelines in the USA, certain individual injury patterns warrant expeditious transport to the highest level trauma center (38). Some of these injuries may be isolated, but severe, resulting in an overall ISS <9. In fact, at least one study (28) has identified HEMS' maximal benefit (a survival OR of 0.66 with 95% CI 0.49–0.88) as occurring within the range of ISS 9–15.

The selected ISS cutoff for this review was somewhat arbitrary, but it was guided by the relatively few cutoffs (other than ISS > 15) encountered in the HEMS literature. One study in this review assessed cases across all ISS levels (Weinlich (34) identified HEMS survival OR of 4.8 with 95% CI of 1.4 to 16.7), but most HEMS studies restrict focus to ISS > 15. Three studies have found significant HEMS survival benefit when accruing cases based on ISS cutoff of 12 (39–41).

A core challenge in the HEMS triage debate revolves around the use of ISS – an *a posteriori* score assigned after injuries are fully characterized – to retrospectively adjudicate transport mode decisions. Due to its retrospective nature, ISS seems misapplied when it is used to understand which patients benefit from HEMS. While it is often used in the literature, ISS at any cutoff is not useful for prospective identification of HEMS-eligible cases. The importance of our findings is perhaps not so much whether the ISS cutoff should be lowered, as to whether there is an identifiable cutoff.

The question of HEMS triage, while critically important, falls outside of this review's scope. However, the triage debate may be usefully informed by our findings: this review of studies using a lower-than-usual ISS > 8 cutoff found that HEMS significantly improved survival in every identified publication.

Two post-intervention factors were likely to influence outcome: crew configuration and the level of receiving trauma center. As discussed previously, crew configuration was not a focus of this review and was not analyzed as an independent variable. HEMS may be viewed from an epidemiological standpoint as a mediator variable – a variable that explains the dependent (outcome) variable and other independent variables (*i.e.*, age, ISS). Hence, we sought to elucidate the presence, not mechanism, of any HEMS effect on trauma survival. Crew configuration information is reported in the ROBINS-I table for informational purposes.

As is the case with crew configuration, receiving center effects are linked to transport mode (HEMS often translates into access to higher-level care). There was adjustment for receiving center level in all of this review's studies except one (33). Such adjustment would plausibly bias against HEMS, given the correlation between HEMS response and highest-level care.

An additional bias is that of missing data. As outlined in the ROBINS-I tool (Supplement 3), only Ageron (27) meticulously evaluated for missing data and used both multiple imputation and sensitivity analyses to demonstrate low likelihood of missing-data bias. For the other studies in the review, there was insufficient information for precise adjudication as to degree of missingness or its potential biasing effect. In addition to the issue of quantity of missing data, there is the issue of the mechanism of missingness. There can be no assurance that data are missing at random, so the missing-data problem is acknowledged as an unresolved limitation for these studies and this review.

There is an additional literature shortcoming that, while not clearly biasing studies, could affect findings. The studies we reviewed have publication dates that stretch back years or even decades. It is acknowledged that HEMS, ground EMS, and hospital care have all evolved since many of these studies were conducted.

The relatively poor certainty ratings associated with GRADE evaluation of non-randomized study designs has been addressed in detail by GRADE experts (13). In short, it is acknowledged that it is unusual (although not impossible) for a non-randomized design to be assigned a "low" or "moderate" (less likely "high") GRADE rating. We note that another recent HEMS trauma meta-analysis of two studies also assigned a "very low" GRADE certainty (42).

With the assumptions that RCT data are not forthcoming and that this review's effect estimates are in the right direction, two developments could result in improvement in GRADE certainty. First, more meticulous methodology can reduce potential bias (*e.g.*, handling missing-data problems with multiple imputation and sensitivity analyses) and aim to remove all plausible bias against ground EMS. Second, HEMS triage could be improved so that the survival OR exceeds the "large effect" threshold (of 2.0) for increasing GRADE's certainty rating.

Given recognized importance of logistics in HEMS trauma triage (27), it is noteworthy that HEMS had a positive association with survival in both (27, 31) of this review's studies that included adjustment for prehospital interval. In fact, as has often been the case in the HEMS scene trauma evidence (ever since Baxt (43) found no HEMS time benefit in the first major HEMS trauma outcomes study) air transport tended to be associated with longer, rather than shorter, prehospital intervals. Detailed information on each study's handling of (and findings regarding) logistics information is found in Supplement 4).

Along with crew configuration and time savings, one of the most often-mentioned issues with the HEMS evidence base is the paucity of functional status (as compared to mortality) outcomes data. With that paucity in mind, this review's *a priori* plan was to focus on survival. One study in this review's main analysis set (Davis (30)) and two in the sensitivity analysis set (Hosomi (35) and Weinlich (34)) assessed functional outcome. For this analysis, the mortality endpoint was the focus. All three results favored HEMS, but statistical significance was only achieved in two reports (30, 34). Detailed information on the functional outcomes results is included in Supplement 4.

One overall lesson from this review is that ISS > 15, the ISS-based cutoff most commonly used in HEMS outcomes studies, may be too restrictive. There was no suggestion that HEMS survival benefit in ISS > 8 was simply due air transport's improving outcomes at the higher end (*e.g.* ISS > 15) of the acuity grouping. Only one study in this review (Hosomi (35), who dichotomized ISS at a cutoff of 50) suggested higher HEMS benefit at the top end of the acuity range. Other studies we assessed found highest HEMS benefit at the lowest ISS category (Andruszkow (28)) or the lowest AIS_{Head} of 3 (Davis (30)). In pediatric patients, Englum (31) reported lack of success in identifying an ISS cutoff (within the ISS > 8 group) below which HEMS' survival benefit dropped.

The study findings noted above should be interpreted as potentially suggestive, not confirmatory, that HEMS benefits were seen at the ISS range's lower (as well as the higher) end. The findings of benefit in the studies we reviewed could still have been driven by benefit in the higher ISS cases. The central tendencies of ISS in the reviewed studies were relatively high (mid-20s), and those studies reporting a proportion of ISS > 15 cases found such proportions were relatively high (40–75%). Furthermore, the Davis (30) study's group of AIS of 3 was not an isolated head injury; there could have been other injuries rendering these cases' ISS well above 15. Finally, the Englum (31) study's lack of finding of a statistically significant difference in HEMS benefit across the ISS range – a finding based on an interaction term – was not consistent with the study's forest plot; that plot indicated mortality benefit only in higher ISS categories.

While our findings support reconsideration of use of ISS > 15 as the "standard" cutoff to define HEMS survival benefit, this review has limitations in addition to those already mentioned. It is possible that some studies' analysis of ISS > 8 subgroups were not detected by our search, and the relatively low number of included studies precluded formal analysis of publication bias. For the studies that were identified, evidentiary quality was not always optimal, even considering drawbacks to non-randomized study methodologies. Finally, high I^2 may be explainable by clinical (rather than methodological) heterogeneity, but the inability to estimate a pooled effect is a limitation to this meta-analysis.

A reasonable conclusion from this meta-analysis would be two-fold. First, all of the available evidence suggests statistically significant HEMS benefit may be accrued in trauma cases other than the usually defined ISS > 15 cutoff. Second, the degree to which HEMS improves survival in patients with a lower cutoff (ISS > 8) patients depends on both regional and patient factors. The evidence suggests that HEMS for scene trauma patients with ISS > 8 cannot be definitely stated to improve outcome, but that HEMS use in this population may be associated with an improved survival OR in the range of 1.15 to 2.04 (*i.e.*, this study set's least-favorable to most-favorable estimates). Just as there is no "standard" HEMS vs ground EMS situation – different settings will differ with respect to patients, crews, logistics, or receiving centers – there is no "standard" HEMS-associated mortality effect. Rather than suggesting that ISS > 8 should be a new cutoff for defining HEMS benefit, the results of this analysis may be most useful in suggesting that it may be time to discontinue rigorous application of an ISS cutoff to define HEMS appropriateness.

Acknowledgments

Thanks are due to Juntendo University's Dr. Youichi Yanagawa for alacrity in providing information regarding his group's analyses of data from the Japan Trauma Data Bank.

Disclosure Statement

The authors report there are no competing interests to declare.

ORCID

Christie L. Fritz http://orcid.org/0000-0002-0206-6164
Sarah Alice Thomas http://orcid.org/0000-0001-7923-5416
Samuel M. Galvagno, Jr. http://orcid.org/0000-0001-5563-4092
Stephen H. Thomas http://orcid.org/0000-0003-2833-5921

References

1. Thomas SH, Cheema F, Wedel SK, Thomson D. Trauma helicopter emergency medical services transport: annotated review of selected outcomes-related literature. Prehosp Emerg Care. 2002; 6(3):359–71. In eng). doi:10.1080/10903120290938508.
2. Thomas SH. Helicopter emergency medical services transport outcomes literature: annotated review of articles published 2000-2003. Prehospital Emergency Care: official Journal of the National Association of EMS Physicians and the National Association of State EMS Directors. 2004;8(3):322–33. In eng).
3. Thomas SH. Helicopter EMS transport outcomes literature: annotated review of articles published 2004-2006. Prehosp Emerg Care. 2007;11(4):477–88. In eng doi:10.1080/10903120701537097.
4. Brown BS, Pogue KA, Williams E, Hatfield J, Thomas M, Arthur A, Thomas SH. Helicopter EMS Transport Outcomes Literature: annotated Review of Articles Published 2007-2011. Emerg Med Int. 2012;2012:876703. doi:10.1155/2012/876703.
5. Lee EJ, Arthur AO, Thomas SH. Helicopter EMS outcomes literature: annotated review of articles published 2012-2013. BJMMR. 2014;4(19):3620–49. doi:10.9734/BJMMR/2014/9127.
6. Thomas SH, Blumen I. Helicopter Emergency Medical Services Literature 2014 to 2016: lessons and Perspectives, Part 2-Nontrauma Transports and General Issues. Air Med J. 2018; 37(2):126–30. doi:10.1016/j.amj.2017.10.005.
7. Garner AA, Mann KP, Fearnside M, Poynter E, Gebski V. The Head Injury Retrieval Trial (HIRT): a single-centre randomised controlled trial of physician prehospital management of severe blunt head injury compared with management by paramedics only. Emerg Med J. 2015;32(11):869–75. In eng doi:10.1136/emermed-2014-204390.
8. Galvagno SM, Sikorski R, Hirshon JM, Floccare D, Stephens C, Beecher D, Thomas S. Helicopter emergency medical services for adults with major trauma. Cochrane Database Syst Rev. 2015; 2015(12):Cd009228. In eng doi:10.1002/14651858.CD009228.pub3.
9. Baker SP, O'Neill B, Haddon W, Jr., Long WB. The injury severity score: a method for describing patients with multiple injuries

and evaluating emergency. J Trauma. 1974;14(3):187–96. https://www.ncbi.nlm.nih.gov/pubmed/4814394).
10. Cochrane Handbook for Systematic Reviews of Interventions (version 6.3). In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane; 2022. Available from www.training.cochrane.org/handbook.
11. Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ. 2021;372:n160. doi:10.1136/bmj.n160.
12. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, Henry D, Altman DG, Ansari MT, Boutron I, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919. doi:10.1136/bmj.i4919.
13. Schünemann HJ, Cuello C, Akl EA, Mustafa RA, Meerpohl JJ, Thayer K, Morgan RL, Gartlehner G, Kunz R, Katikireddi SV, et al. GRADE guidelines: 18. How ROBINS-I and other tools to assess risk of bias in nonrandomized studies should be used to rate the certainty of a body of evidence. J Clin Epidemiol. 2019;111:105–14. doi:10.1016/j.jclinepi.2018.01.012.
14. McGuinness L, Higgins J. Risk-of-bias VISualization (robvis): an R package and Shiny web app for visualizing risk-of-bias assessments. Res Syn Meth. 2021;12(1):55–61. doi:10.1002/jrsm.1411.
15. Aiolfi A, Benjamin E, Recinos G, De Leon Castro A, Inaba K, Demetriades D. Air Versus Ground Transportation in Isolated Severe Head Trauma: a National Trauma Data Bank Study. J Emerg Med. 2018;54(3):328–34. doi:10.1016/j.jemermed.2017.11.019.
16. Boyd CR, Corse KM, Campbell RC. Emergency interhospital transport of the major trauma patient: air versus ground. J Trauma. 1989;29(6):789–93. discussion 793-4. (In eng). doi:10.1097/00005373-198906000-00015.
17. Berns KS, Hankins DG, Zietlow SP. Comparison of air and ground transport of cardiac patients. Air Med J. 2001;20(6):33–6. In eng). doi:10.1016/S1067-991X(01)70008-9.
18. Hesselfeldt R, Steinmetz J, Jans H, Jacobsson M-LB, Andersen DL, Buggeskov K, Kowalski M, Praest M, Øllgaard L, Höiby P, et al. Impact of a physician-staffed helicopter on a regional trauma system: a prospective, controlled, observational study. Acta Anaesthesiol Scand. 2013;57(5):660–8. In eng doi:10.1111/aas.12052.
19. Zakariassen E, Østerås Ø, Nystøyl DS, Breidablik HJ, Solheim E, Brattebø G, Ellensen VS, Hoff JM, Hordnes K, Aksnes A, et al. Loss of life years due to unavailable helicopter emergency medical service: a single base study from a rural area of Norway. Scand J Prim Health Care. 2019;37(2):233–41. doi:10.1080/02813432.2019.1608056.
20. Smith CA, Hardern RD, LeClerc S, Howes RJ. Prehospital analysis of northern trauma outcome measures: the PHANTOM study. Emerg Med J. 2019;36(4):213–8. doi:10.1136/emermed-2017-206848.
21. Galvagno SM, Jr., Thomas S, Stephens C, et al. Helicopter emergency medical services for adults with major trauma. The Cochrane Database of Systematic Reviews. 2013;(3):Cd009228. In eng doi:10.1002/14651858.CD009228.pub2.
22. Thomas SH, Brown KM, Oliver ZJ, Spaite DW, Lawner BJ, Sahni R, Weik TS, Falck-Ytter Y, Wright JL, Lang ES, et al. An Evidence-based Guideline for the air medical transportation of prehospital trauma patients. Prehosp Emerg Care. 2014;18 Suppl 1:35–44. In eng doi:10.3109/10903127.2013.844872.
23. Bramer WM, Rethlefsen ML, Kleijnen J, Franco OH. Optimal database combinations for literature searches in systematic reviews: a prospective exploratory study. Syst Rev. 2017;6(1):245. doi:10.1186/s13643-017-0644-y.
24. Deeks JJ, Higgins JP, Altman DG, et al. Analysing data and undertaking meta-analysis. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane Handbook for Systematic Reviews of Interventions (version 6.3). Cochrane; 2022. Available from www.training.cochrane.org/handbook.
25. Page MJ, Higgins JP, Sterne J, et al. Assessing risk of bias due to missing results in a synthesis. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane Handbook for Systematic Reviews of Interventions (version 6.3). Cochrane; 2022. Available from www.training.cochrane.org/handbook.
26. Bom PRD, Rachinger H. A generalized-weights solution to sample overlap in meta-analysis. Res Synth Methods. 2020;11(6):812–32. doi:10.1002/jrsm.1441.
27. Ageron F-X, Debaty G, Savary D, Champly F, Albasini F, Usseglio P, Vallot C, Galvagno S, Bouzat P, TRENAU Group Association of helicopter transportation and improved mortality for patients with major trauma in the northern French Alps trauma system: an observational study based on the TRENAU registry. Scand J Trauma Resusc Emerg Med. 2020;28(1):35. doi:10.1186/s13049-020-00730-z.
28. Andruszkow H, Schweigkofler U, Lefering R, Frey M, Horst K, Pfeifer R, Beckers SK, Pape H-C, Hildebrand F. Impact of Helicopter Emergency Medical Service in Traumatized Patients: which Patient Benefits Most? PLoS One. 2016;11(1):e0146897. In eng doi:10.1371/journal.pone.0146897.
29. Bläsius FM, Horst K, Brokmann JC, Lefering R, Andruszkow H, Hildebrand F, TraumaRegister DGU® TraumaRegister DGU® Helicopter Emergency Medical Service and Hospital Treatment Levels Affect Survival in Pediatric Trauma Patients. JCM. 2021;10(4):837. doi:10.3390/jcm10040837.
30. Davis DP, Peay J, Serrano JA, Buono C, Vilke GM, Sise MJ, Kennedy F, Eastman AB, Velky T, Hoyt DB, et al. The impact of aeromedical response to patients with moderate to severe traumatic brain injury. Ann Emerg Med. 2005;46(2):115–22. In eng doi:10.1016/j.annemergmed.2005.01.024.
31. Englum BR, Rialon KL, Kim J, Shapiro ML, Scarborough JE, Rice HE, Adibe OO, Tracy ET. Current use and outcomes of helicopter transport in pediatric trauma: a review of 18,291 transports. J Pediatr Surg. 2017;52(1):140–4. In eng doi:10.1016/j.jpedsurg.2016.10.030.
32. Jitsuiki K, Nagasawa H, Muramatsu K-I, Takeuchi I, Ohsaka H, Ishikawa K, Yanagawa Y. The Usefulness of Physician-Staffed Helicopters for Managing Severe Abdominal Trauma Patients. J Emerg Trauma Shock. 2022;15(1):12–6. doi:10.4103/jets.jets_96_21.
33. Kushida Y, Jitsuiki K, Muramatsu K-I, Ikegami S, Nagasawa H, Takeuchi I, Ohsaka H, Oode Y, Omori K, Yanagawa Y, et al. A comparison of physician-staffed helicopters and ground ambulances transport for the outcome of severe thoracic trauma patients. Am J Emerg Med. 2021;45:358–60. doi:10.1016/j.ajem.2020.08.088.
34. Weinlich M, Martus P, Blau MB, Wyen H, Walcher F, Piatek S, Schüttrumpf JP. Competitive advantage gained from the use of helicopter emergency medical services (HEMS) for trauma patients: evaluation of 1724 patients. Injury. 2019;50(5):1028–35. doi:10.1016/j.injury.2018.12.018.
35. Hosomi S, Kitamura T, Sobue T, Nakagawa Y, Ogura H, Shimazu T. Association of Pre-Hospital Helicopter Transport with Reduced Mortality in Traumatic Brain Injury in Japan: A Nationwide Retrospective Cohort Study. J Neurotrauma. 2022;39(1-2):76–85. doi:10.1089/neu.2021.0181.
36. Sterne J, Hernan M, McSAleenan A, Reeves B, Higgins J, et al. Assessing risk of bias in a non-randomized study. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane Handbook for Systematic Reviews of Intervensions (version 6.3). Cochrane; 2022. Available from www.training.cochrane.org/handbook.
37. Schunemann HJ, Vist GE, Higgins JP, et al., et al. Interpreting results and drawing conclusions. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane Handbook for Systematic Reviews of Intervensions (version 6.3). Cochrane; 2022. Available from www.training.cochrane.org/handbook.

38. Newgard CD, Fischer PE, Gestring M, Michaels HN, Jurkovich GJ, Lerner EB, Fallat ME, Delbridge TR, Brown JB, Bulger EM, et al. National guideline for the field triage of injured patients: recommendations of the National Expert Panel on Field Triage, 2021. J Trauma Acute Care Surg. 2022;93(2):e49–e60. doi:10.1097/TA.0000000000003627.
39. Buchanan IM, Coates A, Sne N. Does Mode of Transport Confer a Mortality Benefit in Trauma Patients? Characteristics and Outcomes at an Ontario Lead Trauma Hospital. Cjem. 2016; 18(5):363–9. In eng doi:10.1017/cem.2016.15.
40. Mitchell AD, Tallon JM, Sealy B. Air versus ground transport of major trauma patients to a tertiary trauma centre: a province-wide comparison using TRISS analysis. Can J Surg. 2007;50(2):129–33. In eng.
41. Taylor C, Jan S, Curtis K, Tzannes A, Li Q, Palmer C, Dickson C, Myburgh J. The cost-effectiveness of physician staffed Helicopter Emergency Medical Service (HEMS) transport to a major trauma centre in NSW, Australia. Injury 2012. 2012; 43(11):1843–9. In eng). doi:10.1016/j.injury.2012.07.184.
42. Moore L, Champion H, Tardif P-A, Kuimi B-L, O'Reilly G, Leppaniemi A, Cameron P, Palmer CS, Abu-Zidan FM, Gabbe B, et al. Impact of Trauma System Structure on Injury Outcomes: A Systematic Review and Meta-Analysis. World J Surg. 2018;42(5):1327–39. doi:10.1007/s00268-017-4292-0.
43. Baxt WG, Moody P. The impact of a rotorcraft aeromedical emergency care service on trauma mortality. Jama. 1983;249(22): 3047–51. In eng). doi:10.1001/jama.1983.03330460029027.
44. Lefebvre C, Glanville J, Briscoe S, et al., et al. Searching for and selecting studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane Handbook for Systematic Reviews of Interventions (version 6.3). Cochrane; 2022. Available from www.training.cochrane.org/handbook.
45. Andruszkow H, Lefering R, Frink M, Mommsen P, Zeckey C, Rahe K, Krettek C, Hildebrand F. Survival benefit of helicopter emergency medical services compared to ground emergency medical services in traumatized patients. Crit Care. 2013;17(3): R124. In eng). doi:10.1186/cc12796.
46. Davis DP, Peay J, Good B, Sise MJ, Kennedy F, Eastman AB, Velky T, Hoyt DB. Air medical response to traumatic brain injury: a computer learning algorithm analysis. J Trauma. 2008; 64(4):889–97. In eng. doi:10.1097/TA.0b013e318148569a.
47. Davis DP, Stern J, Sise MJ, Hoyt DB. A follow-up analysis of factors associated with head-injury mortality after paramedic rapid sequence intubation. J Trauma. 2005;59(2):486–90. In eng. doi:10.1097/00005373-200508000-00037.
48. Shi L, Lin L. The trim-and-fill method for publication bias: practical guidelines and recommendations based on a large database of meta-analyses. Medicine (Baltimore). 2019;98(23):e15987. doi: 10.1097/MD.0000000000015987.