

---

**Mémoire, y compris stage professionnalisant[BR]- Séminaires  
méthodologiques intégratifs[BR]- Mémoire : Etude de validation de la fiabilité  
intra et inter-évaluateurs de l'EMS Benefit score au sein du centre médical  
hélicoptère de Bra-sur-Lienne**

**Auteur :** Roussel, Clément

**Promoteur(s) :** STIPULANTE, Samuel; 23682

**Faculté :** Faculté de Médecine

**Diplôme :** Master en sciences de la santé publique, à finalité spécialisée patient critique

**Année académique :** 2023-2024

**URI/URL :** <http://hdl.handle.net/2268.2/19872>

---

*Avertissement à l'attention des usagers :*

*Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.*

*Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.*

---

**Etude de validation de la fiabilité intra et inter-  
évaluateurs de l'EMS Benefit score au sein du  
centre médical hélicoptéré de Bra-sur-Lienne**

Mémoire présenté par **Clément ROUSSEL**  
en vue de l'obtention du grade de  
Master en Sciences de la Santé publique  
Finalité spécialisée en patient critique (SIU)  
Année académique 2023 - 2024

**Etude de validation de la fiabilité intra et inter-  
évaluateurs de l'EMS Benefit score au sein du  
centre médical hélicoptéré de Bra-sur-Lienne**

Mémoire présenté par **Clément ROUSSEL**  
en vue de l'obtention du grade de  
Master en Sciences de la Santé publique  
Finalité spécialisée en patient critique (SIU)  
Année académique 2023 - 2024

Promoteur : **Mr Samuel STIPULANTE**

Co-Promotrice : **Mme Méryl PAQUAY**

## Remerciements

Je tiens à exprimer ma profonde gratitude envers toutes les personnes qui m'ont soutenu et encouragé lors de la réalisation de ce mémoire.

Je tiens tout d'abord à remercier mes promoteurs, monsieur Stipulante et madame Paquay, pour leurs conseils dans les étapes clés de ce projet. Leur expertise ont grandement contribué à la réussite de ce travail de recherche.

Un grand merci également à l'ensemble du personnel du CMH de Bra-sur-Lienne pour leur participation active et l'intérêt dont ils ont fait preuve. Leur contribution a été essentielle pour mener à bien ce projet.

Je tiens à exprimer ma reconnaissance envers mes collègues, notamment Maxime, pour leur patience, leur compréhension et leurs encouragements qui m'ont été d'une grande aide pour surmonter les moments de doute et de difficulté.

Je tiens également à remercier ma famille et mes amis pour leur soutien inconditionnel tout au long de cette aventure.

Enfin, un grand merci à toutes les personnes ayant contribué de près ou de loin à la réalisation de ce mémoire. Leur soutien et leur collaboration ont été précieux.

Merci à tous pour votre soutien et votre encouragement, sans lesquels ce travail n'aurait pas été possible.

# Table des matières

|   |           |
|---|-----------|
| Préambule.....  | 1         |
| <b>1. Introduction.....</b>   | <b>2</b>  |
| <b>L'efficacité du transport médical hélicoptéré: un sujet en évolution .....</b>           | <b>2</b>  |
| <b>L'impact de l'HEMS sur les pathologies critiques .....</b>                               | <b>4</b>  |
| <b>La problématique du triage préhospitalier.....</b>                                       | <b>6</b>  |
| <b>Scores de triage.....</b>  | <b>7</b>  |
| <b>L'EMS Benefit score : EBS .....</b>  | <b>8</b>  |
| <b>2. Matériel &amp; méthode.....</b>   | <b>11</b> |
| <b>Type d'étude .....</b>   | <b>11</b> |
| <b>Population étudiée et taille d'échantillon .....</b>                                     | <b>11</b> |
| <b>Paramètres étudiés et outils de collecte des données.....</b>                            | <b>13</b> |
| a) Paramètres étudiés.....  | 13        |
| b) Organisation de la collecte des données .....  | 14        |
| c) Traitement des données et méthode d'analyse .....  | 16        |
| d) Contrôles et critères de qualité .....   | 17        |
| e) Les aspects réglementaires.....  | 17        |
| <b>3. Résultats .....</b>   | <b>18</b> |
| <b>Présentation descriptives des données .....</b>  | <b>18</b> |
| <b>Analyses de fiabilité.....</b>   | <b>19</b> |
| a) Coefficients de corrélation intra-classe pour l'accord intra-évaluateur.....             | 20        |
| b) Coefficients de corrélation intra-classe pour l'accord inter-évaluateurs infirmier ..... | 21        |
| c) Coefficients de corrélation intra-classe pour l'accord inter-évaluateurs médical.....    | 22        |
| d) Coefficients de corrélation intra-classe pour l'accord inter-évaluateurs global .....    | 23        |
| e) Synthèse des résultats .....   | 23        |
| <b>4. Discussion et Limites .....</b>   | <b>24</b> |
| <b>Discussion &amp; perspectives .....</b>  | <b>24</b> |
| <b>Forces &amp; Limites.....</b>  | <b>27</b> |
| <b>5. Conclusions.....</b>  | <b>29</b> |

## **Lexique**

HEMS : Helicopter Emergency Medical Service

SMUH: Service Médicaux d'Urgence par Hélicoptère

EMS : Emergency Medical Service

SMUR : Service Mobile d'Urgence et de Réanimation

SMU : Soins Médicaux d'Urgence

STEMI : ST-segment Elevation Myocardial Infarction

ALS : Advanced Life Support

ICP : Primary Coronary Intervention

TRISS : Trauma and Injury Severity Score

ISS : Injury Severity Score

ASAPS : American Society of Anesthesiologists

NACA : National Advisory Committee for Aeronautics

HBS : HEMS Benefit Score

EBS : EMS Benefit Score

ASA-PS: American Society of Anesthesiologists Physical Status Classification System

ECOG: Eastern Cooperative Oncology Group Performance Status

ICPC-2: International Classification of Primary Care, 2nd edition

AVC: Accident Vasculaire Cérébral

IDM: Infarctus Du Myocarde

ICC : Coefficient de Corrélacion Intra-classe ou Indice de Coefficient Intra-classe

## Résumé

**Introduction** - La présente étude vise à évaluer la fiabilité de l'EMS Benefit Score (EBS) dans le contexte du transport médical aérien. L'objectif est d'explorer la cohérence des évaluations réalisées par les professionnels de la santé, tant au niveau d'un même évaluateur qu'entre différents évaluateurs, afin de déterminer si cet outil peut être fiable pour évaluer les performances des services médicaux d'urgence.

**Matériel et Méthodes** - L'étude a été menée dans un centre médical hélicoptère en Belgique et a impliqué des médecins et des infirmiers. Les scores EBS ont été encodés à deux moments différents pour évaluer la fiabilité intra-évaluateur et inter-évaluateur. L'analyse statistique a été réalisée en utilisant l'indice de corrélation intra-classe (ICC) pour mesurer la concordance des évaluations.

**Résultats** - Les résultats ont montré une excellente fiabilité de l'EBS, avec des ICC moyens allant de 0.806 à 0.942 pour l'agrément intra-évaluateur et de 0.954 à 0.988 pour l'agrément inter-évaluateur. La fiabilité des évaluations est restée constante même à des intervalles de temps éloignés et les différents professionnels de la santé ont montré une forte concordance dans leurs notations.

**Conclusions** - En conclusion, l'EMS Benefit Score s'est avérée être un outil fiable pour évaluer les performances des services médicaux aérien. Les résultats obtenus suggèrent que cet outil peut contribuer à améliorer la qualité des soins préhospitaliers et à garantir des interventions efficaces pour les patients en situation d'urgence. Il est important de poursuivre les recherches pour explorer davantage les applications pratiques de l'EBS et pour confirmer sa validité.

**Mots clé** : EBS ; HEMS ; Triage préhospitalier ; Fiabilité intra/inter-évaluateurs

## **Abstract**

**Introduction** - This study aims to evaluate the reliability of the EMS Benefit Score (EBS) in the context of medical air transport. The objective is to explore the consistency of assessments made by healthcare professionals, both within the same evaluator and between different evaluators, to determine if this tool can be reliable for evaluating the performance of emergency medical services.

**Materials and Methods** - The study was conducted at a helicopter medical center in Belgium and involved doctors and nurses. EBS scores were encoded at two different time points to evaluate intra-rater and inter-rater reliability. Statistical analysis was performed using the intraclass correlation coefficient (ICC) to measure the agreement of assessments.

**Results** - The results showed excellent reliability of the EBS, with average ICCs ranging from 0.837 to 0.974 for intra-rater agreement and from 0.931 to 0.974 for inter-rater agreement. The reliability of assessments remained consistent even at distant time intervals, and different healthcare professionals showed strong concordance in their ratings.

**Conclusions** - In conclusion, the EMS Benefit Score proved to be a reliable tool for evaluating the performance of air medical services. The results suggest that this tool can contribute to improving the quality of prehospital care and ensuring effective interventions for patients in emergency situations. Further research is important to explore the practical applications of the EBS and to confirm its validity.

**Keywords:** EBS ; HEMS ; Prehospital triage ; Intra/inter-rater reliability



## Préambule

Dans un monde en proie à la raréfaction des ressources, il apparaît de plus en plus nécessaire d'adopter des pratiques efficaces afin d'optimiser l'utilisation de ces biens et répondre aux besoins croissants de la population. Cela concerne tout particulièrement le domaine des soins de santé. Il est donc essentiel de pérenniser et rendre efficace ce système de soins pour garantir un accès équitable à des soins de qualité tout en maîtrisant les coûts.

Pour tendre vers ces impératifs au niveau national, il faut tenir compte de certaines caractéristiques et disparités à l'échelle locale. En effet, le pays est caractérisé par une forte densité de population dans les zones urbaines mais présente également des zones rurales moins peuplées et plus difficiles d'accès. Ces dernières sont confrontées à des défis particuliers en matière de soins de santé et plus particulièrement concernant l'aide médicale urgente.

En effet, l'augmentation de la durée des trajets médicaux en zone rurale est souvent associée à un déclin clinique chez les patients(1), mettant en lumière l'importance cruciale de l'accessibilité aux soins dans ces régions. Il a donc fallu réfléchir à une solution pour assurer une couverture rapide de l'ensemble du territoire et ceci dans le but de limiter les délais d'intervention ainsi que l'intervalle médicale libre.

En ce sens, il est opportun de considérer la possibilité de l'utilisation du transport sanitaire hélicoptéré. Ce vecteur de soins peut s'avérer un moyen efficace pour répondre aux besoins médicaux urgents, en particulier pour certaines pathologies dans les régions éloignées ou difficilement accessibles(2,3).

En tant qu'infirmier SIAMU ayant travaillé aux urgences et étudiant en santé publique, je suis profondément motivé par le désir d'améliorer la qualité des soins d'urgence. À travers mes expériences professionnelles, j'ai pu constater les défis auxquels sont confrontés les équipes médicales, notamment en termes de rapidité et d'efficacité dans la prise en charge des patients en situation critique.

En menant une recherche scientifique sur ce sujet, je souhaite contribuer à la diffusion des connaissances et à la sensibilisation des décideurs politiques concernant les avantages et les opportunités offerts par le transport sanitaire hélicoptère en termes de santé publique.

Ce projet de recherche s'inscrit dans une perspective de valorisation de mes compétences et expériences professionnelles au service de l'amélioration du système de santé dans le domaine des soins d'urgence.

## **1. Introduction**

### **L'efficacité du transport médical hélicoptère: un sujet en évolution**

Le transport sanitaire hélicoptère et sa possible influence sur l'amélioration des résultats demeure encore sujet à débat. Comme pour d'autres discussions scientifiques, il faut évaluer l'utilité des services médicaux d'urgence par hélicoptère (HEMS) en se basant sur des preuves issues de la littérature et des données probantes(4). Dans cette optique, on constate une évolution positive puisque la décennie la plus récente démontre une augmentation marquée du taux de publication de recherches portant sur les résultats des HEMS(5). De plus, on peut noter une diversification des cas d'études puisque si, à l'origine, la plupart étaient centrées sur les victimes de traumatismes, on peut souligner une augmentation notable des études se penchant sur des pathologies cardiaques ou encore sur des cas d'accidents vasculaires cérébraux(5). Notons qu'il est important de démontrer l'apport et l'efficacité du transport médical hélicoptère pour soutenir son utilisation étant donné que les HEMS représentent des ressources limitées et coûteuses avec des risques de sécurité élevés. Une revue systématique anglophone a d'ailleurs tenté de déterminer le coût économique des HEMS. Le coût et l'efficacité des HEMS variaient considérablement entre les études de cette même revue. De manière générale, la tendance montre que les transports médicaux hélicoptères sont généralement plus coûteux que le transport terrestre mais un certain nombre d'études ont démontré que les HEMS seraient rentables(6). L'estimation de cette rentabilité est d'ailleurs en lien avec la pathologie rencontrée lors de l'intervention médicale et semble meilleure en cas de blessures graves.(7). Il apparaît nécessaire de tenir compte de facteurs locaux et inhérent au système de santé de chaque pays pour une meilleure compréhension. De plus, les

coûts liés aux risques de sécurité sont plus important comparé à d'autres modes de transport d'urgence et à d'autres domaines de l'aviation. Selon cette étude de 2021, les taux d'accidents mortels ont été historiquement plus élevés avec l'utilisation de l'HEMS(8).

Au-delà des couts, l'efficacité générale des HEMS par rapport au service médical d'urgence au sol, reste un enjeu qui est souvent débattu, les résultats des études existantes restant mitigés. Les études comparatives sont difficiles à résumer en raison de leurs résultats hétérogènes et à la faible qualité globale des preuves(9). Afin d'asseoir la légitimité de l'HEMS dans la prise en charge préhospitalière, de nombreuses variables ont été étudiées ces dernières années : l'impact de l'HEMS sur la survie des patients, la réduction des temps de transport vers les hôpitaux, l'amélioration des soins médicaux en situation d'urgence, l'efficacité des interventions médicales réalisées à bord de l'hélicoptère, etc. Ces études visent à évaluer les avantages et les inconvénients de l'utilisation de l'hélicoptère médicalisé afin d'optimiser son utilisation.

La survie est la variable la plus étudiée pour évaluer l'efficacité des secours médicaux hélicoptérés. C'est un indicateur clé, objectif et facile d'utilisation. De nombreuses données semblent aller dans ce sens puisqu'elles montrent de manière claire une réduction de la mortalité grâce aux services médicaux d'urgence en hélicoptère par rapport aux services médicaux d'urgence terrestres(7,9–14). La quantité importante d'études scientifiques à ce sujet permet donc de penser que l'hélicoptère pourrait améliorer significativement la survie à court terme dans de nombreux cas, notamment dans certaines pathologies aiguës et chez les patients gravement blessés. C'est ce que soutient cette analyse de régression multivariée qui dresse le profil du patient ayant subi un traumatisme important comme celui qui bénéficierait le plus des avantages que procurent les services médicaux d'urgence par hélicoptère en termes de survie(15). Si une association positive est mis en évidence entre l'implication de l'HEMS et les chances de survie du patient dans un état critique, on ne parvient toutefois pas exactement à identifier comment le bénéfice est obtenu(16). La littérature laisse supposer que l'effet bénéfique proviendrait de l'expertise supplémentaire et des options thérapeutiques apportées sur les lieux.

Il est également important de rester critique et de nuancer ces propos. La qualité des études sur ce sujet varie en fonction des méthodologies utilisées, des populations étudiées et des

critères de jugement évalués. Pour illustrer cela, une étude récente Finlandaise met en avant la persistance d'une surmortalité élevée à long terme chez les patients pris en charge par les services médicaux d'urgence par hélicoptère (17). Une seconde étude, Danoise cette fois, nuance également ces propos en ne démontrant pas de différence significative sur la mortalité à long terme (sur un an) entre l'utilisation de l'HEMS par rapport à l'EMS au sol (18). Dans tous les cas, l'efficacité du transport sanitaire héliporté semble en lien avec la nécessité d'agir rapidement pendant la période préhospitalière, pour les patients gravement malades ou blessés, car les décisions et les interventions effectuées à ce moment-là ont un impact considérable sur les résultats(19).

### **L'impact de l'HEMS sur les pathologies critiques**

Comme évoqué, l'HEMS semble apporter une réponse essentielle principalement aux trois principales pathologies critiques : le patient traumatisé grave, l'infarctus du myocarde (IDM) et l'accident vasculaire cérébral (AVC). En effet, la rapidité de transport vers un centre spécialisé, la possibilité d'interventions médicales avancées en vol et la coordination avec les équipes de différentes spécialités peuvent permettre d'améliorer considérablement les chances de survie et de récupération des patients souffrant de ces pathologies(20).

Dans la prise en charge du patient qui a subi un traumatisme important, l'enjeu des premiers secours est d'optimiser la prise en charge dans la première heure afin de le stabiliser. Le transport aérien dans ce cas précis, semble avoir tout son sens puisqu'il permet une réduction significative du temps d'accès au centre de traumatologie pour les patients grièvement blessés(9). Cela engendre également une diminution de la mortalité chez ces patients ayant subi un traumatisme important (21) comme mis en évidence précédemment. Il est même établi dans certaines situations que l'HEMS est recommandé comme devant être le premier choix pour les patients blessés gravement(22). Cependant, il faut être attentif à ne pas surutiliser le transport par hélicoptère chez tous les patients victime d'un traumatisme plus ou moins important. C'est ce qui est rapporté dans une étude récente qui semble démontrer dans son analyse que la grande majorité des patients traumatisés transportés par hélicoptère n'avaient pas besoin de ce mode de transport(23).

Concernant l'IDM, sa prise en charge par des secours hélicoptérés est très largement étudiée puisqu'on sait avec certitude que le délai avant une intervention coronarienne percutanée (ICP) est un facteur important pour la survie du patient. Tout d'abord, il est mis en exergue que l'utilisation de l'HEMS permet de diminuer le temps de transport dans la prise en charge de patients atteints d'IDM par rapport à un SMUR routier. Cela permet de garantir une prise en charge équivalente en termes de délais dans les zones rurales, comparables à ceux constatés en milieu urbain. (24). Cela peut apporter une réponse pour les interventions en zones rurales et éloignées d'un centre de soins spécialisés et entraîner une amélioration du pronostic chez le patient(2).

Enfin, à propos de la prise en charge de victimes d'AVC, on peut mettre en avant une publication qui relève un impact positif sur le résultat neurologique grâce à l'HEMS en comparaison avec un transport d'urgence standard. Cela sans toutefois arriver à établir une association statistiquement significative entre l'HEMS et la réduction de la mortalité(25). Une revue systématique datant de 2021 a essayé d'explorer l'impact général de l'HEMS, comme moyen de transport pour les AVC ischémiques aigus, sur les processus de soins et les résultats. Elle n'a cependant pas réussi à démontrer un bénéfice significatif en termes de résultats par rapport aux EMS ni d'impact clair sur les processus de soins(26). On constate également ici, un pourcentage important d'interventions pour des patients souffrant d'AIS et ne répondant pas à des critères d'utilisation optimale du transport médical d'urgence hélicoptéré(27). Ceci mettant en lumière un manque de précision du système de triage pour identifier les transports HEMS à haut rendement.

Nous l'avons observé, l'utilisation de l'hélicoptère dans le cadre de l'aide médicale urgente peut contribuer au « juste soin »(28). La littérature scientifique met en évidence l'utilité des services de secours hélicoptérés dans des situations d'urgence médicale où l'accès terrestre est limité ou impossible et dans les pathologies où le temps est une donnée centrale. Cela étant dit, il est crucial d'utiliser ces services de secours hélicoptérés à bon escient, car leur coût est élevé. Les ressources limitées doivent être allouées de manière équitable, en prenant en compte les bénéfices potentiels des interventions hélicoptérées par rapport aux alternatives terrestres disponibles.

## La problématique du triage préhospitalier

Tout d'abord, il faut évoquer l'importance d'utiliser des critères préhospitaliers rigoureux pour justifier la décision d'envoyer un hélicoptère médical. D'après cette recherche, l'utilisation de l'hélicoptère de transport pour effectuer des transferts inter-établissements pourrait parfois être davantage dictée par la disponibilité des ressources employées (à savoir la disponibilité d'un transport aérien) plutôt que par des considérations cliniques (29). Une utilisation inappropriée des HEMS entraîne une augmentation des coûts pour le patient et le système de soins (30). Des initiatives et projets sont actuellement en phase de développement pour diminuer les déploiements inutiles et réduire les coûts. En ce sens, des chercheurs ont mis en place des facteurs pouvant prédire efficacement les cas de patients bénéficiant d'un transport par hélicoptère inapproprié (31). Ou encore l'utilisation d'un outil de décision simple pour identifier les patients pour lesquels le transport terrestre permettrait des économies importantes sans accroître le risque de mortalité (32).

A côté de cela, la littérature souligne l'importance de l'évaluation régulière des services de secours hélicoptérés pour garantir leur efficacité et leur impact positif sur les résultats des patients et cela pour éviter une surutilisation (33). Des études ont montré que la sélection rigoureuse des cas éligibles pour l'utilisation des services aériens ainsi que des protocoles de triage appropriés, sont nécessaires pour optimiser les ressources et éviter une utilisation inutile(34). Ces principes à la base des soins préhospitaliers doivent être appliqués tout en tenant compte des spécificités du transport aérien d'urgence. Les soins préhospitaliers sont des soins médicaux d'urgence prodigués aux patients avant leur arrivée à l'hôpital après activation des services médicaux d'urgence (SMU). Il est donc nécessaire d'évaluer les avantages potentiels des missions des SMU afin de mieux répartir les ressources disponibles et améliorer le système de triage.

Le transport terrestre-EMS doit être préconisé lorsqu'il permet d'assurer les soins nécessaires et un acheminement rapide vers les soins finaux. Il convient d'examiner de manière rigoureuse les pratiques visant à dépêcher systématiquement des équipes HEMS, afin de limiter le risque de recours excessif à ces services. Il faut donc définir des critères spécifiques et des limites d'utilisation des services médicaux aériens adaptés à la réalité du système de santé local et régional, tout en identifiant les patients concernés. Un triage inapproprié peut en effet avoir plusieurs conséquences délétères. Le triage excessif augmente les coûts et peut augmenter le

risque concernant la sécurité des équipes de transport. Un sous-triage peut entraîner une augmentation de la morbidité et de la mortalité chez les patients qui auraient pu bénéficier de son utilisation. Cependant, en raison de la complexité des missions préhospitalières, il est difficile d'évaluer les bénéfices des soins dispensés avant l'arrivée à l'hôpital. Il est donc complexe d'évaluer de manière précise et objective leur efficacité sur le patient. Ces freins peuvent avoir un impact sur la pertinence de l'utilisation de l'HEMS, on peut d'ailleurs noter qu'il semble y avoir une faible proportion de patients qui bénéficie véritablement de l'avantage apporté par le vecteur hélicoptère par rapport aux nombres de sorties effectuées(35). Cette difficulté de mettre au point un critère de triage individuel ou une combinaison de plusieurs de ces critères pour prédire efficacement et de manière uniforme une utilisation excessive du transport aérien est mise en évidence par cette publication de 2019(36).

### **Scores de triage**

Pour améliorer cela, des scores peuvent être utilisés comme outils guidant le triage préhospitaliers afin de rationaliser l'utilisation de l'HEMS. Ces outils permettent d'évaluer de manière objective la gravité des patients et de faciliter la décision d'utiliser les secours aériens hélicoptérés. En intégrant ces scores dans les protocoles d'intervention et en favorisant leur utilisation systématique, nous pouvons maximiser les chances de survie des patients et garantir une utilisation efficiente des ressources. Des scores existent déjà et sont parfois utilisés mais ils sont rarement spécifiques à la phase préhospitalière et au transport hélicoptéré. Citons à ce titre le Trauma and Injury Severity Score (TRISS) qui permet d'évaluer le degré de gravité d'une blessure et estimer la probabilité de survie d'un patient. Il s'agit d'un système d'évaluation couramment utilisé dans le domaine de la traumatologie. C'est un score qui présente certains avantages, comme le rappelle une étude comparative récente dans sa capacité à prédire la mortalité chez les patients polytraumatisés. Elle met en évidence que le score TRISS est le plus performant pour prédire cette variable parmi six systèmes de notation (et notamment en comparaison avec l'Injury Severity Score)(37). Il présente cependant de nombreuses limites(38) et son utilisation à la phase préhospitalière n'apparaît pas envisageable.

D'autres scores et classifications utilisés dans les milieux préhospitaliers existent, tel que le score du comité consultatif national de l'aéronautique (NACA) qui prend notamment en compte la gravité des blessures, les signes vitaux du patient, et la nécessité de soins médicaux urgents. Il est d'ailleurs estimé efficace pour prédire la mortalité dans les 48 heures suivant un événement médical grave(39). Cependant il fait l'objet de nombreuses critiques en raison de sa subjectivité. Une étude a estimé sa fiabilité inter-évaluateurs dans la notation et le codage des patients préhospitaliers en l'analysant grâce à un coefficient de corrélation intra-classe (ICC)(40). Cette fiabilité inter-évaluateurs est importante car elle permet de déterminer la quantité d'erreur ou de variabilité dans les mesures effectuées. Il en ressort un ICC de 0.65, ce qui correspondrait à une fiabilité modérée à bonne mais inférieur à la référence qui est attendue par la suite dans notre étude, puisque l'ICC sera fixé à 0.7. De plus, on peut émettre certaines limites sur ces résultats puisque les scénarios pour les évaluations concernaient principalement des patients classiques et ne posant que peu de problèmes. Cela rend donc l'extrapolation de ces conclusions peu évidentes. De surcroît, il est important de noter que ces scores ne prennent pas en compte l'impact des soins préhospitaliers et n'ont pas été particulièrement conçus pour une utilisation dans ce contexte. Cela rend donc difficile son utilisation à but comparative et dans l'évaluation des bénéfices. Cela indique qu'un score pour être pertinent dans la perspective poursuivie par notre étude doit être non seulement spécifiquement conçu ou adapté pour une utilisation préhospitalière et avoir pour but d'évaluer les avantages des missions des services d'urgence par hélicoptère.

### **L'EMS Benefit score : EBS**

Abordons à présent le thème central de notre recherche. L'HEMS Benefit Score (HBS) est un système de notation à neuf niveaux, développé dans les années 1990, pour mesurer l'efficacité et l'impact positif des interventions d'urgence effectuées par les hélicoptères médicalisés sur la prise en charge des patients en situation critique. Des directives d'application sous forme d'exemple d'interventions sont mises à disposition de l'évaluateur pour le guider dans son choix et diminuer la subjectivité du score. Ce score a été utilisé dans les SMUH finlandais depuis son introduction en 1997 mais les études qui évaluent sa fiabilité sont peu nombreuses et aucune ne s'est penché sur sa validité.



Si une première étude estime sa fiabilité inter-évaluateurs significative et acceptable par rapport aux valeurs de références (40), il faut néanmoins nuancer ces résultats. En effet, une limite principale, tout comme pour le NACA, résidait dans le fait que les missions fictives pour l'évaluation concernaient essentiellement des cas de patients courants et peu problématiques, ce qui ne permet pas d'appliquer ces résultats à la réalité pratique. Une étude plus récente a d'ailleurs démontré une fiabilité inter-évaluateurs bien moindre que la précédente dans la notation et le codage des patients préhospitaliers en utilisant de manière intentionnelle des cas de patients plus complexes (41). De plus, étant donné que la fiabilité du système repose en partie sur les exemples d'interventions, il est possible d'envisager une amélioration en fournissant des exemples plus élaborés et exhaustifs. Cela est d'autant plus vrai que les définitions initiales remontent à près de 20 ans et ne sont plus pertinentes aujourd'hui en raison des nombreux changements et évolutions dans les soins préhospitaliers.

Tous ces aspects ont conduit à une mise à jour de l'HBS. Cette révision a donné l'EMS Benefit Score (EBS). La catégorisation numérique en neuf niveaux héritée de l'HBS original est restée inchangée, mais les exemples d'interventions de chaque niveau ont été complètement revus. Cette refonte a également permis d'étendre le système de notation au-delà de l'utilisation de l'HEMS pour inclure tous les soins d'urgence préhospitaliers, même ceux des unités non HEMS, et de mieux répondre aux besoins actuels. Même si les exemples ont été révisés au travers de l'opinion consensuelle d'experts via une méthodologie Delphi(42) en donnant des lignes directrices pour la sélection de la catégorie la plus adéquate, la notation repose toujours sur le choix subjectif de la personne qui effectue l'encodage. Malgré cette nature subjective de l'EBS, il peut permettre de mettre en lumière des aspects de la prise en charge préhospitalière qui n'ont pas été pris en compte dans les évaluations traditionnelles. En outre, il peut potentiellement contribuer à l'amélioration continue des pratiques préhospitalières en identifiant des domaines d'amélioration et en incitant les équipes à réfléchir de manière critique à leurs actions. Cet indicateur a donc pour vocation de mesurer la qualité de l'évaluation préhospitalière de l'état des patients, d'être utilisé pour améliorer les performances des EMS et pour obtenir plus de données sur où et dans quel type de missions les patients sont susceptibles d'en bénéficier le plus. Cela pourrait donc aider à déterminer si l'utilisation d'un hélicoptère pour les services médicaux d'urgence est justifiée, en prenant en considération les bénéfices pour les patients. Il deviendrait un outil d'évaluation utilisé par les

professionnels de la santé pour aider à prendre des décisions éclairées concernant l'utilisation des SMUH.

Cependant, du fait de sa création récente, aucune recherche ne s'est encore penchée sur l'étude de sa fiabilité ni de sa validité. Étudier la fiabilité et la validité d'une mesure, un score en l'occurrence, est important car cela garantit que cette mesure est précise, cohérente et pertinente. La fiabilité d'un score fait référence à la constance des résultats obtenus à partir des mesures réalisées(43). Une mesure fiable produira des résultats similaires lorsqu'elle est répétée à plusieurs reprises et cela indépendamment des circonstances ou des conditions. La fiabilité est importante car elle permet d'avoir confiance dans les résultats obtenus, en s'assurant que ces résultats ne sont pas dus au hasard ou à des erreurs d'évaluation. La validité, quant à elle, se rapporte à son aptitude à évaluer de manière précise et exacte le phénomène ou la caractéristique qu'elle est censée mesurer(43). Un score valide est une mesure précise et exacte du phénomène évalué. L'étude de la validité permet de s'assurer que la mesure est appropriée pour le contexte dans lequel elle est utilisée et qu'elle fournit des informations pertinentes et utiles. Il est généralement recommandé d'étudier la fiabilité avant la validité dans une recherche car si le score étudié n'est pas fiable, il est peu probable qu'il soit valide. La présente étude se fixe donc pour objectif principal de tester la fiabilité intra et inter-évaluateurs de l'EMS Benefit score. L'objectif secondaire étant de déterminer si la fiabilité est impactée par l'intervalle de temps entre le premier et le second encodage du score, et par la profession de l'évaluateur.

De cet objectif découle la question de recherche suivante :

« Quelle est la répétabilité intra-évaluateur et la reproductibilité inter-évaluateurs de l'EMS Benefit Score? »

## 2. Matériel & méthode

### Type d'étude

Cette recherche adopte une approche quantitative avec un design de type étude de validation pour évaluer et quantifier la fiabilité des mesures de manière objective. L'étude utilise à la fois des collectes rétrospectives et prospectives puisque qu'elle utilise des données préalablement collectées et utilise ces mêmes données pour des évaluations programmées sur une période de temps définie à l'avance. Ce travail se déroule au sein du centre médical hélicoptère de Bra-sur-Lienne et se classe donc comme une étude monocentrique.

Le design de cette étude aura pour ambition de répondre aux hypothèses suivantes :

- La fiabilité intra et inter-évaluateurs sera bonne à excellente, ce qui correspond à un niveau supérieur à l'ICC souhaité, c'est-à-dire supérieur à 0.7.
- Ni l'intervalle de temps entre le premier et le second encodage du score, ni la profession de l'évaluateur, qu'il soit médecin ou infirmier, n'impacteront la bonne fiabilité de ce score.

Les hypothèses de départ ont été définies de manière intuitive et expérimentale par les chercheurs et sont étayées par la littérature scientifique existante.

### Population étudiée et taille d'échantillon

En Belgique, il existe actuellement deux centres de secours hélicoptère : le Centre Médical Hélicoptère (CMH) de Bra-sur-Lienne et le Centre Médical Hélicoptère (CMH) de Brugge. Ce mémoire pour des raisons pratiques ne concerne que la population du centre médical hélicoptère de Bra-sur-Lienne. Ce dernier se compose de 34 membres, pilotes, médecins, infirmiers et membres du personnel administratif. La méthode d'échantillonnage est non probabiliste au volontaire, également appelée échantillonnage par convenance, pour garantir un meilleur engagement des participants.

| Etape de recherche            | Critères d'inclusion   | Critère d'exclusion  |
|-------------------------------|--|--|
| Analyse intra-<br>évaluateur  | <ul style="list-style-type: none"> <li>• Être médecin du CMH</li> <li>• Avoir au moins réalisé 46 interventions depuis la mise en place du score EBS au sein du CMH en juillet 2023. Cela est indispensable pour pouvoir réaliser la totalité des deux séries d'encodage réalisées à deux moments distincts pour l'agrément intra-évaluateur.</li> </ul> | <ul style="list-style-type: none"> <li>• Etre infirmier au CMH. Le choix de ne pas retenir les infirmiers dans cette analyse est dicté par l'utilisation de ce score dans la pratique. En effet, il n'est pour l'instant rempli au quotidien que par les médecins et le design de cette étude ne permet donc pas d'évaluer l'encodage réalisé par des infirmiers à deux moments différents.</li> </ul> |
| Analyse inter-<br>évaluateurs | <ul style="list-style-type: none"> <li>• Être médecin ou infirmier du CMH</li> </ul>   |  |

A partir de cette population, une taille d'échantillon correspondant à un nombre d'encodage nécessaire doit être calculé en amont. Le score étant traité comme une variable quantitative, la taille de l'échantillon se base sur l'ICC souhaité. En se référant à des publications scientifiques récentes et similaires dans leur sujet d'étude et dans l'objectif poursuivi(40,41), la décision fut prise par l'équipe de recherche (réfèrent statistique, promoteurs et étudiant) de fixer l'ICC à 0.7 avec une précision de  $\pm 0,15$  pour l'agrément intra et inter-évaluateur. Tout

cela en fixant un niveau de confiance de 95% ( $\alpha=0,05$ ). Les normes de l'indice de corrélation intra-classe sont les suivantes :

- $ICC < 0.4$  : Faible concordance
- $0.4 \leq ICC < 0.6$  : Concordance modérée
- $0.6 \leq ICC < 0.7$  : Bonne concordance
- $0.7 \leq ICC < 0.8$  : Très bonne concordance
- $0.8 \leq ICC < 0.9$  : Excellente concordance
- $ICC = 1$  : Concordance parfaite

Un ICC de 0.7 est considéré comme substantielle, cela indique une forte concordance entre les mesures répétées et suggère une grande fiabilité des mesures.

- Pour l'accord inter-évaluateurs, en tenant compte de l'ICC, la précision, le niveau de confiance et le nombre d'évaluateurs, chaque médecin a besoin de noter 20 patients et 21 patients pour les infirmiers. Pour pouvoir réaliser une comparaison, les mêmes patients à encoder ont été sélectionnés pour les deux groupes. La sélection de ces patients pour l'ensemble des évaluateurs se réalise sur l'ensemble de la base de données à disposition et de manière aléatoire.
- Pour l'accord intra-évaluateur, en tenant compte de l'ICC, la précision, le niveau de confiance et le nombre d'évaluateurs, chaque médecin a besoin de noter 46 patients. Il s'agit ici d'un ré-encodage de 46 patients déjà vus et notés une première fois par chacun des médecins lors d'une intervention SMUR et présent dans la base de données. La sélection de ces 46 patients se fera également de manière aléatoire mais dans les interventions propre à chaque médecin.
- 

## **Paramètres étudiés et outils de collecte des données**

### **a) Paramètres étudiés**

L'EMS Benefit score, que l'on pourrait traduire par « score d'efficacité » ou « score de performance » des services médicaux d'urgence est le sujet d'étude. C'est un outil récent qui

peut être utilisé dans le domaine des services médicaux d'urgence, notamment par hélicoptère, pour évaluer la pertinence et l'efficacité des interventions médicales. Ce score aide à déterminer si une intervention est justifiée et si elle peut apporter un bénéfice significatif pour le patient.

Le paramètre principal étudié sera la fiabilité de l'EBS. Celle-ci se décompose en deux parties :

- La fiabilité inter-évaluateurs est la mesure de la concordance entre les évaluations faites par différents évaluateurs sur une même série de scores. Elle permet de déterminer dans quelle mesure les résultats obtenus restent cohérents lorsque différentes personnes effectuent une évaluation similaire. Une fiabilité inter-évaluateurs élevée indique que les évaluateurs sont en accord sur les résultats, ce qui renforce la validité des conclusions de l'étude. Il s'agit donc ici de déterminer si la classification des patients avec l'EBS par des évaluateurs différents aboutit à des résultats similaires lorsqu'ils évaluent les mêmes patients.
- La fiabilité intra-évaluateur, quant à elle, est la mesure de la cohérence des évaluations effectuées par le même évaluateur à des moments différents. Il s'agit de la situation de « Test/Retest » par un même évaluateur d'une même série de scores. Elle permet d'évaluer la stabilité des mesures dans le temps et de vérifier que les résultats obtenus ne varient pas en fonction du moment de l'évaluation. Une fiabilité intra-évaluateur élevée indique que l'évaluateur est stable dans ses évaluations, ce qui renforce la précision des résultats de l'étude. Il s'agit donc ici de déterminer si la classification des patients avec l'EBS en T1 est différente ou non de la classification en T2.

#### b) Organisation de la collecte des données

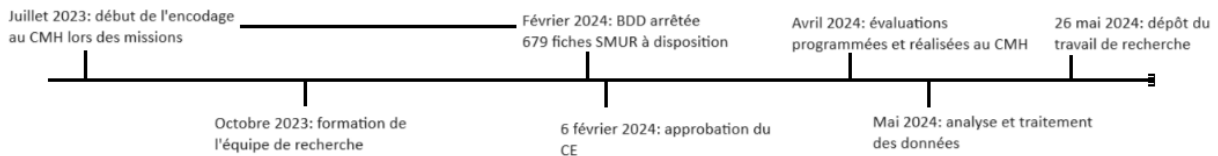
La formation des évaluateurs sur la manière d'utiliser le score de manière précise et cohérente ne fut pas nécessaire concernant les médecins. Ils bénéficient effectivement d'une formation continue puisqu'ils utilisent ce score dans leur travail quotidien depuis juillet 2023 et ont été sensibilisés et formés à son utilisation lors de sa mise en place. Concernant les infirmiers, bien qu'ils ne s'en servent pas eux-mêmes, l'objectif secondaire de cette recherche est justement de déterminer si les infirmiers en l'état évaluent les patients de manière aussi fiable que les médecins. Les former en amont aurait pu donc venir impacter les résultats.

Le premier encodage de l'EBS a été réalisé au CMH par les médecins après chaque mission depuis juillet 2023. Les scores étaient reportés sur la fiche de sortie SMUR. Cela fournit la base de données nécessaire pour la sélection des cas de patients à encoder ou réencoder pour notre étude. Ces encodages sont programmés durant le mois d'avril 2024 pour permettre au personnel de les réaliser et d'obtenir une collecte de données complète.

L'ensemble des fiches SMUR de la base de données récoltés depuis juillet 2023 sont anonymisés avant le second encodage programmé, toutes les informations concernant l'identité du patient ont été supprimées. Il ne reste sur la fiche que les informations relatives à l'anamnèse réalisée par le médecin lors de la sortie originelle, à savoir l'examen physique, le monitoring, la médication ainsi que les remarques potentielles. A partir de là, une randomisation a été effectuée à partir du logiciel Excel. Pour procéder à cette randomisation, chaque fiche SMUR anonymisée s'est vue attribuer un numéro aléatoire à l'aide de la fonction « ALEA ». Cette fonction octroie un nombre réel aléatoire distribué de manière symétrique supérieur ou égal à 0 et inférieur à 1. Puis dans un second temps, et grâce à une fonction « RANG » qui permet de classer les nombres de cette liste dans un ordre choisi, les numéros aléatoires ont été reclassés dans un ordre croissant. Il ne restait plus qu'à sélectionner les 46 premiers numéros dans la liste propre à chaque médecin et les 21 premiers numéros dans l'ensemble de la base de données pour les médecins et infirmiers. Pour finir, l'ensemble des fiches SMUR qui ont été tirées aléatoirement pour l'agrément intra et inter-évaluateurs ont été analysées et résumées dans un tableau croisé dynamique et ceci pour éviter tout risque de doublons pour les médecins. Après avoir retiré certains d'entre eux et les avoir remplacés de façon aléatoire par des fiches non tirées précédemment, des listes définitives furent arrêtées.

Pour aider les médecins et infirmiers dans cette tâche, un modus operandi a été réalisé et fourni dans chacune des fardes distribuées. Il se composait d'une introduction rappelant les enjeux et objectifs de l'étude ainsi que le mode opératoire à proprement parlé. Se trouvait également présent le système de notation EBS en annexe pour permettre un encodage plus facile et précis. Les fardes destinées aux médecins étaient individuelles et nominales, condition nécessaire pour évaluer la fiabilité intra-évaluateur. La décision fut prise d'opter pour une version papier pour augmenter l'implication des participants.

## Vue d'ensemble temporelle de l'organisation de la collecte des données



### c) Traitement des données et méthode d'analyse

L'ensemble des résultats obtenus via la collecte des données sont examinés grâce aux logiciels d'analyse IBM SPSS et R studio pour l'analyse statistique et la visualisation des données.

Concernant l'analyse statistique de fiabilité par l'indice de corrélation intraclasse :

- L'analyse choisie est l'agrément puisque ce dernier va véritablement étudier la cohérence de l'accord entre les variables. Une analyse de la consistance ne permet d'étudier que la corrélation d'une variable par rapport à l'autre mais de manière linéaire. Ces variables peuvent être linéairement liées mais sans nécessairement être en accord ni cohérentes entre elles, voilà pourquoi ce choix a été opéré.
- Le modèle sélectionné est un modèle mixte à deux facteurs qui détermine d'où provient l'effet aléatoire. Ce modèle permet de fixer les évaluateurs en ce qui concerne la variance et de limiter cet effet aléatoire aux mesures effectuées.
- Dans toutes les analyses effectuées, l'intervalle de confiance est fixé à 95%. Il y a donc 95% de toutes les mesures, de tous les échantillons qui se trouveront compris entre la borne inférieure de l'ICC et la borne supérieure.
- Concernant les résultats, la mesure qui nous intéresse principalement est la moyenne de l'ICC. L'ICC moyen permet en effet de percevoir dans quelle mesure les évaluations successives sont cohérentes les unes par rapport aux autres et cela est donc plus intéressant dans les situations de « Test/Retest » et d'inter-évaluation.
- La p-valeur retenue pour déterminer si les tests d'hypothèses sont statistiquement significatifs est de 0.05. Cela signifie qu'il y a moins de 5% de chances que les résultats observés soient dus au hasard. Ce choix est réalisé par convention dans le domaine de la recherche en sciences et en statistique. Cela limite le risque de faux positif tout en minimisant les risques de conclure à tort qu'il y a un effet.



#### d) Contrôles et critères de qualité

Pendant la collecte et le traitement des données, diverses mesures ont été prises pour garantir leur qualité. Dans l'ensemble, la grille STROBE a été utilisée pour vérifier les critères de qualité. Tout cela dans le but de faciliter la reproductibilité de l'étude et de renforcer la confiance dans les conclusions tirées.

De plus, un contrôle continu a été mise en place pour assurer une approche méthodologique rigoureuse grâce à une équipe de recherche compétente.

- Un référent statistique pour concevoir le plan d'analyse statistique, en déterminant les tests statistiques appropriés à utiliser en fonction des objectifs de recherche et des hypothèses formulées.
- Un ingénieur informatique pour s'assurer d'une randomisation adéquate et optimale dans le but de diminuer les biais de sélection.
- Des promoteurs pour l'approbation des différentes étapes de ce travail en s'assurant qu'il réponde aux normes académiques et aux attentes fixées.
- Mis en place par le chercheur d'un mécanisme de double vérification lors de l'encodage des données, dans le but de minimiser les erreurs humaines liées à la saisie manuelle.
- 

#### e) Les aspects réglementaires

Soumission de l'étude au comité d'éthique du collège des Enseignants du Master en Sciences de la Santé Publique de l'université de Liège. Le protocole d'étude a reçu l'accord du Comité d'Éthique du Centre Hospitalier Universitaire de Liège le 06 février 2024 avec la référence 2024/41. Aucune objection éthique à la réalisation de cette étude n'a été relevée. Ce projet a fait l'objet d'un *call for abstract* au congrès de l'Association Francophone des Infirmiers d'Urgence en date du 26 mars 2024 à Braine-l'Alleud.

### 3. Résultats

Les résultats des analyses concernant la fiabilité intra-évaluateur et inter-évaluateurs de l'EMS benefit score seront détaillés dans la section suivante. Ces analyses permettront d'évaluer la cohérence et l'accord des évaluations effectuées par les médecins et les infirmiers, à la fois au niveau individuel et en comparaison les uns avec les autres.

La base de données, contenant les notations EBS, commencée en juillet 2023 et arrêtée en février 2024, contient 679 fiches SMUR. Le nombre de volontaires inclus et participant à cette étude fut de 10 médecins et 8 infirmiers.

#### Présentation descriptives des données

Pour l'agrément intra-évaluateur, 10 séries de deux fois 46 encodages réalisées à deux moments distincts ont été analysées. Nous nous trouvons ici dans la situation d'un « Test/Retest » de la même série par un même évaluateur, en l'occurrence un médecin. Il s'agit de comparer l'accord et la cohérence de la première série par rapport à l'autre.

Un seul médecin a été inclus en n'ayant effectué que 43 sorties, ces données manquantes ont été jugées acceptables par l'équipe de recherche. Deux médecins furent exclus en raison du nombre insuffisant de missions et d'encodages réalisés (26 et 28 respectivement). Dans sa globalité, nous évaluons donc 444 scores EBS sur 460 prévues, outre les 3 données exclues à priori, certaines données sont manquantes. Ce qui renvoie à un total de 96.5% de données valides. L'analyse en tient compte dans l'ajustement de l'indice de corrélation intra-classe en excluant ces données manquantes.

Chaque série de donnée est étudiée séparément dans un premier temps puis dans un second, l'ensemble des variables sont regroupées dans un tableau récapitulatif.

Tableau 1 . Récapitulatif de traitement des informations

| Intra-évaluateur |     |       |
|------------------|-----|-------|
| Observations     | N   | %     |
| Valide           | 444 | 96.52 |
| Exclu            | 16  | 3.48  |
| Total            | 460 | 100   |

Pour l'agrément inter-évaluateurs, huit séries de 21 scores identiques ont été analysées, ainsi que dix séries de ces mêmes 20 scores correspondant respectivement aux données encodées par les infirmiers et par les médecins. Il s'agit de mesurer l'accord et la concordance des séries entre elles. Dans sa globalité, nous évaluons donc 377 scores EBS sur 378 prévues, 1 seule donnée est manquante. Ce qui renvoie à un total de 99.73% de données valides. Les séries encodées par les infirmiers sont d'abord analysées entre elles puis celles des médecins. La distinction est effectuée dans le but d'opérer une comparaison et de répondre à l'objectif secondaire. Les deux lots de données sont ensuite regroupés et comparés dans un tableau final.

Tableau 2 . Récapitulatif de traitement des informations

| Inter-évaluateurs |     |       |
|-------------------|-----|-------|
| Observations      | N   | %     |
| Valide            | 377 | 99.73 |
| Exclu             | 1   | 0.27  |
| Total             | 378 | 100   |

## Analyses de fiabilité

Passons maintenant à l'analyse de la fiabilité pour mieux comprendre la qualité des mesures et l'accord entre les évaluateurs.

a) Coefficients de corrélation intra-classe pour l'accord intra-évaluateur

Tableau 3. Mesures moyennes des coefficients de corrélation intra-classe pour l'accord intra-évaluateur

| Médecin       | Coefficients de corrélation intra-classe,<br>Mesures moyennes |                  | Intervalle de confiance à 95% |         |
|---------------|---|------------------|-------------------------------|---------|
|               | Borne inférieure  | Borne supérieure | Moyennes                      | P-value |
| 1             | 0.915   | 0.974            | 0.953                         | <0.001  |
| 2             | 0.788   | 0.935            | 0.883                         | <0.001  |
| 3             | 0.728   | 0.917            | 0.850                         | <0.001  |
| 4             | 0.933   | 0.980            | 0.964                         | <0.001  |
| 5             | 0.766   | 0.931            | 0.873                         | <0.001  |
| 6             | 0.923   | 0.977            | 0.958                         | <0.001  |
| 7             | 0.711   | 0.919            | 0.847                         | <0.001  |
| 8             | 0.806   | 0.947            | 0.900                         | <0.001  |
| 9             | 0.704   | 0.910            | 0.837                         | <0.001  |
| 10            | 0.789   | 0.935            | 0.883                         | <0.001  |
| <b>Totaux</b> | 0.8063  | 0.9425           | 0.8948                        | <0.001  |

Au niveau de l'agrément intra-évaluateur, les valeurs de l'ICC, pour chaque évaluateur pris individuellement, sont comprises entre 0.837 et 0.964. Cela signifie que la fiabilité des évaluations varie de très bonne à excellente. Dans ce cas, les évaluateurs ont montré individuellement une cohérence et un accord élevé dans leurs évaluations du score EBS à deux moments différents. Si l'on interprète la valeur de l'ICC moyen de l'ensemble de ces évaluations qui est de 0.894, cela indique qu'en moyenne, les deux évaluations successives démontrent une concordance élevée l'une par rapport à l'autre. Le fait que la borne inférieure de l'intervalle de confiance soit de 0.806 atteste que même le scénario le moins favorable reste élevé et acceptable. Il est important de noter que ces chiffres ne représentent pas la

variabilité entre les évaluateurs, mais plutôt la cohérence et l'accord de chaque évaluateur avec lui-même lors des évaluations successives. De plus, la p-valeur étant inférieure à 0.001, cela confirme que nos résultats sont statistiquement significatifs et renforce la fiabilité de cette mesure. En résumé, l'interprétation des résultats suggère que l'évaluation intra-évaluateur de l'EBS est fiable et que les mesures prises par un même évaluateur sont fortement concordantes.

#### b) Coefficients de corrélation intra-classe pour l'accord inter-évaluateurs infirmier

Tableau 4. Mesures moyennes des coefficients de corrélation intra-classe pour l'accord inter-évaluateurs infirmier

| <u>Coefficients de corrélation intra-classe,</u> | Intervalle de confiance à 95% |                  |                          |                  |
|--|-------------------------------|------------------|--------------------------|------------------|
|  | Borne inférieure              | Borne supérieure | Corrélation intra-classe | P-value          |
| Mesures uniques                                  | 0.468                         | 0.790            | 0.628                    | <b>&lt;0.001</b> |
| Mesures moyennes                                 | 0.876                         | 0.968            | 0.931                    | <b>&lt;0.001</b> |

Au niveau de l'agrément inter-évaluateurs concernant les infirmiers, on remarque un ICC moyen de 0.931 quand on compare la cohérence des différentes variables entre elles, c'est-à-dire des différents infirmiers dans leurs notations. Cet ICC indique une très bonne concordance entre les évaluateurs dans leurs notations. La limite inférieure se situe à un taux de 0.876, ce qui est très acceptable et cela même dans le pire des cas mais cette valeur s'éloigne légèrement de l'ICC moyen, on peut également le constater pour la borne supérieure. Cette différence entre la limite inférieure et supérieure pourrait indiquer qu'il y a une certaine variabilité parmi les évaluateurs, avec quelques écarts dans les notations.

### c) Coefficients de corrélation intra-classe pour l'accord inter-évaluateurs médical

Tableau 5. Mesures moyennes des coefficients de corrélation intra-classe pour l'accord inter-évaluateurs médical

| <u>Coefficients de corrélation intra-classe,</u> | Intervalle de confiance à 95% |                  |                          |                  |
|--|-------------------------------|------------------|--------------------------|------------------|
|  | Borne inférieure              | Borne supérieure | Corrélation intra-classe | P-value          |
| Mesures uniques                                  | 0.596                         | 0.859            | 0.733                    | <b>&lt;0.001</b> |
| Mesures moyennes                                 | 0.937                         | 0.984            | 0.965                    | <b>&lt;0.001</b> |

Au niveau de l'agrément inter-évaluateurs concernant les médecins, on remarque un coefficient de corrélation de 0.965 en moyenne qui indique une excellente concordance entre les évaluations des différents médecins dans leurs notations. Cela signifie que les évaluateurs sont en accord dans leurs jugements et que les évaluations sont fiables. La limite inférieure se situe à un taux de 0.937, ce qui est excellent et indique que la corrélation est très proche de l'ICC moyen. La limite supérieure de 0.984 est également proche de cette valeur. Cela indique que la variabilité des évaluations est relativement faible et renforce l'idée que les évaluations des médecins sont en accord et donc fiables.

Quand on compare ces deux tableaux entre eux, on constate que l'ICC moyen de 0.965 au niveau médical atteste d'une concordance encore plus élevée entre les évaluateurs qu'au niveau infirmier (0.931). Ces deux indices de corrélation intra-classe restent néanmoins considérés comme très élevés et montrent une grande fiabilité des mesures peu importe la profession de l'évaluateur. Cependant, on peut toutefois noter une variabilité plus élevée dans les évaluations faites par les infirmiers en comparaison de celles effectuées par les médecins. Cela peut refléter une certaine discordance entre les évaluateurs infirmiers dans certains cas. Les médecins possèdent un accord élevé même dans les cas les plus difficiles à évaluer et une concordance presque parfaite dans la plupart des cas.

#### d) Coefficients de corrélation intra-classe pour l'accord inter-évaluateurs global

Tableau . Mesures moyennes des coefficients de corrélation intra-classe pour l'accord inter-évaluateurs global

|                  | Intervalle de confiance à 95% |                  |                          |                  |
|------------------|-------------------------------|------------------|--------------------------|------------------|
|                  | Borne inférieure              | Borne supérieure | Corrélation intra-classe | P-value          |
| Mesures uniques  | 0.533                         | 0.819            | 0.674                    | <b>&lt;0.001</b> |
| Mesures moyennes | 0.954                         | 0.988            | 0.974                    | <b>&lt;0.001</b> |

Pour l'analyse de la fiabilité inter-évaluateurs de l'ensemble des évaluateurs, médecins et infirmiers, on peut constater un ICC moyen de 0.974, ce qui indique une excellente concordance dans leurs évaluations. Les scores EBS attribués par chacun d'entre eux sont très cohérents. L'intervalle de confiance de 95% compris entre 0.954 et 0.988 confirme cette conclusion en indiquant que la vraie valeur de l'ICC se situe avec une certitude de 95% dans ces bornes.

#### e) Synthèse des résultats

Tableau 2 . Récapitulatif des résultats

| Coefficient de corrélation intra-classe |            |                  |                  |         |
|---|------------|------------------|------------------|---------|
| Fiabilité du score EBS                  |            |                  |                  |         |
|   | ICC moyens | Borne inférieure | Borne supérieure | P-value |
| Intra-évaluateur                        | 0.894      | 0.806            | 0.942            | <0.001  |
| Inter-évaluateurs                       | 0.974      | 0.954            | 0.988            | <0.001  |

Les résultats finaux semble indiquer une excellente fiabilité à la fois intra-évaluateur (ICC moyen de 0.894), dans les évaluations répétées par le même évaluateur, et inter-évaluateurs (ICC moyen de 0.974), entre différents évaluateurs, de l'EMS benefit score. La fiabilité inter-

évaluateurs se rapprochant même d'une valeur de 1 et donc d'une corrélation parfaite. Au niveau de l'ensemble des tests statistiques, la p-valeur est jugée comme hautement significative, ce qui atteste d'une grande robustesse et fiabilité dans les résultats obtenus lors de ces tests. Cela renforce la confiance dans les conclusions de l'étude et suggère que les variables testées ont un effet réel sur le phénomène étudié. Les résultats obtenus nous ont permis de constater une fiabilité intra-évaluateur excellente, avec un coefficient de corrélation intra-classe (ICC) de 0.9. Cela signifie que les évaluations réalisées par un même évaluateur sont cohérentes et fiables. De plus, la fiabilité inter-évaluateurs s'est avérée quasi-parfaite, avec un ICC de 0.97. Ce résultat souligne la robustesse de l'EMS benefit score en tant qu'outil de comparaison des performances entre différents intervenants médicaux.

## **4. Discussion et Limites**

### **Discussion & perspectives**

L'étude a répondu de manière satisfaisante aux objectifs principaux et secondaires en s'appuyant sur les hypothèses formulées. L'objectif principal visait à tester la fiabilité de l'EMS Benefit Score. Avec des coefficients de corrélation intra-classe élevés et significatifs, l'étude a réussi à démontrer la fiabilité de l'EBS comme outil servant à apprécier la qualité de l'évaluation préhospitalière. Les résultats ont montré des ICC moyens élevés, allant de 0.894 pour l'intra-évaluateur à 0.974 pour l'inter-évaluateurs, ce qui confirme la cohérence et l'accord élevé des notations. L'objectif secondaire consistait à déterminer si la fiabilité de ce score était affectée par le délai entre les encodages et par la profession de l'évaluateur. Les résultats ont montré une évaluation constante à des intervalles de temps éloignés de plusieurs mois (juillet 2023-avril 2024). Concernant la profession de l'évaluateur, on constate un indice de corrélation intra-classe moyen modérément plus élevé du côté des médecins en comparaison des infirmiers (0.965 contre 0.931) pour l'accord inter-évaluateurs ainsi qu'une variabilité légèrement plus élevée dans les évaluations infirmières par rapport à celles des médecins. Puisqu'il n'existe actuellement aucun gold standard pour l'évaluation d'un score préhospitalier, la validation de notre méthode d'utilisation de ce score est basée sur l'estimation de sa fiabilité intra et inter-évaluateurs.



En comparaison avec les études antérieures(40,41) sur l'HBS et d'autres scores et codes utilisés dans des contextes préhospitaliers (NACA, ASA-PS, ICPC-2, ECOG), l'EBS semble offrir une meilleure fiabilité inter-évaluateurs, avec des ICC moyens plus élevés. Cela suggère que l'EBS est un outil plus robuste et plus cohérent pour évaluer l'efficacité des interventions préhospitalières. On peut supposer que cette supériorité est liée au fait que les exemples d'interventions de chaque niveau aient été complètement revus lors de la création de l'EBS. Cela a pu permettre une meilleure compréhension et une application plus précise du score par les évaluateurs. Des exemples plus élaborés et exhaustifs peuvent avoir contribué à réduire la subjectivité et à améliorer la cohérence des évaluations. Ces données ouvrent la voie à de nombreuses perspectives de recherche et d'application pratique de l'EMS Benefit Score (EBS). Tout d'abord, il serait intéressant d'étendre ce champ d'étude à d'autres vecteurs EMS pour évaluer cette validation à une plus grande échelle. Cela permettrait de confirmer la fiabilité de cet outil dans différents contextes et d'obtenir des données supplémentaires sur son utilisation. Le score est en effet conçu et destiné pour être appliqué à tous les vecteurs EMS en fournissant une évaluation standardisée des avantages fournis par les diverses interventions préhospitalières. En utilisant ce score, les praticiens peuvent comparer et évaluer les performances des différentes interventions et déterminer laquelle est la plus efficace en fonction des besoins des patients. De plus, notre étude se concentrait principalement sur la fiabilité de l'EBS, mais d'autres aspects de sa validité pourraient également être étudiés pour obtenir une image complète de cet outil d'évaluation. Il serait donc intéressant d'approfondir ces aspects dans des études futures pour affiner l'utilisation de l'EBS dans la pratique clinique. On peut souligner à cet égard que l'étude européenne Delphi de 2021(42), qui est à l'origine de la réactualisation de l'EBS, a choisi de réviser les exemples de catégories de prestation sans toutefois toucher aux catégories de notation numérique, car il a été considéré comme imprudent de les modifier en même temps. Cette catégorisation numérique ne repose sur aucun fondement scientifique puisqu'elle fut créée à l'origine sur la base de l'expérience pratique. Il pourrait donc être intéressant de le prendre en considération dans le futur et envisager une révision pour perfectionner le système de notation. Cela pourrait faire l'objet d'une étude de validité de contenu. Par ailleurs, il pourrait être intéressant d'envisager une étude comparative entre l'EMS Benefit Score et d'autres outils d'évaluation des services médicaux d'urgence. Cela permettrait de déterminer si l'EBS est plus efficace ou pertinent que d'autres méthodes existantes et d'identifier ses avantages

et inconvénients par rapport à ces outils. Cela pourrait rentrer dans une étude de validité de critère. Certaines études comparatives concernant l'HBS ont été réalisées par le passé et peuvent servir d'inspiration pour l'EBS dans le futur(40,41). Compte tenu de la réactualisation récente de l'EBS, il n'existe pas à l'heure actuelle de publications scientifiques recensées qui se sont fixées pour objectif l'étude de cette validité(42).

Les résultats obtenus démontrent certaines tendances, notamment une stabilité des évaluations sur des intervalles de temps éloignés et une variabilité légèrement plus élevée des évaluations du côté infirmier. Pour aller plus loin, il pourrait être intéressant de mener des études plus approfondies afin de comparer la fiabilité intra-évaluateur de l'EBS, en utilisant divers intervalles de temps pour identifier celui qui permet de réduire au minimum le biais de mémoire(44). Les variations légères entre les médecins et les infirmiers peuvent s'expliquer, en plus de la profession, par des données socio-démographiques de nos évaluateurs, telles que leur âge ou leur nombre d'années d'expérience. En effet, ces variables peuvent avoir un impact significatif sur la qualité et la précision des évaluations réalisées. Il serait intéressant de collecter ces informations pour une analyse plus approfondie et une interprétation plus précise des résultats de notre recherche puisque l'on connaît le rôle crucial de l'expérience de l'évaluateur dans la fiabilité des évaluations(45). Ici, on pourrait expliquer ces différences modérées dans la manière d'évaluer les patients par le bénéfice engendré par l'utilisation de l'EBS au quotidien s'agissant des médecins et leur degré de formation à l'utilisation de ce score. La comparaison entre les deux groupes permet de mettre en lumière et de souligner l'importance d'une formation continue et d'une standardisation des pratiques(46). Malgré cela, notre recherche ne met pas en évidence de différences significatives dans la fiabilité de l'évaluation, qu'elle soit réalisée par des médecins ou des infirmiers. Ces résultats encourageants peuvent ouvrir la possibilité à une utilisation plus large de ce score dans la pratique quotidienne. L'emploi de l'EBS par les infirmiers pourrait permettre d'augmenter la quantité de données recueillies. Cela favoriserait également le développement de leurs compétences en matière d'évaluation préhospitalière et permettrait un allègement de la charge de travail médical. Cette approche collaborative pourrait contribuer à long terme à améliorer la qualité des soins préhospitaliers et à une prise en charge plus efficace. Afin de soutenir cette vision, il serait opportun d'envisager l'intégration d'une utilisation quotidienne de l'EBS par les infirmiers du CMH. Puis dans un second temps, il serait pertinent d'évaluer les

infirmiers sur la répétabilité de leur mesure et d'obtenir une vision globale sur la fiabilité de leurs notations(47).

Concernant les perspectives offertes par cette recherche dans le domaine des soins préhospitaliers. Il faut rappeler que la fiabilité des outils d'évaluation est cruciale dans ce domaine pour garantir des interventions appropriées et efficaces. En diminuant la nature subjective de l'EBS, son utilisation pourrait s'avérer utile pour obtenir des informations sur certaines composantes des missions préhospitalières sans toutefois être exhaustive. L'évaluation de l'efficacité des soins d'urgence préhospitaliers est en effet complexe et aucun système de notation ne couvre entièrement cette question. L'utilisation de l'EBS, combinée à d'autres manœuvres, peut soutenir l'amélioration du système en permettant un meilleur contrôle de la qualité, ce qui favoriserait également le développement de l'efficacité de l'HEMS. Ces pratiques pourraient alors avoir un impact significatif sur les performances en permettant d'obtenir plus de données exploitables pour une amélioration continue. En outre, il serait pertinent d'explorer les perspectives d'application de l'EMS Benefit Score et de réfléchir à son utilisation dans la pratique clinique. Comment cet outil pourrait être intégré de manière optimale dans les protocoles et les pratiques des équipes médicales? Comment pourrait-il contribuer à améliorer la qualité des soins préhospitaliers et à garantir des interventions plus efficaces pour les patients en situation d'urgence? Il n'existe pas actuellement à notre connaissance d'études spécifiques sur l'impact de l'EMS Benefit Score sur la mortalité. Cependant, il est possible que des études futures examinent l'impact de ce score sur les résultats à long terme des patients, y compris la mortalité.

### **Forces & Limites**

Il est essentiel que les études de validation possèdent une cohérence interne élevée dans ses résultats(44). Cette cohérence interne est ici assurée par un alpha de Cronbach élevé (> 0.8) accompagnant l'ensemble de nos résultats. Cela permet d'assurer la fiabilité et la validité des données recueillies, tout en facilitant la comparaison de ces résultats pour les chercheurs. Il faut souligner que l'ensemble des notations EBS effectuées pendant cette étude ont eu lieu sur des fiches tirées d'interventions réelles et non de scénarios fictifs. Les données fictives peuvent effectivement parfois être biaisées ou ne pas refléter fidèlement les tendances et les comportements observés dans la pratique. Les scénarios tirés d'expériences réelles

permettent de renforcer la légitimité des résultats obtenus et leur impact potentiel. Des mesures ont également été prises pour diminuer le risque de biais. L'ensemble des fiches d'interventions de la base de données ont été anonymisées avant d'être présentées aux évaluateurs pour diminuer le risque de biais de recherche. De plus, la randomisation de ces fiches à évaluer ainsi que la suppression des cas répétitifs ont eu pour but de diminuer la possibilité de biais de sélection. Enfin, une taille d'échantillon ainsi qu'un nombre d'évaluateurs suffisants ont été assurés lors de la conduite de cette recherche pour garantir une étude de variabilité inter-évaluateurs de qualité selon les recommandations (48).

Malgré les résultats encourageants de notre étude, certaines limites doivent être prises en compte. Tout d'abord, l'échantillon utilisé était limité à un seul centre médical hélicoptéré, ce qui peut limiter la généralisation des résultats. Il serait donc nécessaire de reproduire cette étude dans d'autres contextes pour consolider et étendre les conclusions de cette recherche. La nature volontaire de la participation risquerait d'introduire un biais de sélection, car les participants pourraient être ceux qui se sentent plus confiants dans leurs évaluations. Il pourrait être pertinent dans des études futures de garantir une participation plus aléatoire pour éviter ce biais potentiel. Enfin, une limite inhérente au score lui-même a été mise en évidence lors de cette étude. Le fait de négliger certaines annotations explicatives des médecins, parfois présentes pour justifier leurs évaluations, peut également conduire à une perte d'informations importantes sur leur raisonnement et leur perception du cas. Cela est dû à la nécessité d'avoir un score numérique et objectif qui ne tient pas compte de ces données supplémentaires. De même que les ratures, changements d'avis et hésitations des évaluateurs peuvent créer des zones d'ombre dans les évaluations, rendant difficile la compréhension et l'interprétation de certains résultats.

## 5. Conclusions

Les objectifs de cette étude étaient de tester la fiabilité intra-évaluateur et inter-évaluateurs de l'EMS Benefit Score ainsi que de déterminer si cette fiabilité était impactée par l'intervalle de temps entre les évaluations et par la profession de l'évaluateur. Nous avons démontré la fiabilité de l'EMS Benefit Score (EBS) dans le contexte du transport médical aérien en soulignant une concordance élevée dans les évaluations médicales et infirmières. Les résultats ont confirmé le potentiel de cet outil pour évaluer les performances des services médicaux d'urgence et garantir des interventions appropriées et efficaces.

Il est important de poursuivre les recherches dans ce domaine pour explorer davantage les applications pratiques de l'EBS et pour améliorer sa validité. Il serait intéressant d'étendre cette étude à d'autres vecteurs EMS pour confirmer sa fiabilité dans différents contextes. De plus, des études sur les aspects de validité de l'EBS pourraient permettre une utilisation plus complète et précise de cet outil. L'EMS Benefit Score est un outil prometteur pour évaluer les performances des services médicaux d'urgence. L'amélioration continue de cet outil et son intégration dans la pratique clinique pourraient contribuer à une meilleure qualité de soins pour les patients en situations d'urgence.

Enfin, cette étude soulève également des questions sur l'utilisation optimale des scores d'évaluation dans le domaine des soins préhospitaliers. Comment pouvons-nous garantir que ces outils restent pertinents et fiables dans un environnement complexe et en évolution constante ? Comment pouvons-nous continuer à améliorer la qualité des soins d'urgence en utilisant des outils d'évaluation efficaces et pertinents ?

## Références :

1. Kai TR, Broady MJ, Davenport DL, Bernard AC. The effect of emergency medical system transport time on in route clinical decline in a rural system. *J Trauma Acute Care Surg.* juin 2020;88(6):734-41.
2. Ishiyama M, Kurita T, Takasaki A, Takamura T, Masuda J, Ishikura K, et al. Impact of Helicopter Transport on Reperfusion Times and Long-Term Outcomes in Acute Myocardial Infarction Patients in Rural Areas: A Report From the Mie Acute Coronary Syndrome Registry. *Air Medical Journal.* sept 2021;40(5):337-43.
3. Rzońca P, Świeżewski SP, Jalali R, Gotlib J, Gałązkowski R. Helicopter Emergency Medical Service (HEMS) Response in Rural Areas in Poland: Retrospective Study. *IJERPH.* 30 avr 2019;16(9):1532.
4. Brown BS, Pogue KA, Williams E, Hatfield J, Thomas M, Arthur A, et al. Helicopter EMS Transport Outcomes Literature: Annotated Review of Articles Published 2007–2011. *Emergency Medicine International.* 2012;2012:1-21.
5. Rudman JS, Fritz CL, Thomas SA, McCartin M, Price J, Blumen IJ, et al. Helicopter Emergency Medical Services Outcomes Research 1983 to 2022: Evidence Overview and Longitudinal Trends. *Air Medical Journal.* nov 2023;42(6):429-35.
6. Taylor CB, Stevenson M, Jan S, Middleton PM, Fitzharris M, Myburgh JA. A systematic review of the costs and benefits of helicopter emergency medical services. *Injury.* janv 2010;41(1):10-20.
7. Taylor C, Jan S, Curtis K, Tzannes A, Li Q, Palmer C, et al. The cost-effectiveness of physician staffed Helicopter Emergency Medical Service (HEMS) transport to a major trauma centre in NSW, Australia. *Injury.* nov 2012;43(11):1843-9.
8. Simonson RJ, Keebler JR, Chaparro A. A Bayesian Approach on Investigating Helicopter Emergency Medical Fatal Accidents. *aerosp med hum perform.* 1 juill 2021;92(7):563-9.
9. Hesselfeldt R, Steinmetz J, Jans H, Jacobsson M -L. B, Andersen DL, Buggeskov K, et al. Impact of a physician-staffed helicopter on a regional trauma system: a prospective, controlled, observational study. *Acta Anaesthesiol Scand.* mai 2013;57(5):660-8.
10. Ringburg AN, Thomas SH, Steyerberg EW, Van Lieshout EMM, Patka P, Schipper IB. Lives Saved by Helicopter Emergency Medical Services: An Overview of Literature. *Air Medical Journal.* nov 2009;28(6):298-302.
11. Phillips M, Arthur AO, Chandwaney R, Hatfield J, Brown B, Pogue K, et al. Helicopter Transport Effectiveness of Patients for Primary Percutaneous Coronary Intervention. *Air Medical Journal.* mai 2013;32(3):144-52.
12. Desmettre T, Yeguiayan JM, Coadou H, Jacquot C, Raux M, Vivien B, et al. Impact of emergency medical helicopter transport directly to a university hospital trauma center on mortality of severe blunt trauma patients until discharge. *Crit Care.* 2012;16(5):R170.
13. Sborov KD, Gallagher KC, Medvecz AJ, Brywczyński J, Gunter OL, Guillaumondegui OD, et al. Impact of a New Helicopter Base on Transport Time and Survival in a Rural Adult Trauma Population. *Journal of Surgical Research.* oct 2020;254:135-41.

14. 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. *Journal of Neurotrauma*. 1 janv 2022;39(1-2):76-85.
15. Andruszkow H, Schweigkofler U, Lefering R, Frey M, Horst K, Pfeifer R, et al. Impact of Helicopter Emergency Medical Service in Traumatized Patients: Which Patient Benefits Most? Burney RE, éditeur. *PLoS ONE*. 15 janv 2016;11(1):e0146897.
16. Frankema SPG, Ringburg AN, Steyerberg EW, Edwards MJR, Schipper IB, Van Vugt AB. Beneficial effect of helicopter emergency medical services on survival of severely injured patients. *British Journal of Surgery*. 20 oct 2004;91(11):1520-6.
17. Björkman J, Laukkanen-Nevala P, Olkinuora A, Pulkkinen I, Nurmi J. Short-term and long-term survival in critical patients treated by helicopter emergency medical services in Finland: a registry study of 36 715 patients. *BMJ Open*. févr 2021;11(2):e045642.
18. Alstrup K, Rognås L, Sollid S, Johnsen SP, Valentin JB, Petersen JAK. Association of Helicopter vs Ground Emergency Medical Transportation With 1-Year Mortality in Denmark. *JAMA Netw Open*. 11 janv 2021;4(1):e2033318.
19. Wilson MH, Habig K, Wright C, Hughes A, Davies G, Imray CHE. Pre-hospital emergency medicine. *The Lancet*. déc 2015;386(10012):2526-34.
20. Golden AP, Odoi A. Emergency medical services transport delays for suspected stroke and myocardial infarction patients. *BMC Emerg Med*. déc 2015;15(1):34.
21. Andruszkow H, Lefering R, Frink M, Mommsen P, Zeckey C, Rahe K, et al. Survival benefit of helicopter emergency medical services compared to ground emergency medical services in traumatized patients. *Crit Care*. 2013;17(3):R124.
22. Weinlich M, Martus P, Blau MB, Wyen H, Walcher F, Piatek S, et al. Competitive advantage gained from the use of helicopter emergency medical services (HEMS) for trauma patients: Evaluation of 1724 patients. *Injury*. mai 2019;50(5):1028-35.
23. Horwood CR, Ricci K, Sobol CG, Evans D, Eiferman D. Stop Flying the Patients! Evaluation of the Overutilization of Helicopter Transport of Trauma Patients. *Journal of Surgical Research*. déc 2020;256:290-4.
24. Moens D, Stipulante S, Donneau A, Hartstein G, Pirotte O, D'orio V, et al. Air versus ground transport of patients with acute myocardial infarction: experience in a rural-based helicopter medical service. *European Journal of Emergency Medicine*. août 2015;22(4):273-8.
25. Florez-Perdomo WA, Garcia-Ballestas E, Konar SK, Ramos-Gomez L, Al-Mufti F, Sursal T, et al. Effect of Helicopter Transportation of Acute Ischemic Stroke Patients on Mortality and Functional Outcomes: A Systematic Review and Meta-Analysis. *Air Medical Journal*. sept 2022;41(5):476-83.
26. Tal S, Mor S. The impact of helicopter emergency medical service on acute ischemic stroke patients: A systematic review. *The American Journal of Emergency Medicine*. avr 2021;42:178-87.
27. Adcock AK, Minardi J, Findley S, Daniels D, Large M, Power M. Value Utilization of Emergency Medical Services Air Transport in Acute Ischemic Stroke. *The Journal of Emergency Medicine*. nov 2020;59(5):687-92.

28. Chollet C. Activité et organisation des interventions médicales hélicoptérées en France. *Bulletin de l'Académie Nationale de Médecine*. déc 2023;207(9):1242-50.
29. Stewart K, Garwe T, Oluborode B, Sarwar Z, Albrecht RM. Association of Interfacility Helicopter versus Ground Ambulance Transport and in-Hospital Mortality among Trauma Patients. *Prehospital Emergency Care*. 3 sept 2021;25(5):620-8.
30. Miles MVP, Beasley JR, Reed HE, Miles DT, Haiflich A, Beckett AR, et al. Overutilization of Helicopter Emergency Medical Services in Central Gulf Coast Region Results in Unnecessary Expenditure. *Journal of Surgical Research*. mai 2022;273:211-7.
31. Brown C, Irfan W, Schoen JE, Marr AB, Stuke LE, Cavalea AC, et al. Predictors of Inappropriate Helicopter Transport. *The American Surgeon*. févr 2021;87(2):248-52.
32. Udekwu P, Schiro S, Toschlog E, Farrell M, McIntyre S, Winslow J. Trauma system resource preservation: A simple scene triage tool can reduce helicopter emergency medical services overutilization in a state trauma system. *J Trauma Acute Care Surg*. août 2019;87(2):315-21.
33. Bledsoe BE, Wesley AK, Eckstein M, Dunn TM, O'Keefe MF. Helicopter Scene Transport of Trauma Patients with Nonlife-Threatening Injuries: A Meta-Analysis: *The Journal of Trauma: Injury, Infection, and Critical Care*. juin 2006;60(6):1257-66.
34. Brown JB, Smith KJ, Gestring ML, Rosengart MR, Billiar TR, Peitzman AB, et al. Comparing the Air Medical Prehospital Triage Score With Current Practice for Triage of Injured Patients to Helicopter Emergency Medical Services: A Cost-effectiveness Analysis. *JAMA Surg*. 1 mars 2018;153(3):261.
35. Kurola J, Wangel M, Uusaro A, Ruokonen E. Paramedic helicopter emergency service in rural Finland – do benefits justify the cost? *Acta Anaesthesiol Scand*. août 2002;46(7):779-84.
36. Miyagi H, Evans DC, Werman HA. Are there Field Triage Criteria that Can Predict Low-Yield Air Medical Transports? *Prehosp Disaster med*. déc 2019;34(6):596-603.
37. Milton M, Engelbrecht A, Geysler M. Predicting mortality in trauma patients - A retrospective comparison of the performance of six scoring systems applied to polytrauma patients from the emergency centre of a South African central hospital. *African Journal of Emergency Medicine*. déc 2021;11(4):453-8.
38. Demetriades D, Chan LS, Velmahos G, Berne TV, Cornwell EE, Belzberg H, et al. TRISS methodology in trauma: the need for alternatives. *British Journal of Surgery*. 22 janv 2003;85(3):379-84.
39. Darioli V, Taffé P, Carron PN, Dami F, Vallotton L, Yersin B, et al. Evaluation of the discriminative performance of the prehospital National Advisory Committee for Aeronautics score regarding 48-h mortality. *European Journal of Emergency Medicine*. oct 2019;26(5):366-72.
40. Raatiniemi L, Liisanantti J, Tommila M, Moilanen S, Ohtonen P, Martikainen M, et al. Evaluating helicopter emergency medical missions: a reliability study of the HEMS benefit and NACA scores. *Acta Anaesthesiol Scand*. mai 2017;61(5):557-65.
41. Heino A, Laukkanen-Nevala P, Raatiniemi L, Tommila M, Nurmi J, Olkinuora A, et al. Reliability of prehospital patient classification in helicopter emergency medical service missions. *BMC Emerg Med*. déc 2020;20(1):42.



42. Heino A, Raatiniemi L, Iiro T, Meriläinen M, Liisanantti J, Tommila M, et al. The development of emergency medical services benefit score: a European Delphi study. *Scand J Trauma Resusc Emerg Med.* déc 2021;29(1):151.
43. Kimberlin CL, Winterstein AG. Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacy.* 1 déc 2008;65(23):2276-84.
44. Rey E, Carballo-Fazanes A, Varela-Casal C, Abelairas-Gómez C, on behalf of ALFA-MOV Project collaborators. Reliability of the test of gross motor development: A systematic review. *Montazeri A, éditeur. PLoS ONE.* 16 juill 2020;15(7):e0236070.
45. Porumb (Chifor) I, Leucuta DC, Nigoghossian M, Culic B, Lucaciu PO, Culic C, et al. Caries Lesion Assessment Using 3D Virtual Models by Examiners with Different Degrees of Clinical Experience. *Medicina.* 13 déc 2023;59(12):2157.
46. Shivali S, Thiagarajan P. A practical guide to the American Society of Anesthesiologists-physical status classification (ASA-PS). *Indian J Anaesth.* 2022;66(4):299.
47. Akinci D'Antonoli T, Cavallo AU, Vernuccio F, Stanzione A, Klontzas ME, Cannella R, et al. Reproducibility of radiomics quality score: an intra- and inter-rater reliability study. *Eur Radiol.* 21 sept 2023;34(4):2791-804.
48. Quinn L, Tryposkiadis K, Deeks J, De Vet HCW, Mallett S, Mookkink LB, et al. Interobserver variability studies in diagnostic imaging: a methodological systematic review. *BJR.* août 2023;96(1148):20220972.