**Wie effektiv unterstützt das Framework ehrapy die Identifikation von Risikofaktoren für akutes Nierenversagen bei Kindern nach Herzoperationen anhand eines angereicherten Routinedatensatzes?**

**Abstract**

Hintergrund: Akutes Nierenversagen (AKI) nach Kinderherzoperationen tritt häufig früh (0–48 h) auf und verschlechtert die Prognose. An der MHH-PICU steht ein AKI-gelabelter Routinedatensatz (>1.300 OPs, prä/intra/postoperative Daten) aus einem klinischen CDSS zur Verfügung.

Ziel: Es wurde geprüft, wie effektiv ehrapy (AnnData-basiert) die Identifikation von AKI-Risikofaktoren aus diesem angereicherten Routinedatensatz unterstützt und reproduzierbare Analysen ermöglicht.

Methoden: Aus dem Gesamtdatensatz wurde ein Analyse-Subset mit 1.209 OP-Episoden gebildet. Endpunkt war AKI 0–7 Tage (KDIGO). Der Datensatz wurde in AnnData/.h5ad harmonisiert. Ein logistisches GLM mit natürlichen Splines für Alter und cluster-robusten Standardfehlern (Clustering: Patient) wurde geschätzt. Kovariaten: OP-Dauer [h], Re-Operation (ja/nein), Geschlecht; getestet wurde Dauer×Re-OP. Zusätzlich wurden Onset-Analysen (Zeit bis AKI) und eine 5-fach Cross-Validation (ROC-AUC, PR-AUC, Brier) durchgeführt; der Code lief reproduzierbar über 00\_run.py/config.yaml.

Ergebnisse: Längere OP-Dauer war mit höheren AKI-Odds assoziiert (OR/h ≈ 1,147; 95 %-KI 1,077–1,222; p < 0,001). Re-Operation zeigte niedrigere Odds (OR ≈ 0,236; 0,145–0,384; p < 0,001). Dauer×Re-OP war signifikant (p ≈ 3,9 × 10⁻⁴). Geschlecht zeigte keinen unabhängigen Effekt. Der Onset lag überwiegend 0–48 h nach OP. In der Cross-Validation erzielte das Modell ROC-AUC ≈ 0,637, PR-AUC ≈ 0,538 und Brier ≈ 0,238.

Schlussfolgerung & Fördernutzen: ehrapy/AnnData ermöglicht eine klare, wiederholbare Identifikation zentraler Risikofaktoren aus klinischen Routinedaten; prozessnahe Kennzahlen (v. a. OP-Dauer) sind entscheidend, Alter zeigt nichtlineare Effekte, Geschlecht keinen unabhängigen. Die Förderung unterstützt prospektive Validierung, die Erweiterung um intraoperative Signale (z. B. NIRS) und die Dissemination als offenes Analyse-Protokoll für Kinderherzzentren.

Schlagwörter: akutes Nierenversagen; Kinderherzchirurgie; Routinedaten; ehrapy; AnnData; KDIGO; logistisches GLM; Reproduzierbarkeit.

2: Hintergrund: Akutes Nierenversagen (AKI) manifestiert sich als eine häufige und schwerwiegende Komplikation nach Kinderherzoperationen, die typischerweise in den ersten 48 Stunden nach dem Eingriff auftritt und die Prognose signifikant verschlechtert. An der Medizinischen Hochschule Hannover (MHH) steht ein umfassender, AKI-gelabelter Datensatz aus einem klinischen Entscheidungsunterstützungssystem (CDSS) zur Verfügung, der prä-, intra- und postoperative Routinedaten von über 1.300 Operationen enthält. Die effektive Analyse dieser komplexen Daten stellt eine zentrale Herausforderung dar.

Die Zielsetzung der vorliegenden Arbeit besteht in der Evaluation der Effektivität des AnnData-basierten EHRAPY-Frameworks hinsichtlich der transparenten und reproduzierbaren Identifikation von AKI-Risikofaktoren aus einem angereicherten klinischen Routinedatensatz.

Methode: Die vorliegende Untersuchung stützt sich auf ein Analyse-Subset von 1.209 OP-Episoden. Der Endpunkt der Untersuchung war AKI in den ersten sieben Tagen nach der Operation (definiert nach KDIGO-Kriterien). Die vorliegende Datensequenz wurde in der AnnData-Datei im HDF5-Format harmonisiert. Es wurde ein logistisches GLM mit natürlichen Splines für das Alter, cluster-robusten Standardfehlern (Clustering: Patient) und untersuchten Kovariaten wie OP-Dauer [h], Re-Operation (ja/nein) sowie Geschlecht analysiert. Darüber hinaus wurde die Interaktion zwischen der Dauer der Operation und dem erneuten operativen Eingriff getestet. Ergänzend wurden Onset-Analysen (Zeit bis zum AKI-Ereignis) und eine fünffache Cross-Validation durchgeführt. Die vollständige Reproduzierbarkeit des Analyseprotokolls wurde durch Skripte in 00\_run.py/config.yaml sichergestellt.

Die Ergebnisse der Untersuchung lassen sich wie folgt zusammenfassen: Eine längere Operation war hochsignifikant mit einer erhöhten Wahrscheinlichkeit für das Auftreten von AKI assoziiert (OR/h ≈ 1,147; 95%-KI 1,077–1,222; p < 0,001). Eine Re-Operation zeigte entgegen der Annahme signifikant niedrigere Odds (OR ≈ 0,236; 0,145–0,384; p < 0,001), wobei die Interaktion mit der OP-Dauer signifikant war (p ≈ 3,9 × 10⁻⁴). Es konnte festgestellt werden, dass das Geschlecht keinen unabhängigen Effekt aufweist. Der überwiegende Anteil der AKI-Fälle manifestierte sich demnach innerhalb der ersten 48 Stunden nach dem operativen Eingriff. In der Cross-Validation erzielte das Modell eine solide Performance (ROC-AUC ≈ 0,637; PR-AUC ≈ 0,538; Brier ≈ 0,238).

Schlussfolgerung & Fördernutzen: Die Resultate der Untersuchung demonstrieren die Fähigkeit des EHRAPY-Frameworks, eine eindeutige und wiederholbare Identifikation zentraler Risikofaktoren aus klinischen Routinedaten zu ermöglichen. Prozessnahe Kennzahlen wie die OP-Dauer erweisen sich als entscheidend, während das Alter nichtlineare Effekte veranschaulicht; das Geschlecht hingegen hat keinen unabhängigen Effekt. Eine potenzielle Förderung würde die prospektive Validierung dieser Ergebnisse, die Erweiterung der Datenbasis um intraoperative Signale (z. B. NIRS) und die Dissemination des offenen Analyse-Protokolls für andere Kinderherzzentren signifikant unterstützen.