Review of: "Machine learning-based prediction of acute kidney injury after nephrectomy in patients with renal cell carcinoma"

Vera von Dossow[1]

1 Herz- und Diabeteszentrum Nordrhein-Westfalen

Janis Fliegenschmidt, Vera von Dossow

Institute of Anesthesiology and Pain Therapy

Heart and Diabetes Center Bad Oeynhausen

Ruhr-University Bochum

Georgstrasse 11, 32545 Bad Oeynhausen, Germany

Re: Machine learning-based

prediction of acute kidney injury

after nephrectomy in patients

with renal cell carcinoma

Acute kidney injury (AKI) is among the most important postoperative complications as it is associated with an increased risk of mortality. Therefore, AKI prediction is of particular clinical importance, especially in renal cancer patients undergoing planned nephrectomy. The authors designed an interesting investigation comparing a logistic regression model versus machine-learning based prediction models of acute kidney injury after nephrectomy in patients with renal cell carcinoma. The principal finding is that machine learning models showed improved performance compared to a parallelly designed logistic regression model as well as an established risk score (SPARK index). This is of outstanding clinical importance as it allows for early intervention through perioperative prevention strategies.

There are several concerns to be addressed:

1. In the introduction section the authors report an incidence of 2-54 % for AKI after surgery. Even if this range is reported for several surgery types, the authors should focus on the reported incidence of AKI in urological cancer patients in the literature. Otherwise, it is impossible to interpret the importance of the findings of this study.

2. It is not clear from the report at which point during the patient's clinical course the machine-learned model is applied. Also, it is not clear from the description whether the machine learning algorithm has been applied to time series data or a curated data set. Risk stratification for serious adverse events should be implemented as early as possible: Intuitively, the range of applicable perioperative risk mitigation

strategies is dependent on the readiness of reliable information. Using time series data and generating predictions on the evolving image of the patient's condition could provide stronger insight into the model's predictive power.

3. When applying machine learning algorithms retrospectively on time-dependently evolving data, feature leak is a major concern. Feature leak describes constellations in which the machine learning algorithm has access to data which is only available due to the retrospective nature of the study but would not have been available at the point of time of the postulated evaluation. Since the algorithm is not being trained to retrospectively identify AKI from the medical record but rather to predict AKI, the authors should address how they dealt with this issue in the description of their methods.

These notes do not undercut the general value of this research paper. Artificial intelligence is a promising approach in preoperative risk stratification, and it can be expected to shape future decision processes and guide clinician attention. The applicability of the described models will have to be demonstrated in a prospectively designed study. Advancing the mathematical problem of labeling a set of patient data to be indicative of AKI vulnerability or not into a clinically relevant prediction of AKI is an important step in the more abstract quest of developing beneficial applications for artificial intelligence in medicine.

, Vera von Dossow1

Re: Machine learning-based
prediction of acute kidney injury
after nephrectomy in patients
with renal cell carcinoma

Acute kidney injury (AKI) is among the most important postoperative complications as it is associated with an increased risk of mortality. Therefore, AKI prediction is of particular clinical importance, especially in renal cancer patients undergoing planned nephrectomy. The authors designed an interesting investigation comparing a logistic regression model versus machine-learning based prediction models of acute kidney injury after nephrectomy in patients with renal cell carcinoma. The principal finding is that machine learning models showed improved performance compared to a parallelly designed logistic regression model as well as an established risk score (SPARK index). This is of outstanding clinical importance as it allows for early intervention through perioperative prevention strategies.

There are several concerns to be addressed:

1. In the introduction section the authors report an incidence of 2-54 % for AKI after surgery. Even if this range is reported for several surgery types, the authors should focus on the reported incidence of AKI in urological cancer patients in the literature. Otherwise, it is impossible to interpret the importance of the findings of this study.

2. It is not clear from the report at which point during the patient's clinical course the machine-learned model is applied. Also, it is not clear from the description whether the machine learning algorithm has

been applied to time series data or a curated data set. Risk stratification for serious adverse events should be implemented as early as possible: Intuitively, the range of applicable perioperative risk mitigation strategies is dependent on the readiness of reliable information. Using time series data and generating predictions on the evolving image of the patient's condition could provide stronger insight into the model's predictive power.

3. When applying machine learning algorithms retrospectively on time-dependently evolving data, feature leak is a major concern. Feature leak describes constellations in which the machine learning algorithm has access to data which is only available due to the retrospective nature of the study but would not have been available at the point of time of the postulated evaluation. Since the algorithm is not being trained to retrospectively identify AKI from the medical record but rather to predict AKI, the authors should address how they dealt with this issue in the description of their methods.

These notes do not undercut the general value of this research paper. Artificial intelligence is a promising approach in preoperative risk stratification, and it can be expected to shape future decision processes and guide clinician attention. The applicability of the described models will have to be demonstrated in a prospectively designed study. Advancing the mathematical problem of labeling a set of patient data to be indicative of AKI vulnerability or not into a clinically relevant prediction of AKI is an important step in the more abstract quest of developing beneficial applications for artificial intelligence in medicine.