Ambiguous baseline definitions affect automated AKI diagnosis at emergency department

Michael S A Niemantsverdriet[1,2], Meriem Khairoun[3], Ayman El Idrissi[3], Romy R Koopsen[3], Imo E Hofer[1], Wouter W Van Solinge[1], Jan Willem Uffen[3], Wouter M Tiel-Groenestege[1], Karin A H Kaasjager[3], Saskia Haitjema[1]

[1]*Central Diagnostic Laboratory, University Medical Center Utrecht, Utrecht University, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*
[2]*SkylineDx, Lichtenauerlaan 40, 3062 ME Rotterdam, The Netherlands*
[3]*Department of Internal Medicine, University Medical Center Utrecht, Utrecht University, Heidelberglaan 100, 3584 CX Utrecht, The Netherlands*

Background

Incorrect labeling of patients can hamper model development resulting in sub-optimal clinical decision support systems (CDSS) that can misclassify patients. During model development, patients are labeled, using guidelines postulated in literature, based on their retrospective routine electronic health record (EHR) data. Acute kidney injury (AKI) is identified by the Kidney Disease Improving Global Outcomes (KDIGO) criteria using changes in serum creatinine measurements. Creatinine measured at the emergency department (ED) is compared to a `baseline` measurement extracted from the patient's medical history. Depending on the definition of 'baseline' patients can be misclassified by the model. We evaluated the effect of multiple baseline-definitions on AKI prevalence in the ED.

Methods

47.190 ED-visits (19.956 patients) in the UMC Utrecht with a prior creatinine measurement between 2011-2019 were included from the Utrecht Patient-Orientated Database. An increase of 26,5 μmol/L creatinine between baseline-value and the ED-value was used as AKI definition (KDIGO). We analyzed four baseline-definitions: lowest, mean, median and most recent value from the patient's EHR. Multiple time intervals were used (≤365 days prior ED-presentation) to determine AKI-prevalence.

Results

The longest interval (365 days prior presentation) in combination with the lowest value as baseline resulted in the highest AKI-prevalence (12,65%) compared to the mean (4,23%), median (4,8%) and the most recent value (4,5%). Iteratively reducing the time window for extracting the creatinine measurement only showed extreme differences when using the lowest value as baseline. In comparison with the shortest interval (45 days) the longest interval increased the prevalence with 10,92% (5.151/47.190 additional AKI labels).

Conclusions

Using a specific definition of baseline, results in significantly different AKI prevalence in the ED. Adequate translation of guidelines to diagnose disease is crucial for accurate patient labeling to reduce misclassification by the model and to improve CDSS's accuracy to better support clinical decision making by treating physicians.