Table 2. STARDaki checklist for the reporting of diagnostic accuracy studies in acute kidney injury.

Item No.	STARD 2015 Original Text	STARDaki criteria
1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC).	- Identification as a study of diagnostic accuracy using at least one measure of accuracy (e.g., sensitivity, specificity, predictive values, AUC, or AUPRC). This item primarily applies to studies assessing the diagnostic accuracy for the presence of AKI, as defined by the KDIGO criteria. It particularly focuses on tools and methods for the early identification of AKI. While the item may also be relevant for other AKI related outcomes (such as development, duration, progression, and resolution) and related acute renal outcomes (e.g., components of major adverse kidney events [MAKE] or development of CKD, its core application refers to the diagnosis of AKI.
2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts).	unchanged
3	Scientific and clinical background, including the intended use and clinical role of the index test.	**Elaborated**: - If known, the proposed pathophysiology of the biomarker (e.g., podocyte damage, local reactive oxygen species, altered medullary blood flow) should be described. If the mechanism is not fully characterized, current hypotheses based on preclinical or clinical evidence should be summarized.

Item No.	STARD 2015 Original Text	STARDaki criteria
		- The intended clinical role of the biomarker should be explained (e.g., risk stratification, prognostic assessment and etiologic classification).
4	Study objectives and hypotheses.	unchanged
5	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study).	Whether participant enrollment and data collection were initiated before (prospective study) or after (retrospective study) the index test and reference standard were performed. For retrospective studies, it is recommended to briefly note measures taken to minimize bias in AKI adjudication (e.g., use of standardized AKI diagnostic criteria for retrospective review, independent adjudication by two investigators).

Item No.	STARD 2015 Original Text	STARDaki criteria
6	Eligibility criteria.	**Elaborated**: Clarify population selection based on clinical role: - Diagnostic studies (Early AKI detection): Enroll patients with clinical suspicion of AKI (e.g., acute illness, nephrotoxins) without prior AKI diagnosis. Define reference standard and specify baseline creatinine methods (detailed requirements see in 10b). - Prognostic studies (Predicting AKI outcomes): Enroll patients with confirmed AKI (KDIGO criteria). Define CKD inclusion/exclusion (e.g., eGFR <60 mL/min/1.73m^2 for >3 months) and analytical adjustments (stratification, multivariable models). The study population must comprise a single group of individuals sharing common AKI risk factors (e.g., critically ill patients, post-operative patients, or those exposed to nephrotoxins). Exploratory or hypothesis-developing studies comparing separate case/control cohorts–e.g., AKI patients vs. healthy controls–are not suitable for this purpose as they cannot correctly assess prognostic utility. If the study population includes predefined, clinically meaningful main subgroups (not minor or overlapping subgroups) that differ in core AKI risk factors (e.g., post-cardiac surgery vs. post-major abdominal surgery; contrast-induced AKI vs. nephrotoxin-induced AKI), each main subgroup should be clarified separately (e.g., list key inclusion criteria and sample size for each).
7	On what basis potentially eligible participants were identified (such	**Elaborated**: - Describe methods for participant identification: EHR screening (e.g., sepsis/ICU admission flags), lab

Item No.	STARD 2015 Original Text	STARDaki criteria
	as symptoms, results from previous tests, inclusion in registry).	monitoring (creatinine/biomarker thresholds), AKI registries, or clinician referral. - If applicable, specify the tools or algorithms used for screening (e.g., KDIGO criteria alerts in EHR systems).
8	Where and when potentially eligible participants were identified (setting, location and dates).	**Elaborated**: - Specify the location of the patients at time of enrolment (e.g., emergency department, cardiac critical care unit, general wards, etc.) - Report enrollment dates (start/end) and total study duration. - Clearly report the time of the exposure to AKI risk/triggering events (e.g., time of surgery, time of sepsis diagnosis). - If terms like "community-acquired" or "hospital-acquired" are used, explicitly state their operational definitions (e.g., "AKI diagnosed within 24h of admission was considered community-acquired"). - State how patients with unclear timing of occurrence or longitudinal timeframe for AKI events (diagnosis, progression, prognosis, etc.) were managed (e.g., excluded, analyzed as a separate subgroup, or classified based on best available evidence).

Item No.	STARD 2015 Original Text	STARDaki criteria
9	Whether participants formed a consecutive, random or convenience series.	unchanged
10a	Index test, in sufficient detail to allow replication.	**Elaborated**: - Clearly report assay platform (e.g., Research-grade vs. FDA/CE-certified clinical kits), sample handling (e.g., storage temperature, transport time, replicate testing), intra-/inter-batch CV% (e.g., ≤15% for urinary biomarkers), Lot-to-lot variability assessment (if using clinical kits), sampling-to-assay time window (e.g., urine samples refrigerated ≤2 hours). For urinary markers, indicate whether test results were indexed to urine creatinine or not.
10b	Reference standard, in sufficient detail to allow replication.	**Elaborated**: - Clearly report creatinine measurement (e.g., standardized pre-designed lab testing with a certain time interval). - Specify baseline and reference creatinine determination. How was baseline creatinine (historical values) determined? How was reference creatinine [creatinine used for AKI adjudication) determined? Were KDIGO recommendations for missing baseline creatinine used (the lower value of a. back-calculation of baseline using an eGFR equation and an eGFR of 75 (adults) or 120 (children); b. admission serum creatinine]? -Clearly report urine output recording methods (e.g., standardized pre-designed measurement with a

Item No.	STARD 2015 Original Text	STARDaki criteria
		certain time interval ≥24h). - Specify whether adjudication of AKI was blinded, ie performed by an independent expert panel (no target biomarker result access). - Clearly describe window for AKI diagnosis relative to biomarker sampling.
11	Rationale for choosing the reference standard (if alternatives exist).	Describe the rationale for choosing the AKI criteria, including specific stages (e.g., any stage vs stage 2-3) as the reference standard (including alternatives, if applicable). If AKI criteria are not used as the reference standard (e.g., in studies focusing on dynamic biomarkers such as AUC of daily creatinine clearance), justify the alternative reference standard and its relevance to AKI pathophysiology.
12a	Definition of and rationale for test positivity cutoffs or result categories of the index test, distinguishing pre-specified from exploratory.	**Elaborated**: Declare study objective: validation (testing pre-specified cutoffs defined a priori) versus exploratory (deriving new cutoffs). -If validating: state source (literature/manufacturer/biological rationale) and anchor to clinical outcomes (e.g., AKI stage-specific mortality). -If exploratory to derive cutoffs: specify derivation method (e.g., Youden index) and note requirement for external validation. -Technical specifications (for both scenarios): report handling of values <LLoD (e.g., 0.5 × LLoD imputation)

Item No.	STARD 2015 Original Text	STARDaki criteria
		and >ULoD (e.g., truncation rate); declare dynamic thresholds if used.
12b	Definition of and rationale for test positivity cutoffs or result categories of the reference standard, distinguishing pre-specified from exploratory.	unchanged
13a	Whether clinical information and reference standard results were available to the performers/readers of the index test.	unchanged
13b	Whether clinical information and index test results were available to the assessors of the reference standard.	**Elaborated**: -Mandatory blinding: reference standard assessors must be unaware of biomarker results.
14	Methods for estimating or comparing measures of diagnostic	**Elaborated**: The analytical methods for estimating diagnostic accuracy should be appropriate to the specific study

Item No.	STARD 2015 Original Text	STARDaki criteria
	accuracy.	design and research question. The following list provides examples of key methodological considerations that, **WHERE APPROPRIATE** to the study, should be properly designed and explicitly reported: - Longitudinal data handling, if available: Time-dependent analysis (e.g., landmark analysis for biomarker-AKI timing), generalized estimating equations (GEE) for repeated measures, cluster bootstrapping (resampling by patient, not by timepoint). - Subgroup analysis, if available: Pre-specified subgroups (e.g., sepsis-AKI, post-cardiac surgery AKI), Interaction term adjustment (e.g., AKI stage × biomarker interaction), Multiplicity control (Hierarchical testing: primary cohort → subgroups).
15	How were the indeterminate index test or reference standard results handled?	unchanged
16	How were missing data on the index test and reference standard handled?	**Elaborated**: - Clarify Creatinine imputation (e.g., linear interpolation ≤36h) or discard (e.g.,>36h gaps between protocol-driven measurements). - Clarify Urine output imputation (e.g., minimum 2-hour measurements) or discard (e.g., six measurement gaps).

Item No.	STARD 2015 Original Text	STARDaki criteria
17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory.	**Elaborated**: - Report subgroup variability, if available (e.g., sepsis vs. postoperative AKI sensitivity). - Interaction test results (e.g., CKD status impact).
18	Intended sample size and how it was determined.	**Elaborated**: - Present sample size formula with parameters (expected sensitivity/Specificity, margin of error, power). - For any pre-specified subgroup analyses, report the planned method for evaluating diagnostic accuracy in these subgroups (e.g., planned sample size or power calculation for key subgroups, or a acknowledgment of the exploratory nature of the subgroup analysis).
19	Flow of participants, using a diagram.	unchanged
20	Baseline demographic and clinical characteristics of participants.	**Elaborated**: - Specify risk exposures: Nephrotoxic (e.g., contrast volume, aminoglycoside duration), trauma, severe infection, major cardiac operation, etc. - Specify susceptibility and clinical status: Age, sex, baseline kidney function, comorbid conditions, and in applicable populations (e.g., critically ill), key hemodynamic parameters (e.g., mean arterial pressure less than a certain age-specific percentile or not, use or not of vasopressors).

Item No.	STARD 2015 Original Text	STARDaki criteria
21a	Distribution of severity of disease in those with the target condition.	**Elaborated**: - Clearly define the spectrum of the target condition (i.e., which KDIGO stages are included) and report the distribution of AKI severity accordingly. - Specify the number and proportion of participants in each included KDIGO stage (AKI Stage 1, 2, 3). - If each AKI stage has been divided in greater detail (e.g., Stage 1s, 1a, 1b), specification should also be provided. - Specify how many patients with the target condition received renal replacement therapy.
21b	Distribution of alternative diagnoses in those without the target condition.	**Elaborated:** - Clearly describe the clinical and diagnostic characteristics of participants without AKI (the control or comparator group). - Report the distribution of key alternative conditions or reasons for enrollment in these participants (for studies where alternative conditions exist), which may include, but are not limited to: Patients with stable CKD; Patients with AKI due to decreased effective blood volume; Patients with other acute illnesses but no evidence of AKI (e.g., sepsis without AKI, post-operative without AKI); Healthy controls, if applicable.

Item No.	STARD 2015 Original Text	STARDaki criteria
22	Time interval and any clinical interventions between the index test and reference standard.	**Elaborated**: - Define time windows based on the biomarker's clinical role: 1) Early detection: Time from biomarker sampling to AKI diagnosis (e.g., ≤24h for early prediction before serum creatinine rise). 2) Prognostic assessment: Time from biomarker sampling to clinical endpoint (e.g., 7 days for predicting renal replacement therapy or 30-day mortality). 3) Etiologic classification: Time from exposure/insult (e.g., nephrotoxic drug administration, sepsis onset) to biomarker sampling (e.g., six hours post-cardiopulmonary bypass to capture ischemia-reperfusion injury). - Report relevant clinical interventions (e.g., fluid resuscitation, diuretics) occurring between biomarker sampling and reference standard assessment that may affect interpretation. - Report Cr/UO data coverage: e.g., must span from 24h pre-sampling to 48h post-sampling. - Clarify endpoint independence: Blinded adjudication timing is irrelevant (analysis based on data collection period, not adjudication date). - Clarify protocol violations: e.g., exclude patients with incomplete Cr/UO data within the defined window.
23	Cross-tabulation of the index test results (or their distribution) by the results of the reference	unchanged

Item No.	STARD 2015 Original Text	STARDaki criteria
	standard.	
24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals).	**Elaborated**: - Report estimates of diagnostic accuracy with 95% confidence intervals, including sensitivity, specificity, AUC, PPV, NPV, LR+, and LR-. - Additionally, for studies where the target condition is relatively rare, consider reporting metrics relevant to imbalanced data, such as Precision (Positive Predictive Value, PPV), Recall (Sensitivity), the F1-score (harmonic mean of precision and recall), and Youden's Index (C-Statistic, to quantify overall discriminative ability and balance misdiagnosis/missed diagnosis risks). For studies using dynamic diagnostic models (e.g., continuous monitoring data) or requiring time-dependent evaluation, also consider reporting the Area Under the Activity Monitoring Operating Characteristic curve (AMOC) to reflect performance changes over time/data accumulation.
25	Any adverse events from performing the index test or the reference standard.	**Elaborated**: - Explicitly report sampling-related adverse events (even if zero).
26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability.	**Elaborated**: Discuss study limitations relevant to AKI diagnostic research, which should include, but are not limited to: - Reference standard related: Potential biases in clinical adjudication of AKI (e.g., lack of blinding to the

Item No.	STARD 2015 Original Text	STARDaki criteria
		index test, inter-adjudicator variability, subjectivity in baseline creatinine determination). - Data quality related: Sources of error or missingness in key variables (e.g., baseline creatinine estimation methods, gaps in serum creatinine or urine output monitoring). - Temporal and interventional confounding: Clinical events or interventions occurring between the index test sampling and the reference standard assessment that may confound the results. - Generalizability: The applicability of the findings to specific clinical settings or patient populations beyond the study cohort.
27	Implications for practice, including the intended use and clinical role of the index test.	**Elaborated**: As is. Additionally, for studies with sufficient evidence and where appropriate, authors are encouraged to discuss: - Actionable thresholds and their intended clinical use (e.g., a specific biomarker level that might trigger a nephrology consultation). - Potential health-economic implications (e.g., cost-effectiveness).

Item No.	STARD 2015 Original Text	STARDaki criteria
28	Registration number and name of registry.	As is.
29	Where the full study protocol can be accessed.	As is.
30	Sources of funding and other support; role of funders.	As is.

Abbreviation: AKI, Acute Kidney Injury; KDIGO, Kidney Disease: Improving Global Outcomes; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration; EHR, Electronic Health Record; ICU, Intensive Care Unit; GEE, Generalized Estimating Equations; VIF, Variance Inflation Factor; LLoD, Lower Limit of Detection; ULoD, Upper Limit of Detection; Cr, Creatinine; UO, Urine Output; AUC, Area Under the ROC Curve; PPV, Positive Predictive Value; NPV, Negative Predictive Value; LR+, Likelihood Ratio (Positive); LR-, Likelihood Ratio (Negative); NGAL, Neutrophil Gelatinase-Associated Lipocalin; KIM-1, Kidney Injury Molecule-1; FDA, Food and Drug Administration; CE, Conformité Européenne (European Conformity).