PROTOCOL

Review title: Diagnostic Utility of Copeptin in Pediatric Polyuria-Polydipsia Syndrome: A Systematic Review and Meta-Analysis

Objectives: to evaluate the diagnostic accuracy and utility of copeptin baseline dosing, copeptin diagnostic test after saline intravenous solution infusion, and arginine stimulation tests in pediatric populations presenting with polyuria-polydipsia syndrome

Databases searched include PubMed, Cochrane Library, Web of Science, ScienceDirect, Scopus, and Google Scholar, using specific terms related to polyuria-polydipsia syndrome and diagnostic tests.

Search date 03.08.2024

Limits - publication date between 2018-2024

URL to search strategy.
https://www.crd.york.ac.uk/PROSPEROFILES/576715_STRATEGY_20240822.pdf

Background:

Polyuria-polydipsia syndrome represents a diagnostic challenge for clinicians, especially in the pediatric population. This syndrome can be caused by different underlying conditions such as central diabetes insipidus (CDI), nephrogenic diabetes insipidus (NDI), and primary polydipsia (PP). Currently, the gold standard for diagnosis is the water deprivation test, which is followed by the administration of a synthetic antidiuretic hormone (ADH), aiming to evaluate the renal response to arginine-vasopressin (AVP) secretion.

This standard diagnostic test is uncomfortable and difficult to be performed in children. Recent studies suggest that measuring copeptin levels, which is a stable biomarker - representing the C-terminal part of the AVP precursor - can enhance diagnostic accuracy and reduce patient discomfort. Assessing the basal copeptin dosing and arginine or saline stimulation tests, is essential to establish their diagnostic accuracy in clinical settings for pediatric patients.

Population: We included studies performed with children and adolescents, that present the following clinical symptoms: polyuria, polydipsia, nocturia, electrolytic imbalances, and other related symptoms of central diabetes insipidus, nephrogenic diabetes insipidus, or primary polydipsia.

The exclusion criteria were mainly for studies that refer exclusively to disorders not directly related to diabetes insipidus, or studies assessing only adult patients. And also, patients with hyperosmolar polyuria, such as those with diabetes mellitus, will again be excluded to preserve the specificity of this review.

Index test: We considered the following index tests:

• Copeptin dosing: we selected the studies that involve measuring baseline copeptin levels, and studies that asses copeptin levels after hypertonic saline infusion or arginine stimulation. The protocol should aim to achieve a target serum sodium level of approximately 150 mmol/L, corresponding to a serum osmolality of around 300 mOsm/kg.

• Water Deprivation Test: we considered all studies that utilize a standardized water deprivation protocol, which involves depriving the patient of fluid for up to eight hours or until a 3% loss in body weight is achieved. During the test, plasma osmolality is measured at regular intervals to ensure an adequate rise that stimulates endogenous vasopressin release. Urine volume and osmolality are also monitored throughout the test. Additionaly, the patient is given desmopressin, then the urine volume and osmolality are measured to assess the response to exogenous vasopressin.

• Other relevant alternative diagnostic tests for patients presenting with diabetes insipidus related symptoms.

• All patients were eligible if presenting with hippo-osmolar polyuria, associated with secondary polydipsia, due to inadequate secretion of ADH or an abnormal renal response, or other related symptoms, excluding patients with hyperosmolar polyuria, such as in diabetes mellitus.

Reference standards:

• For copeptin dosing, studies must have included specific determination of baseline copeptin levels. Another test that will be assessed is represeented by the copeptin levels after intravenous infusion of 3% NaCl solution, as well as arginine stimulation test, or other copetin stimulation methods. These mechanisms are meant to achieve a target serum sodium level of approximately 150 mmol/L, corresponding to a serum osmolality of around 300 mOsm/kg.

• Additionally, studies comparing copeptin with other traditional diagnostic methods such as the water deprivation test. This test is usually extended for 8 hours, and followed by desmopressin administration, to assess the urinary reponce, od ADH. The study should have recorded serum sodium levels ranging from 145 mmol/L to 150 mmol/L and urinary osmolality ranging from 300 mOsm/kg to 1200 mOsm/kg.

These reference standards are essential to be included in the testing protocols, because they are considered optimal for stimulating hypothalamic osmoreceptors, leading to ADH release and subsequent renal urinary concentration as a physiological response.

Eligible Study Designs

• Comparative Primary Study Designs: Mainly we consider including comparative primary study designs where all patients undergo all tests, or patients are randomised to different tests.

• Non-comparative Primary Studies: Due to the limited number of eligible studies, the decision was made to extend the eligibility criteria to include other types of studies. Non-comparative primary studies where only one of the index tests has been investigated were also included. These studies were selected based on similar populations, diagnostic pathways, and described reference standards to minimize bias in our comparative DTA review.

• Other studies included: Due to the reduced number of such studies especially in pediatric population we also considered - Randomized controlled trials (RCTs), cohort studies, case-control studies, cross-sectional studies, observational studies, and case series.

We excluded studies focusing solely on conditions unrelated to diabetes insipidus, not involving pediatric populations or not using specific tests.

Our primary outcome is to assess the diagnostic accuracy of copeptin, in establishing the differential diagnostic of PPS. These measurements will be assessed using diagnostic accuracy metrics.

- Sensitivity to measures the proportion of actual positives patients with PPS, correctly identified by the test (copeptin), who have the actual condition, such as CDI;

- Specificity for the proportion of actual negatives patients correctly identified by the test who do not have the condition, such as ruling out the NDI or PP;

- Positive Predictive Value (PPV): representing the proportion of positive test results for copeptin that are true positives, meaning that the test correctly identifies those out of all positive results given;

- Negative Predictive Value (NPV) for the proportion of negative test results that are true negatives, meaning the copeptin test correctly identifies patients who do not have the condition out of the negative results given by the test;

- Likelihood Ratios (LR) with LR+ and LR- indicating the test's ability to correctly identify the condition from the patients with positive or negative tests;

- and Area Under the Receiver Operating Characteristic Curve (AUC) for measuring the test's overall ability to discriminate between those with and without the condition.

Meta-analyses are used, based on fixed-effects models, for the assessment of measures of diagnostic accuracy because the studies had shown low heterogeneity. Heterogeneity was assessed using the I^2 statistic, Cochran's Q test and τ^2 (τ^2) test. The statistical analyses were performed following standard meta-analysis protocols, to ensure robust and reliable estimates, using forest plots to visually compare diagnostic accuracy. Furtehrmore, we used HSROC model to assess threshold variability and obtained a deeper insight into the diagnostic accuracy across studies.

Secondary Outcomes:

We aimed to evaluate the consistency of diagnostic accuracy metrics across different diagnostic methods used in the studies reviewed. In this way we could assessed copeptin performance as a diagnostic test among all main diagnostic methods in pediatric patients with polyuria-polydipsia syndrome. Specifically, we analysed sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+), negative likelihood ratio (LR-), and area under the curve (AUC) for baseline copeptin levels, copeptin after stimulation (hypertonic saline and arginine), and the water deprivation test.

Another secondary outcome is the assessment of clinical implications, in selecting the most suitable diagnostic test, based on specific efficiency metrics, as well as establishing if copeptin is a reliable biomarker to be used as alternative tool for standard WDT in pediatric population.

Also we aim to identify the main gaps that will be a future target for further research, especially regarding copeptin testing protocols in pediatric patients, standard thresholds, for reducing variability among different diagnostic test, and introducing copeptin on a regular basis as a reliable biomarker for evaluation of children with PPS.

Adapted measure methods have been used, according to the final results. For instance, we conducted a sensitivity subgroup analysis for baseline copeptin levels, copeptin after stimulation with hypertonic saline solution or arginine stimulation and the water deprivation test. For this analysis we used a non-parametric test – Kruskal-Wallis - for comparing the subgroup since we cannot assume a normal distribution of the data. And further subtle variability was visually assessed using boxplots for each diagnostic method.

Study Selection:

1. First stage of screening: Two reviewers will independently use the search results to screen titles, abstracts, and all studies for relevance.

2. Full Text Review: Full text articles will be screened for eligibility based on the inclusion and exclusion criteria.

3. Agreement: In the event of disagreement, this will be addressed through discussion or, where there is a need for it, a third reviewer.

Data Extraction:

Before full data extraction, the extraction form was piloted on two studies, not included in this review to ensure clarity and consistency in data collection.

The **data extraction form** was extracted in duplicate and included the following key elements:

- Identification of the study: authors, year of publication, journal, country.

-Characteristics of participants: age, sex, clinical presentation.

- Index test details: copeptin dosing protocols, details for a water deprivation test.

- Reference standards: Criteria and measurements used.

- Diagnostic accuracy measures: sensitivity, specificity, PPV, NPV, LR+, LR−, AUC.

Results might include: raw data (e.g., individual 2 x 2 tables), summary statistics, subgroup analyses.

- Quality assessment: Risk of bias via the QUADAS-2 tool.

Process of extraction:

- Data extraction will be done independently by two reviewers.

- Extracted data will be compared and any discrepancy resolved by discussion or by a third reviewer.

Information will be recorded on standardised forms and transferred to a computer database for analysis.

Risk of Bias/Quality Assessment:

QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2) is used to assess the quality of diagnostic accuracy studies. It evaluates the risk of bias and applicability concerns in four key domains:

- Patient Selection: assess how participants were selected and the appropriateness of the selection criteria in projecting the possibility of selection bias

- Index Test: ssess the way how the test is conducted and interpretated.

- Reference Standard: This domain examines the validity and applicability of the reference standard that was used to classify subjects with regard to the target condition.

- Flow and Timing: In the methodology assessment it was examined whether investigations had been carried out into appropriate timing between index test and reference standard. Retrospective study designs often yield variability in timing, while prospective designs very rarely are.

- Applicability Concerns: Relevance of patient selection, index test, and reference standard to the review question.

Process:

Two independent reviewers will assess the risk of bias for each study using the QUADAS-2 tool.

Discrepancies will be resolved through discussion or by consulting a third reviewer.

Results of the risk of bias assessment will be synthesized and reported in a tabular format, highlighting areas of potential bias and applicability concerns.

Data Synthesis Methods used in this review:

Qualitative Synthesis:

-Narrative Summary: A comprehensive narrative synthesis will be provided to summarize the findings from the included studies. This will include descriptions of study characteristics, diagnostic methods used, and outcomes.

Quantitative Synthesis (Meta-Analysis) included:

-Models Used: fixed-effects models will be used due to the low variability among studies.

-Diagnostic Accuracy Measures: Pooled estimates for sensitivity, specificity, PPV, NPV, LR+, LR-, and AUC will be calculated.

-Heterogeneity Assessment: Statistical heterogeneity will be assessed using the I^2 statistic and Cochran's Q test. An I^2 value greater than 50% will indicate substantial heterogeneity.

-Subgroup Analyses: Where data allow, subgroup analyses will be performed based on different diagnostic methods (e.g., copeptin baseline, copeptin after stimulation).

-Sensitivity Analyses: Sensitivity analyses will be conducted to assess the robustness of the findings by excluding studies with high risk of bias

Additional HSROC model will be used to assess undetectable viability between studies and reinforce the findings of the standard meta-analysis.

Software: data analysis will be conducted using Python in IDLE (Python 3.14) environment, which offered a reliable number of statistical capabilities. Key libraries used included NumPy and Pandas for data manipulation, SciPy for statistical analysis, Matplotlib and Seaborn for data visualization. Advanced AI language models helped us in code generation to enhance precision and efficiency, with expert reviewing all the process to guarantee validity. To ensure result reliability, key analyses were cross-checked using RStudio software or IBM SPSS, confirming the consistency of findings derived from Python scripts.

For **data presentation** we will include:

- Forest Plots: Forest plots will be used to present the pooled estimates and confidence intervals for sensitivity, specificity, and other diagnostic accuracy measures.

-Summary Tables: Tables will summarize the characteristics of included studies, their risk of bias assessments, and the results of the meta-analyses. Funnel plots used for publication bias. Using HSROC model - provides summary ROC plot, for overall diagnostic performance of copeptin. In sensitivity analisis - boxplots will be used for visualization assessment of variability between diagnostic methods.

By employing these methods, the review aims to provide a rigorous and comprehensive synthesis of the diagnostic accuracy of copeptin and other tests in pediatric polyuria-polydipsia syndrome.

Planned Subgroup Investigations

A sensitivity analysis, is planned, aiming to evaluate the consistency of diagnostic accuracy metrics across different diagnostic methods used in the studies reviewed. In this way we could assessed copeptin performance as a diagnostic test among all main diagnostic methods in pediatric patients with polyuria-polydipsia syndrome. Specifically, the focus is on sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+), negative likelihood ratio (LR-), and area under the curve (AUC) for baseline copeptin levels, copeptin after stimulation (hypertonic saline and arginine), and the water deprivation test.

A non-parametric test – Kruskal-Wallis - for comparing the subgroup is used since we cannot assume a normal distribution and traditional tests, such as ANOVA, are not applicable in our case.

In the subgroup analysis we also want to draw attention to the subtle differences that may not be captured by statistical tests alone. As consequence, especially due to the fact that the studies regarding diagnostic test in pediatric population for PPS are limited, therefor is a small sample size for the Kruskal-Wallis test, we assure further assessment of the potential differences in diverse diagnostic methods using visual representations with boxplots of each diagnostic metrics.

These subgroup analyses aim to identify variations in diagnostic accuracy based on different diagnostic methods and patient characteristics, providing insights into the most effective diagnostic approaches for specific patient populations