RESEARCH ARTICLE

Arginine vs. Hypertonic Saline-Stimulated Copeptin to Diagnose AVP Deficiency

Julie Refardt[1-3], MD; Cihan Atila[1,2], MD; Irina Chifu[4], MD; Emanuele Ferrante[5], MD; Zoran Erlic[6], MD; Juliana B. Drummond[7], MD; Rita Indirli[5,8], MD; Roosmarijn Drexhage[3], MD; Clara O. Sailer[1,2], MD PhD; Andrea Widmer[1,2], MD; Susan Felder[1,2], MD; Andrew S. Powlson[9], MD; Nina Hutter[1,2], nurse; Deborah R. Vogt[2], PhD; Mark Gurnell[9,10], Prof; Beatriz Santana Soares Rocha[7], Prof; Johannes Hofland[3], MD, PhD; Felix Beuschlein[6,11,12], Prof; Martin Fassnacht[4,13], Prof; Bettina Winzeler[1,2], MD; Mirjam Christ-Crain[1,2], Prof;

[1]Departments of Endocrinology, Diabetology and Metabolism University Hospital Basel; Basel, Switzerland. [2]Department of Clinical Research, University Hospital Basel, University of Basel; Basel, Switzerland. [3]Department of Internal Medicine, Section of Endocrinology, Erasmus Medical Center, Rotterdam, The Netherlands. [4]Division of Endocrinology and Diabetes, Department of Internal Medicine I, University Hospital, University of Wuerzburg, Wuerzburg Germany. [5]Endocrinology Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy. [6]Department of Endocrinology, Diabetology, and Clinical Nutrition, University Hospital Zurich (USZ) and University of Zurich (UZH), Zürich Switzerland. [7]Department of Internal Medicine, Medical School of the Federal University of Minas Gerais, Belo Horizonte, MG, Brazil. [8]University of Milan, Department of Clinical Sciences and Community Health, Milan, Italy. [9]Wellcome-MRC Institute of Metabolic Science, University of Cambridge & Addenbrooke's Hospital, Cambridge Biomedical Campus, Cambridge, UK. [10]Cambridge NIHR Biomedical Research Centre, Cambridge University Hospitals, Cambridge, UK. [11]Medizinische Klinik und Poliklinik IV, Klinikum der Universität, Ludwig-Maximilians-Universität München, Munich, Germany. [12]The LOOP Zurich - Medical Research Center, Zurich, Switzerland. [13]Central Laboratory, University Hospital Wuerzburg, Wuerzburg, Germany

Short title: Stimulated copeptin to diagnose AVP-deficiency

Key words: diabetes insipidus, polyuria polydipsia syndrome, arginine vasopressin deficiency, primary polydipsia, hypertonic saline, AVP

Word count manuscript: 2953

Corresponding author:

Dr. Julie Refardt

Department of Endocrinology, University Hospital Basel, Basel, Switzerland

Petersgraben 4, 4031 Basel, Switzerland

Julie.refardt@usb.ch

ABSTRACT

Background

Distinguishing arginine-vasopressin-deficiency (AVP-D) from primary polydipsia (PP) is challenging. Hypertonic saline-stimulated copeptin diagnoses AVP-D with a high accuracy, but requires close sodium monitoring. Arginine-stimulated copeptin has shown similar diagnostic accuracy, but with a simpler test protocol.

We performed a head-to-head comparison hypothesizing arginine-stimulation to be non-inferior to hypertonic-saline stimulation in diagnosing AVP-D.

Methods

Randomized multicenter study conducted between 2018-2022. Patients underwent diagnostic evaluation with hypertonic saline and arginine stimulation. Two endocrinology experts independently made the final diagnosis with use of clinical information, treatment response and the hypertonic saline test results. The primary endpoint was overall diagnostic accuracy using pre-defined copeptin cut-offs 3.8pmol/L after 60 minutes for arginine and 4.9 pmol/L once sodium was >149 mmol/L for hypertonic saline stimulation.

Results

Of the 158 patients who underwent both tests, 69 (44%) were diagnosed with AVP-D and 89 (56%) with PP. The diagnostic accuracy [95% CI] according to the final diagnosis to differentiate AVP-D from PP patients was 74.4% [67.0, 80.6] for arginine- compared to 95.6% [91.1, 97.8] for hypertonic saline-stimulated copeptin. Accordingly, arginine stimulation was inferior to hypertonic saline stimulation (estimated difference -21.2% [-28.7, -14.3]).

Adverse events were mild for both tests, 72% (n=103/143) of patients preferred arginine to hypertonic saline stimulation. Arginine-stimulated copeptin ≤3.0 pmol/L diagnosed AVP-D with a specificity [95%CI] of 90.9% [81.7, 95.7], while levels >5.2 pmol/L diagnosed PP with a specificity of 91.4% [83.7, 95.6].

Conclusion

In the diagnostic evaluation of AVP-D, copeptin upon hypertonic saline stimulation was superior to arginine stimulation.

ClinicalTrials.gov Number: NCT03572166

Arginine vasopressin deficiency (AVP-D, formerly known as central diabetes insipidus) and AVP resistance (AVP-R, formerly known as nephrogenic diabetes insipidus) must be differentiated from primary polydipsia (PP), defined by excessive fluid intake despite initial adequate AVP secretion and renal response[1,2]. AVP-D is characterized by inadequate AVP release, whereas AVP-R results from renal insensitivity to AVP[3,4]. Complete and partial dysfunction have been described in both forms[5]. Differentiation from PP is critical, since treatment differs and potential misdiagnosis carries the risk of severe complications[1]. While the indirect water deprivation test was once considered the diagnostic standard[5], several studies have shown that water deprivation has low diagnostic accuracy and places a high clinical burden on patients[6,7]. After the establishment of copeptin as a stable and quick osmo-sensitive surrogate marker for AVP[8,9], the direct test approach was rediscovered[10]. Whereas unstimulated copeptin levels using a cut-off of >21.4 pmol/L can be used to diagnose AVP-R[11,12], a stimulated copeptin is required to differentiate between AVP-D and PP. A large multicenter trial showed a high diagnostic accuracy (96.5%) for hypertonic saline-stimulated copeptin to diagnose AVP-D[6]. A downside of that approach is the need for frequent sodium monitoring to avoid overstimulation and the discomfort due to induced hypernatremia.

An alternative test has been suggested through the use of arginine-stimulated copeptin, which showed similar diagnostic accuracy (93%) with a simpler, well-tolerated test protocol[13]. Based on these results, arginine-stimulated copeptin would seem preferable to hypertonic saline-stimulated copeptin as a standard test to differentiate between AVP-D and PP, but a prospective head-to-head comparison is lacking.

This randomized international multicenter trial aimed to compare the diagnostic accuracy of these tests. We hypothesized that the diagnostic accuracy of arginine-stimulated copeptin is non-inferior to hypertonic saline-stimulated copeptin.

METHODS

Study Design and Participants

This international multicenter study was conducted at seven tertiary medical centers in Europe and Brazil from September 2018 to September 2022 with a three-month follow-up completed in December 2022. The local ethics committees of all centers approved the study protocol, and written informed consent was obtained from all patients prior to any study procedure.

Adult patients with polydipsia (>3L/day, self-reported) and hypotonic polyuria (>50ml/kg body weight in a 24-hour urine collection and urine osmolality <800mOsm/kg) or patients with a known diagnosis of AVP-D were recruited. Patients with AVP-R or polyuria-polydipsia secondary to other causes (diabetes mellitus, hypercalcemia or hypokalemia) were excluded. Additional exclusion criteria were epilepsy requiring treatment; uncontrolled arterial hypertension, heart failure; liver cirrhosis; uncorrected adrenal or thyroid hormone deficiency; pregnancy or breastfeeding; or any relevant acute or terminal illness. Further details are provided in the study protocol at *NEJM.org*.

Procedures

Baseline Assessment

After obtaining a detailed medical history, a standardized clinical and biochemical evaluation was performed. Pituitary magnetic resonance imaging (MRI) was recommended for all study participants. MRIs were assessed for general allusions of the pituitary (i.e., pituitary lesions, post-operative changes) and typical AVP-D characteristics[14,15].

Participants were randomized to undergo either arginine or hypertonic saline stimulation first, on two different days. Tests were performed in the morning after an overnight fast, fluid intake was allowed until 6 a.m. Desmopressin treatment was paused 24 hours before the tests, or for a minimum of 12 hours in severely symptomatic AVP-D patients. Treatment was restarted after completion of the given test. Patients who were receiving hydrocortisone received an individualized stress dose.

Arginine Stimulation test

Infusion of 0.5 g per kg body weight (maximum 40 g) of L-Arginine Hydrochloride (21%) diluted in 500ml normal saline (NaCl 0.9%) was given over 30 minutes. Blood samples for copeptin measurement were collected before, 60 and 90 minutes after starting the infusion.

The diagnosis was made at the end of the study according to predefined copeptin cut-off levels at 60 minutes[13]-- Copeptin < 2.4 pmol/L = complete AVP-D; Copeptin 2.4 – 3.8 pmol/L = partial AVP-D; Copeptin > 3.8 pmol/L = PP.

Hypertonic Saline Infusion Test

Participants received appropriate venous access on both arms-- one for the infusion and one for blood sampling. After a 250 ml bolus of hypertonic saline (NaCl 3%), the infusion was continued at a rate of 0.15 ml per kg body weight per minute. Blood samples were drawn every 30 minutes. Sodium levels were additionally monitored by rapid venous blood gas analysis (vBGA). The infusion was stopped once sodium level in the vBGA reached >149 mmol/L, followed by immediate copeptin measurement. Once sampling was completed, patients received an oral water load (30ml per kg bodyweight) and 500 ml infusion of 5% glucose within 60 minutes. Patients were discharged once normonatremia was reached.

The diagnosis was made at the end of the test according to pre-defined copeptin cut-off levels[6] measured at a sodium level >149 mmol/L -- Copeptin < 2.7 pmol/L = complete AVP-D; Copeptin 2.7 – 4.9 pmol/L = partial AVP-D; Copeptin > 4.9 pmol/L = PP.

Assessment of Test Burden and Adverse Events

Test burden and predefined clinical symptoms (i.e., thirst, vertigo, headache, nausea, malaise), were assessed and rated according to a visual analog scale (VAS) (0 = no sensation/burden to 10 = maximum sensation/burden).

Preliminary Diagnosis and Assessment of Treatment Response

After completing both tests, patients were discharged with a provisional diagnosis and treatment. At the three-month follow-up visit, treatment response and clinical outcome were assessed and the preliminary diagnosis re-evaluated.

Final Diagnosis

The final diagnosis was independently made by two endocrine experts after consideration of the patient's medical history and clinical symptoms, laboratory and imaging data, the results of the hypertonic saline-stimulation test and the therapeutic response at three-month follow-up. The experts were blinded to the arginine stimulation test results and their diagnoses were not bound to the hypertonic saline stimulation test results. In the event of a discordant assessment (n=6), a third endocrine expert was consulted.

Once patients were classified as AVP-D or PP, distinction into partial or complete AVP-D was mainly based on the predefined hypertonic saline stimulated copeptin cut-off values[6], but could be overruled based on clinical information.

Laboratory Measurements

Laboratory measurements were performed by automated biochemical analyses in study center laboratories. Serum sodium levels were analyzed by indirect Ion selective electrode (ISE) method and vBGAs were measured by direct ISE method.

Hypertonic saline-stimulated copeptin levels were measured immediately following a completed test by the according study center; arginine-stimulated serum copeptin levels were measured centrally at the end of the study. All copeptin measurements were performed using the BRAHMS Copeptin proAVP assay (Thermo Fisher Scientific, Hennigsdorf, Germany; details in *Supplementary Appendix*).

Sample size

Sample size was estimated as 139 patients to achieve at least 80% power at a two-sided 5% type I error rate to show the non-inferiority (non-inferiority margin: -10%) of the overall diagnostic accuracy of the arginine stimulation test to the hypertonic saline stimulation test with assumed values of 93%[13] and 96.5%[6], respectively. To address an assumed drop-out rate of 8%, the recruitment goal was 152 participants. For details see *Supplementary Appendix*.

Outcomes

The primary outcome was overall diagnostic accuracy in differentiating AVP-D from PP using the pre-defined copeptin cut-offs, i.e., the ratio of correctly diagnosed patients (true positives + true negatives) to all patients with a final diagnosis.

Secondary outcomes were test tolerability and preference and the diagnostic performance of previously derived and predefined copeptin cut-offs: a) AVP-D versus PP: 3.7 pmol/L after 60 minutes and 4.1 pmol/L after 90 minutes for arginine stimulation[13]; 6.5 pmol/L for hypertonic saline stimulation[6]; b) complete versus partial AVP-D: 2.4 pmol/L after 60 minutes and 2.6 pmol/L after 90 minutes for arginine stimulation[13]; 2.7 pmol/L for hypertonic saline stimulation[6].

Statistical Analysis

Overall diagnostic accuracy, sensitivity, specificity, positive and negative predictive values with 95% Wilson confidence interval (CI) are indicated for each examined copeptin cut-off and the area under the receiver-operator characteristics curve (ROC AUC) was estimated with bootstrap 95% CI for each test procedure. The difference in the diagnostic accuracy between arginine and hypertonic stimulation was calculated with 95% CI by applying the “Tango” method for matched pairs[16] and compared to the non-inferiority margin.

The diagnostic potential of arginine-stimulated copeptin was explored by deriving “best” cut-offs using Youdens J statistic (jointly maximizing sensitivity and specificity).

Demographic information, laboratory parameters and test tolerability were described using median [IQR] or frequency and percentages. All diagnostic analyses were performed on two

modified intention-to-treat sets: mITT1 = all patients with a final diagnosis; mITT2 = mITT1 excluding patients with severe nausea / vomiting (post-randomization exclusion). Safety analyses were performed on the intention-to-treat set (ITT), including all patients starting at least one diagnostic test. The widths of the CIs have not been adjusted for multiplicity and may not be used in place of hypothesis testing.

All analyses were predefined and conducted using the statistical software package R (version 4.2.3)[17].

Author Contributions

Study design: JR, DV, BW, MCC. Data gathering: all authors. Data analysis: JR, CA, DV; all authors vouch for the data and analysis. First manuscript draft: JR; revision and decision to publish: all authors. Confidentiality agreements of the data existed between the sponsor and all authors until publication.

RESULTS

Patient flow and characteristics

In all, 177 patients were included, of whom 13 were excluded post-randomization (*Supplementary Fig. S1*). Of the remaining164 patients (ITT), six patients withdrew consent after the first diagnostic test (five after arginine stimulation, one after hypertonic saline stimulation). These patients were included in the safety but not the primary analysis. The remaining 158 patients underwent both diagnostic tests, received a final diagnosis and were evaluated for the primary outcome (mITT1). The pre-defined mITT2 set excluded 22 patients experiencing severe nausea (VAS ≥8) and / or vomiting during the tests. The median [IQR] interval between both tests was four days [1-8].

Our study population corresponds to the general published patient population (*Supplementary Table S1*). Of the 158 patients (67% female) with a final diagnosis, 69

(44%) were diagnosed with AVP-D -- 41 (59%) with complete and 28 (41%) with partial deficiency - and 89 (56%) with PP (*Table 1, Table S2*).

The main etiologies for AVP-D were post-surgical (30%), hypothalamic-pituitary lesions (26%), hypophysitis (12%) and idiopathic AVP-D (12%). Twenty-nine (42%) AVP-D patients also had anterior pituitary deficiencies.

Patients with complete AVP-D had higher quantities of polyuria und polydipsia compared to patients with partial AVP-D and PP (*Table 1*). Similar observations were made for baseline copeptin and urine osmolality levels, which were lowest for complete AVP-D patients. Patient characteristics randomized to arginine stimulation (n=78) or hypertonic saline infusion first (n=80) were comparable (*Table S3*)

Pituitary MRI was performed in 68% (n=108) of patients. Characteristics typical for AVP-D were observed in 67 (62%) cases, of which 58 were later diagnosed with AVP-D (*Table 1*).

Primary Outcome

The overall diagnostic accuracy [95% CI] to differentiate patients with AVP-D from those with PP was 74.4% [67.0, 80.6] for the arginine stimulation and 95.6% [91.1, 97.8] for the hypertonic saline stimulation test (estimated difference: -21.2% [-28.7, -14.3]; *Fig. 1, Table 2, Fig. S2*). Accordingly, arginine stimulation was inferior to hypertonic saline stimulation. The ROC AUC [95% CI] for arginine-stimulated copeptin was 0.85 [0.80, 0.91] and for hypertonic saline-stimulated copeptin 0.99 [0.98, 1.00] (*Fig. S3*).

Test performance was comparable when excluding patients with severe nausea or vomiting (mITT2) with a diagnostic accuracy of 72.6% [64.5, 79.4] for arginine stimulation and 96.3% [91.7, 98.4] for hypertonic saline stimulation.

Arginine-stimulated copeptin also performed worse compared to hypertonic saline-stimulated copeptin in the differentiation between partial AVP-D and PP (*Table 2, Fig. 1 and Fig. S2*).

Misclassified patients are described in the *Supplementary Appendix* including *Table S4*.

Secondary Outcomes

Other pre-defined copeptin cut-offs resulted in similar overall diagnostic accuracies. Arginine-stimulated copeptin: 75.0% [67.7, 81.1] for 3.7 pmol/L after 60 minutes and 79.2% [72.1, 84.9] for 4.1 pmol/L after 90 minutes; hypertonic saline-stimulated copeptin: 96.2% [92.0, 98.2] for 6.5 pmol/L (*Table S5*).

The pre-defined hypertonic saline-stimulated copeptin cut-off of 2.7 pmol/L differentiated between complete and partial AVP-D with a diagnostic accuracy of 88.4% [78.8, 94.0] (sensitivity 92.7% [80.6, 97.5], specificity 82.1% [64.4, 92.1]).

Exploratory analyses of data-derived “best” cut-offs did not reveal a convincing performance (*Fig. S4)*. However, arginine-stimulated copeptin levels ≤3.0 pmol/L diagnosed AVP-D with a specificity of 90.9% [81.7, 95.7] (sensitivity 59.5% [49.1, 69.1]), while levels >5.2 pmol/L diagnosed PP with a specificity of 91.4% [83.7, 95.6] (sensitivity 56.4% [44.6, 67.4]); details in Table 3 and *Fig. S2*).

Applying these two cut-offs to our cohort (156 ITT patients set with copeptin measures available) allowed a correct test result in 91/156 patients (58.3% [50.5, 65.8]), while leading to an inclusive or incorrect test result in 48/156 (30.8% [24.1, 38.4] and 17/156 (10.9% [6.9, 16.8]) patients, respectively.

Safety Outcomes

Generally, participants reported that both tests were acceptable *(Table 4)*, (details on test duration in the *Supplementary Appendix)*. Nearly all patients reported severe thirst (median [IQR] VAS 8 [7-9]) at the end of the arginine stimulation test, followed by mild headache (37%, VAS 3 [2-5.5]) and malaise (32%, VAS 3.5 [2-5.5]) as frequent adverse effects. Severe thirst was also the main adverse effect of the hypertonic saline stimulation test (98%, VAS 9 [8-10]), followed by mild headache (59%, VAS 4 [3-7]) and malaise (51%, VAS 5 [3-7]).

The overall intensity of adverse effects was low in both tests but occurred with more frequency and higher intensity during the hypertonic saline stimulation test.
No severe adverse event occurred during either test. No adverse events were noted in the six participants who withdrew consent after the first test. Overall, the majority of patients (72%) preferred the arginine to the hypertonic saline infusion test.

DISCUSSION

Arginine-stimulated copeptin was inferior to hypertonic saline-stimulated copeptin in the diagnosis of AVP-D and showed a greater overlap between AVP-D and PP patients. Hypertonic saline-stimulated copeptin thus remains the test with higher diagnostic accuracy and safety. Nevertheless, arginine stimulation was preferred by patients, and arginine-stimulated copeptin levels ≤3.0 pmol/L and >5.2 pmol/L showed high specificity in correctly diagnosing over half of patients with AVP-D and PP.

The diagnostic performance of arginine stimulation in this cohort was lower than the previously described 93%[13,18]. The prior accuracy was derived from a smaller monocentric cohort with a similar distribution of AVP-D and PP patients (n=96, AVP-D: 40%, PP 60%[13]). According to the mITT2 set, severe nausea or emesis – which are non-osmotic copeptin stimuli[19,20] - were responsible for copeptin-overstimulation in three patients. The worse performance of arginine and weaker than expected copeptin stimulation may have several

reasons. First, symptom severity of the PP patients was more accentuated in the current cohort. Although the amount of polyuria / polydipsia was similar, PP patients in the current cohort had a lower baseline serum and urinary osmolality. Possibly, diagnostic accuracy of arginine stimulation could be improved by raising serum osmolality by overnight water deprivation. Second, the 40g upper limit arginine dosage used in this protocol may have led to weaker stimulative potency in the three obese patients. Lastly, arginine stimulation had a stronger comparator in the current study. Whereas in the previous evaluation[13] the water deprivation test (known diagnostic accuracy 70-77%[6,7]) was part of the expert diagnosis, here it was the hypertonic saline test.

This highlights the second important finding of this study: The high diagnostic accuracy of the hypertonic saline test was validated with 95.6%[6]. Of note, hypertonic saline-stimulated copeptin also differentiated reliably between partial and complete AVP-D.

The adverse effects were only mild to moderate and limited to the duration of the infusion. Regular rapid sodium measurements avoided sodium overstimulation and guaranteed safety of the test. This emphasizes the utility and reliability of the hypertonic saline test as the gold standard for the diagnosis of AVP-D.

However, the hypertonic saline test has some limitations: it can only be performed in patients in whom appropriate venous accesses can be placed and in settings with the availability of constant surveillance and rapid sodium measurements. In addition, there are only limited safety data in patients older than 65 years and several comorbidities (exclusion criteria in this study) prevent patients from this diagnostic evaluation.
Arginine stimulation has been known for decades for the evaluation of the anterior pituitary[21,22]. Most clinicians are familiar with its protocol which can be performed in the out-patient setting. Compared to hypertonic saline stimulation, it is shorter, was better tolerated and preferred by patients. Accordingly, its use can be recommended as a simple well-

tolerated initial diagnostic test. In this regard, arginine stimulation is also preferable to the water deprivation test.

While arginine stimulation did not result in a single optimal copeptin cut-off value, it showed high specificity in diagnosing AVP-D and PP patients using the copeptin cut-offs of ≤3.0 and >5.2 pmol/L. Patients with copeptin levels between the above cut-offs or experiencing severe nausea or vomiting during the arginine stimulation should, however, undergo hypertonic saline stimulation for further evaluation.

The importance of a reliable diagnostic test was emphasized by overlapping clinical and laboratory characteristics of partial AVP-D and PP patients, who showed no difference in the amount of polyuria and polydipsia nor urine osmolality. Pituitary MRI was performed in two thirds of the patients. In these patients with a high pre-test probability, findings typical for AVP-D were seen in 58 of 67 cases. Conversely, there were several AVP-D patients without abnormalities and PP patients with false positive results. Accordingly, MRI findings should always be evaluated in the clinical context.

The main limitation of our study is the absence of a clear diagnostic standard for AVP-D. While the diagnoses were based on careful review of all patient data, it also included the outcome of the hypertonic saline stimulation. To overcome this incorporation bias, treatment response at three months was integrated into the final diagnosis. Nevertheless, the diagnostic value of hypertonic saline-stimulated copeptin may be overestimated. The strength of this study is the randomized international design and large sample size of well-characterized patients with AVP-D and PP.

In conclusion, in the diagnosis of AVP-D, hypertonic saline-stimulated copeptin was superior to arginine-stimulated copeptin.

ACKNOWLEDGEMENTS

We thank all patients for their participation and the medical and laboratory staff at all study sites for their contribution to this study. Special thanks go to all study nurses and staff for their invaluable support.

DISCLOSURES AND FUNDING

JR was supported by a grant from the Swiss National Science Foundation (P2BSP3-181720) and the Goldschmidt Jacobson Foundation. CA received the Young Talents in Clinical Research grant from the Swiss Academy of Medical Sciences and the G&J Bangerter-Rhyner Foundation. JD was supported by a grant of the Minas Gerais Research Support Foundation (FAPEMIG-APQ-01521-21). MG was supported by the NIHR Cambridge Biomedical Research Centre (the views expressed are those of the author and not necessarily those of the NIHR or the Department of Health and Social Care). MCC was supported by a grant of the Swiss National Science Foundation SNF-162608. All remaining authors have nothing to disclose. Laboratory measurement of copeptin was funded by Thermo Fisher Scientific.

The funders had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

DATA SHARING STATEMENT

De-identified individual participant data that underlie the results reported in this article will be shared upon publication to researchers who provide a methodologically sound proposal to achieve the aims in the approved proposal. Proposals should be directed to the corresponding author. To gain access, data requestors will need to sign a data access agreement.

REFERENCES

1. Christ-Crain M, Winzeler B, Refardt J. Diagnosis and management of diabetes insipidus for the internist: an update. J Intern Med 2021;290:73-87.
2. Arima H, Cheetham T, Christ-Crain M, et al. Changing the name of diabetes insipidus: a position statement of The Working Group for Renaming Diabetes Insipidus. Eur J Endocrinol 2022;187:P1-p3.
3. Christ-Crain M, Bichet DG, Fenske WK, et al. Diabetes insipidus. Nat Rev Dis Primers 2019;5:54.
4. Bockenhauer D, Bichet DG. Pathophysiology, diagnosis and management of nephrogenic diabetes insipidus. Nat Rev Nephrol 2015;11:576-88.
5. Miller M, Dalakos T, Moses AM, Fellerman H, Streeten DH. Recognition of partial defects in antidiuretic hormone secretion. Ann Intern Med 1970;73:721-9.
6. Fenske W, Refardt J, Chifu I, et al. A Copeptin-Based Approach in the Diagnosis of Diabetes Insipidus. N Engl J Med 2018;379:428-39.
7. Fenske W, Quinkler M, Lorenz D, et al. Copeptin in the differential diagnosis of the polydipsia-polyuria syndrome--revisiting the direct and indirect water deprivation tests. J Clin Endocrinol Metab 2011;96:1506-15.
8. Morgenthaler NG, Struck J, Alonso C, Bergmann A. Assay for the measurement of copeptin, a stable peptide derived from the precursor of vasopressin. Clin Chem 2006;52:112-9.
9. Christ-Crain M, Morgenthaler NG, Fenske W. Copeptin as a biomarker and a diagnostic tool in the evaluation of patients with polyuria-polydipsia and hyponatremia. Best Pract Res Clin Endocrinol Metab 2016;30:235-47.
10. Zerbe RL, Robertson GL. A comparison of plasma vasopressin measurements with a standard indirect test in the differential diagnosis of polyuria. N Engl J Med 1981;305:1539-46.
11. Timper K, Fenske W, Kühn F, et al. Diagnostic Accuracy of Copeptin in the Differential Diagnosis of the Polyuria-polydipsia Syndrome: A Prospective Multicenter Study. J Clin Endocrinol Metab 2015;100:2268-74.
12. Fenske WK, Christ-Crain M, Horning A, et al. A copeptin-based classification of the osmoregulatory defects in the syndrome of inappropriate antidiuresis. J Am Soc Nephrol 2014;25:2376-83.
13. Winzeler B, Cesana-Nigro N, Refardt J, et al. Arginine-stimulated copeptin measurements in the differential diagnosis of diabetes insipidus: a prospective diagnostic study. Lancet 2019;394:587-95.
14. Fujisawa I. Magnetic resonance imaging of the hypothalamic-neurohypophyseal system. J Neuroendocrinol 2004;16:297-302.
15. Kinoshita Y, Taguchi A, Tominaga A, Sakoguchi T, Arita K, Yamasaki F. Predictive factors of postoperative diabetes insipidus in 333 patients undergoing transsphenoidal surgery for non-functioning pituitary adenoma. Pituitary 2022;25:100-7.
16. Tango T. Equivalence test and confidence interval for the difference in proportions for the paired-sample design. Stat Med 1998;17:891-908.
17. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria 2021.
18. Binder G, Weber K, Peter A, Schweizer R. Arginine-stimulated copeptin in children and adolescents. Clin Endocrinol (Oxf) 2023;98:548-53.

19. Baylis PH. Osmoregulation and control of vasopressin secretion in healthy humans. Am J Physiol 1987;253:R671-8.
20. Brooks E, Bachmeier C, Vorster J, et al. Copeptin is increased by nausea and vomiting during hypertonic saline infusion in healthy individuals. Clin Endocrinol (Oxf) 2021;94:820-6.
21. Boguszewski CL, Boguszewski M, de Herder WW. The science behind the relations among cancer, height, growth patterns, and growth hormone axis. Endocr Relat Cancer 2023;30.
22. Yuen KCJ, Johannsson G, Ho KKY, Miller BS, Bergada I, Rogol AD. Diagnosis and testing for growth hormone deficiency across the ages: a global view of the accuracy, caveats, and cut-offs for diagnosis. Endocr Connect 2023;12.

TABLE AND FIGURES

Table 1 Patient Characteristics

	AVP-Deficiency			Primary Polydipsia
Characteristics	**Complete** n=41, 26%	**Partial** n=28, 18%	**All** n=69, 44%	n=89, 56%
Age, years	38 [31, 47]	50 [39, 58]	42 [32, 54]	37 [28, 50]
Female sex, n (%)	24 (59)	14 (50)	38 (55)	68 (76)
Body mass index, kg/m^2	29.5 [24.2, 33.8]	27.0 [25.0, 30.1]	27.6 [24.5, 33.0]	23.8 [21.0, 28.5]
Clinical symptoms at time of diagnosis				
Polydipsia, liters/day	7.0 [5.0, 9.0]	5.3 [3.9, 6.0]	6.0 [4.0, 8.0]	5.0 [4.0, 7.0]
Polyuria, liters/day	8.0 [6.0, 9.4]	4.8 [3.5, 6.2]	6.0 [4.2, 8.5]	4.8 [4.0, 6.5]
Emictions per day	15 [8, 20]	11 [8, 14]	12 [8, 15]	10 [9, 15]
Nocturia, n (%)	32 (78)	24 (86)	56 (81)	68 (76)
Nocturia, times/night	4 [3, 5]	3 [2, 4]	4 [3, 5]	3 [2, 3]
Drinking at night, n (%)	32 (78)	19 (68)	51 (74)	60 (67)
Drinking at night, liters/night	1.5 [1.0, 2.5]	0.8 [0.5, 1.1]	1.0 [0.5, 2.0]	0.7 [0.5, 1.0]
Medical history, n (%)				
Anterior pituitary insufficiency	16 (39)	13 (46)	29 (42)	5 (6)
Adrenocorticotropic hormone	14 (34)	11 (39)	24 (35)	2 (3)
Thyrotropic hormone	14 (34)	13 (46)	27 (39)	3 (4)
Growth hormone	4 (10)	4 (14)	8 (12)	1 (1)
Gonadotropins	13 (32)	10 (36)	23 (33)	3 (4)
Pituitary lesions	21 (51)	18 (64)	5 (7)	10 (11)
History of pituitary surgery	12 (29)	10 (36)	22 (32)	6 (7)
History of pituitary apoplexy	0 (0)	1 (1)	1 (1)	1 (1)
Psychiatric disorder	3 (7)	5 (18)	8 (12)	24 (27)
Cardiovascular disease	3 (7)	3 (11)	6 (9)	4 (5)
Cerebrovascular disease	3 (7)	2 (7)	5 (7)	1 (1)
Other	28 (68)	20 (71)	48 (84)	53 (60)
AVP-D etiology, n (%)				
Post-surgical AVP-D	10 (24)	11 (39)	21 (30)	NA
Hypothalamic-pituitary lesions	11 (27)	7 (25)	18 (26)	NA
Trauma	3 (7)	2 (7)	5 (7)	NA
Empty sella or hypoplasia	3 (7)	2 (7)	5 (7)	NA
Vascular (e.g., apoplexy, Sheehan syndrome)	0 (0)	1 (1)	1 (1)	NA
Hypophysitis	4 (10)	4 (14)	8 (12)	NA
Idiopathic AVP-D	7 (17)	1 (4)	8 (12)	NA

Familial AVP-D	3 (7)	0 (0)	3 (4)	NA
Laboratory data at baseline				
Serum sodium, mmol/L	142 [140, 143]	142 [140, 143]	142 [140, 143]	140 [138, 141]
Serum osmolality, mOsm/kg	293 [289, 296]	292 [290, 295]	293 [290, 296]	287 [283, 291]
Serum Copeptin, pmol/L	1.8 [1.4, 2.1]	2.7 [2.3, 3.4]	2.2 [1.6, 2.4]	2.6 [2.0, 3.9]
Urine osmolality, mOsm/kg	137 [90, 216]	230 [168, 312]	181 [108, 299]	222 [156, 431]
MRI characteristics, n (%)	39 (95)	25 (89)	64 (93)	44 (49)
Pituitary stalk enlarged	8 (20)	5 (20)	13 (20)	2 (5)
Bright spot absent	27 (71)	16 (64)	43 (68)	6 (13)
Enlargement of the posterior pituitary	7 (18)	3 (12)	10 (16)	2 (5)
Allusion to adenohypophysis / hypophysitis	4 (11)	6 (24)	10 (16)	0 (0)
Other findings (e.g. adenoma)	12 (31)	15 (63)	27 (43)	14 (32)

Table 1 Characteristics of all included patients who underwent both tests and received a final diagnosis (modified intention-to-treat-set 1).

Data presented as frequency (percentage) and median [IQR].

n=number, IQR=interquartile range, AVP-Deficiency=arginine vasopressin deficiency, HRT=hormonal replacement therapy, MRI = magnetic resonance imaging

Table 2 Diagnostic Measures of the Diagnostic Tests

Test	Accuracy	Sensitivity	Specificity	PPV	NPV	n
AVP-D vs PP						
Arginine-stimulated copeptin 3.8 pmol/L	74.4 [67.0, 80.6]	75.4 [64.0, 84.0]	73.6 [63.4, 81.7]	69.3 [58.2, 78.6]	79.0 [68.9, 86.5]	156
Hypertonic saline-stimulated copeptin 4.9 pmol/L	95.6 [91.1, 97.8]	91.3 [82.3, 96.0]	98.9 [93.9, 99.8]	98.4 [91.7, 99.7]	93.6 [86.8, 97.0]	158
partial AVP-D vs PP						
Arginine-stimulated copeptin 3.8 pmol/L	68.7 [59.7, 76.5]	53.6 [35.8, 70.5]	73.6 [63.4, 81.7]	39.5 [25.6, 55.3]	39.5 [25.6, 55.3]	115
Hypertonic saline-stimulated copeptin 4.9 pmol/L	95.2 [89.3, 97.9]	83.3 [64.1, 93.3]	98.8 [93.3, 99.8]	95.2 [77.3, 99.2]	95.2 [88.4, 98.1]	117

Diagnostic measures shown as % and 95% confidence interval of the two tests.

AVP-D=arginine vasopressin deficiency; PP=primary polydipsia; PPV=positive predictive value; NPV=negative predictive value; n=sample size

Table 3 Performance of Different Arginine-Stimulated Copeptin Values

To diagnose AVP-deficiency			To diagnose primary polydipsia		
Threshold	**Specificity**	**Sensitivity**	**Threshold**	**Specificity**	**Sensitivity**
2.2	98.9 [92.8, 99.9]	36.3 [27.1, 46.6]	4.2	78.3 [68.7, 85.6]	70.2 [58.5, 79.7]
2.3	97.8 [91.0, 99.5]	40.6 [31.0, 51.0]	4.4	79.8 [70.3, 86.8]	69.0 [57.3, 78.7]
2.4	97.8 [91.0, 99.5	42.1 [32.4, 52.5]	4.5	81.2 [71.9, 88.0]	69.0 [57.3, 78.7]
2.5	96.6 [89.3, 99.0]	46.4 [36.4, 56.7]	4.6	82.7 [73.5, 89.1]	69.0 [57.3, 78.7]
2.6	95.5 [87.7, 98.4]	49.3 [39.2, 59.5]	4.7	85.6 [76.8, 91.4]	66.7 [55.0, 76.7]
2.7	95.5 [87.7, 98.4	52.2 [42.0, 62.3]	4.9	87.0 [78.5, 92.5]	64.4 [52.6, 74.7]
2.8	94.3 [86.1, 97.8]	53.7 [43.4, 63.7]	5.0	88.5 [80.2, 93.6]	63.3 [51.5, 73.7]
2.9	90.9 [81.7, 95.7]	56.6 [46.2, 66.4]	5.1	89.9 [81.9, 94.6]	58.7 [46.9, 69.5]
3.0	**90.9 [81.7, 95.7]**	**59.5 [49.1, 69.1]**	**5.2**	**91.4 [83.7, 95.6]**	**56.4 [44.6, 67.4]**
3.1	87.4 [77.6, 93.3]	60.9 [50.5, 70.4]	5.4	92.8 [85.5, 96.6]	54.1 [42.4, 65.3]
3.2	86.3 [76.2, 92.5]	66.7 [56.4, 75.6]	5.5	92.8 [85.5, 96.6]	51.8 [40.2, 63.2]
3.3	84.0 [73.5, 90.8]	66.7 [56.4, 75.6]	5.6	94.3 [87.3, 97.5]	50.6 [39.1, 62.1]
3.4	81.7 [70.9, 89.0]	66.7 [56.4, 75.6]	5.9	94.3 [87.3, 97.5]	49.5 [38.0, 61.0]
3.5	81.7 [70.9, 89.0]	69.6 [59.4, 78.2]	6.2	94.3 [87.3, 97.5]	46.0 [34.8, 57.7
3.6	79.4 [68.4, 87.2]	71.1 [60.9, 79.5]	6.3	94.3 [87.3, 97.5]	42.6 [31.6, 54.3]
3.7	75.9 [64.6, 84.5]	72.5 [62.5, 80.7]	6.4	94.3 [87.3, 97.5]	41.4 [30.6, 53.2
3.8	75.9 [64.6, 84.5]	74.0 [64.0, 82.0]	6.6	94.3 [87.3, 97.5]	39.1 [28.5, 50.9]
3.9	73.6 [62.2, 82.6]	75.4 [65.5, 83.2]	6.7	95.7 [89.3, 98.4]	39.1 [28.5, 50.9]
4.0	73.6 [62.2, 82.6]	76.9 [67.1, 84.4]	6.9	95.7 [89.3, 98.4]	38.0 [27.5, 49.8]
4.1	71.3 [59.7, 80.6]	78.3 [68.7, 85.6]	7.0	95.7 [89.3, 98.4]	33.4 [23.4, 45.1]

Table 3 Arginine-stimulated copeptin levels and their according specificities and sensitivities [95% confidence interval] to diagnose AVP-deficiency and primary polydipsia accordingly. The proposed cut-offs with >90% specificity (based on point estimates) for each diagnosis are marked in bold / grey.

Note: The specificity to diagnose AVP-deficiency is the sensitivity to diagnose primary polydipsia and vice versa

Table 4 Adverse Events

Stimulation test	Arginine		Hypertonic saline	
	n (%)	VAS score	n (%)	VAS score
Adverse effects				
Thirst	158/160 (99)	8 [7-9]	155/158 (98)	9 [8-10]
Vertigo	42/160 (26)	3.5 [2-5]	75/158 (47)	5 [3-6.5]
Headache	59/160 (37)	3 [2-5.5]	94/158 (59)	4 [3-7]
Nausea	40/160 (25)	3.5 [1-6]	50/158 (32)	3.5 [2-7]
Malaise	52/160 (32)	3.5 [2-5.5]	81/158 (51)	5 [3-7]
Overall symptom burden		2 [0-3]		4 [2-7]
Adverse events				
Neuromuscular symptoms (agitation, blurred vision, muscle spasms, paresthesia, shivering, tremor)	6 (4)		23 (15)	
Emesis	11 (7)		9 (6)	
Symptomatic hypoglycemia (3.2 mmol/L)	1 (0)			
Pulmonal symptoms (dyspnoea, coughing)			3 (2)	
Skin symptoms (rash, urticaria)	1 (0)		1 (1)	
Weakness (heavy eyelids, drowsiness)	1 (0)		2 (1)	
Diarrhea the following day			1 (1)	
Back pain			1 (1)	
Presyncope after venous canulation	1 (0)			
Patients assessment				
Preference [A]	103 (72)		17 (12)	

Table 4 shows the occurrence of adverse events in all patients who underwent at least one stimulation test (n=164). Scores on the visual-analogue scale (VAS) range from 0 to 10, with 0 indicating no symptoms and 10 indicating the most severe symptoms imaginable.

Data presented as frequency (percentage) and median [interquartile range].

[A] Data on preference was available from 143 patients; of those 23 patients indicated no preference

Figure 1 Stimulated copeptin values after each test according to final diagnosis, differentiating between

A) AVP-deficiency (orange), primary polydipsia (blue)

B) complete AVP-deficiency (brown), partial AVP-deficiency (yellow), primary polydipsia (blue)

Y-axis is on log-scale for better visualization. The horizontal line in each box represents the median, the lower and upper boundaries of the boxes the interquartile range, the ends of the whisker lines the minimum and maximum values within 1.5 times the interquartile range, and the dots outliers.

SUPPLEMENTARY APPENDIX OF THE RESEARCH ARTICLE:

Arginine or Hypertonic Saline–Stimulated Copeptin to Diagnose AVP Deficiency

TABLE OF CONTENTS

List of Authors and Affiliations

Julie Refardt[1-3], Cihan Atila[1,2], Irina Chifu[4], Emanuele Ferrante[5], Zoran Erlic[6], Juliana B. Drummond[7], Rita Indirli[5,8], Roosmarijn C. Drexhage[3], Clara O. Sailer[1,2], Andrea Widmer[1,2], Susan Felder[1,2], Andrew S. Powlson[9], Nina Hutter[1,2], Deborah R. Vogt[2], Mark Gurnell[9,10], Beatriz S. Soares[7], Johannes Hofland[3], Felix Beuschlein[6,11], Martin Fassnacht[4,12], Bettina Winzeler[1,2], Mirjam Christ-Crain[1,2]

[1]Departments of Endocrinology, Diabetology and Metabolism University Hospital Basel; Basel, Switzerland.

[2]Department of Clinical Research, University Hospital Basel, University of Basel; Basel, Switzerland.

[3]Department of Internal Medicine, Section of Endocrinology, Erasmus Medical Center, Rotterdam, The Netherlands.

[4]Division of Endocrinology and Diabetes, Department of Internal Medicine I, University Hospital, University of Wuerzburg, Wuerzburg Germany.

[5]Endocrinology Unit, IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy.

[6]Department of Endocrinology, Diabetology, and Clinical Nutrition, University Hospital Zurich (USZ) and University of Zurich (UZH), Zürich Switzerland.

[7]Department of Internal Medicine, Medical School of the Federal University of Minas Gerais, Belo Horizonte, MG, Brazil.

[8]University of Milan, Department of Clinical Sciences and Community Health, Milan, Italy.

[9]Wellcome-MRC Institute of Metabolic Science, University of Cambridge & Addenbrooke's Hospital, Cambridge Biomedical Campus, Cambridge, UK.

[10]Cambridge NIHR Biomedical Research Centre, Cambridge University Hospitals, Cambridge, UK.

[11]Medizinische Klinik und Poliklinik IV, Klinikum der Universität, Ludwig-Maximilians-Universität München, Munich, Germany.

[12]The LOOP Zurich - Medical Research Center, Zurich, Switzerland.

[13]Central Laboratory, University Hospital Wuerzburg, Wuerzburg, Germany.

Details on Copeptin Assay Characteristics

BRAHMS Copeptin proAVP assay (Thermo Fisher Scientific, Hennigsdorf, Germany), is a CE certified automated immunoassay with a coefficient of variation (CV) within-laboratory precision of approximately 9.8%, and a CV for reproducibility of 7%.

Details on Sample Size Calculation

Sample size was estimated as 139 patients to achieve at least 80% power to show the non-inferiority of the overall diagnostic accuracy of the arginine stimulation test to the hypertonic saline stimulation test with a type I error rate of 5% as based on the two-sided 95% confidence interval on the difference. Based on our previous studies, we assumed an overall diagnostic accuracy of 93% for the arginine stimulation[1] and 96.5% for the hypertonic saline stimulation test[2]. Based on clinical relevance considerations we pre-specified a non-inferiority margin of -10%. Sample size range (N 20-220) was evaluated by simulating 999 independent samples each from a bivariate binomial distribution with probabilities 0.965 and 0.930 and correlation ρ=0.5. For each sample, the lower level of the 95% confidence interval for the difference in proportions was calculated based on Tango's asymptotic score[3,4] and compared against the non-inferiority margin. To address an assumed drop-out rate of 8%, the recruitment goal was 152 participants.

Details on Misclassified Patients

For arginine stimulation, 17 patients with AVP-deficiency (AVP-D) where misclassified as primary polydipsia patients (PP), with 13 having partial AVP-D (*Supplementary Table S4 and Figure S2*). Three patients suffered from severe nausea of which two vomited around the time of copeptin collection. No triggering factor was identified for the other patients. Twenty-three PP patients were wrongly diagnosed with AVP-D, of which three patients had a BMI above 30 kg/m^2.

For the seven wrongly categorized patients through hypertonic saline stimulation, six had a false-negative result (i.e. PP instead of AVP-D). Of these, all were classified as partial AVP-D

patients. One reported severe malaise at the end of the test and one was later diagnosed with transient AVP-D. No other explanatory factors were identified. The one patient with false-positive copeptin response (i.e. AVP-D instead of PP) had insufficient hypertonic saline stimulation (maximum sodium level 145 mmol/L) due to malaise during the test.

Details on Test Duration

While by protocol the arginine infusion test has a set start and end point (60 rsp. 90 minutes after start of infusion), the duration of the hypertonic saline test is variable. In our cohort, the median [IQR] duration of the hypertonic saline infusion was 90 minutes [60-90], after which patients had to stay for an additional hour until sodium levels normalized. Sodium re-normalization was reached in 87.3% of patients after this time. Twenty patients required further sodium re-lowering due to persistent hypernatremia of 147 mmol/L [146,148], prolonging the total test duration.

Supplementary Table S1: Demographic Information on the Broader Patient Population

	Global region	**Sample size**	**Age** Years, median [IQR] or mean (SD)		**Sex / Gender** Female (%)		**Race / Ethnicity** Caucasian (%)	
			AVP-Deficiency	**Primary Polydipsia**	**AVP-Deficiency**	**Primary Polydipsia**	**AVP-Deficiency**	**Primary Polydipsia**
Fenske W., Refardt J. et al.[2]	Europe South America	141	45 [33-53]	32 [24-44]	64 %	67 %	93 %	96 %
Winzeler B. et al.[1]	Europe	98	40 (12) 51 (15)	35 (14) 34 (10)	58 %	72 %	97 %	98 %
de Leon J. et al.[5]	North America	150	NA	42 (13)	NA	48 %	NA	84 %
de Leon J. et al.[6]	North America	61	NA	42 (12)	NA	33 %	NA	87 %
Timper K. et al.[7]	Europe	55	44 [38-58] 45 [29-56]	36 [26-46]	89 %	72 %	NA	NA
Fenske W. et al.[8]	Europe	50	41 (14) 39 (11)	35 (12)	NA	NA	NA	NA
Maghnie M. et al.[9]	Europe	79	7 [0.7 to 25]	NA	53 %	NA	NA	NA
Atila C. et al.[10]	Web-based international	1034	42 [31-53]	NA	77 %	NA	NA	NA
Winzeler B., Sailer CO. et al.[11]	Europe	34	NA	30 [26-39]		68 %	NA	NA
Hadjizacharia P. et al.[12]	North America	60	37 (20)	NA	22 %	NA	NA	NA
Dilrukshi M. et al.[13]	Europe	109	42 [24-60]	NA	54 %	NA	NA	NA
Sjöström A. et al.[14]	Europe	153	48 [12-81]	50 [17-79]	57 %	66 %	NA	NA
Hawken E. et al.[15]	North America	48	NA	40 (10) 33 (12) 24 (10)	NA	25 %	NA	NA
Iraqi HM. et al.[16]	Europe Asia	92	35 (21)	NA	64 %	NA	NA	NA
Iraqi HM. et al.[17]	Europe Asia	70	47 (15)	NA	57 %	NA	NA	NA
Meta-analyses & systemic review articles Angelousi A.[18] et al. & Mu D[19]. et al. not listed.								

Supplementary Table S1: Demographic information on the Broader Patient Population

Studies were identified according to a PubMed search from the inception of the database to July 15, 2023, for articles published in English using the terms “diabetes insipidus,” “arginine vasopressin deficiency”, “primary polydipsia”, “demographics”, “cohort study” and “epidemiology” and all studies with well-described samples were selected. In total, 17 studies - including seven prospective studies and one meta-analysis - were identified and the results summarized in this table.
Data from international studies are lacking. No signals on differences between races or ethnicities. Manifestation at any age possible according to underlying etiology. For AVP deficiency, no major sex or gender differences. Available studies included slightly higher numbers of female patients. For Primary Polydipsia, tendency of more females than males affected for the general population. For psychiatric patients balanced between both sexes.

AVP=arginine vasopressin. NA = information not available

Supplementary Table S2

Medication, n (%)				
Desmopressin	20 (49)	11 (39)	31 (45)	2 (2)
Hydrocortisone	14 (34)	13 (46)	27 (39)	7 (8)
Levothyroxine	16 (39)	10 (23)	26 (38)	9 (10)
Testosterone	5 (12)	4 (10)	9 (13)	2 (2)
Hormonal contraceptive / HRT	3 (7)	2 (7)	5 (7)	6 (7)
Growth hormone	2 (5)	1 (4)	3 (4)	0 (0)
Antidepressants	4 (10)	4 (10)	8 (12)	16 (18)
other medication	27 (66)	21 (75)	48 (70)	60 (67)

Table S2 Medications of all included patients who underwent both tests and received a final diagnosis (modified intention-to-treat-set 1).
Data presented as number (n) and frequency (percentage).

Supplementary Table S3

	Arginine Stimulation First N=78	**Hypertonic Saline Infusion First** N=80
Characteristics		
Age, years	38 [31, 50]	42 [32, 53]
Female sex, n (%)	50 (64)	56 (70)
Body mass index, kg/m^2	25.6 [22.0, 29.4]	25.8 [21.6, 30.0]
Clinical symptoms at time of diagnosis		
Polydipsia, liters/day	6.0 [4.6, 8.0]	5.4 [4.0, 8.0]
Polyuria, liters/day	5.7 [4.0, 7.9]	5.1 [4.0, 7.8]
Emictions per day	12 [9, 18]	10 [8, 15]
Nocturia, n (%)	58 (74)	66 (83)
Nocturia, times/night	3 [2, 4]	3 [2, 5]
Drinking at night, n (%)	50 (64)	61 (76)
Drinking at night, liters/night	1.0 [0.5, 1.5]	1.0 [0.5, 1.5]
Medical history, n (%)		
Anterior pituitary insufficiency	19 (24)	15 (19)
Adrenocorticotropic hormone	13 (17)	13 (16)
Thyrotropic hormone	15 (19)	15 (19)
Growth hormone	4 (5)	5 (6)
Gonadotropins	13 (17)	13 (16)
Pituitary lesions		
History of pituitary surgery	15 (19)	13 (16)
History of pituitary apoplexy	1 (1)	1 (1)
Psychiatric disorder	15 (19)	17 (21)
Cardiovascular disease	3 (4)	7 (9)
Cerebrovascular disease	2 (3)	4 (5)
Other	52 (67)	49 (61)
AVP-D etiology, n (%)		
Post-surgical AVP-D	11 (14)	10 (13)
Adenoma	5 (6)	6 (46)
Craniopharyngioma	4 (5)	1 (1)
Meningioma	1 (1)	3 (4)
Other	1 (1)	0 (0)
Hypothalamic-pituitary lesions	10 (13)	16 (20)
Adenoma	6 (8)	6 (8)
Rathke cleft cyst	2 (3)	4 (5)
Germinoma	1 (1)	0 (0)
Meningioma	0 (0)	1 (1)

Sarcoidosis	0 (0)	1 (1)
Other	1 (0)	4 (5)
Trauma	3 (4)	2 (3)
Empty sella or hypoplasia	1 (1)	5 (6)
Vascular (e.g., apoplexy, Sheehan syndrome)	1 (1)	1 (1)
Hypophysitis	3 (4)	5 (6)
Idiopathic AVP-D	2 (3)	6 (8)
Familial AVP-D	0 (0)	3 (4)
Medication, n (%)		
Desmopressin	16 (21)	17 (21)
Hydrocortisone	18 (23)	16 (20)
Levothyroxine	17 (22)	18 (23)
Testosterone	9 (12)	2 (3)
Hormonal contraceptive / HRT	5 (6)	6 (8)
Growth hormone	1 (1)	2 (3)
Antidepressants	13 (17)	11 (14)
other medication	53 (68)	55 (69)
Laboratory data at baseline		
Serum sodium, mmol/L	140 [139, 142]	141 [139, 143]
Serum osmolality, mOsm/kg	290 [283, 293]	291 [287, 294]
Serum Copeptin, pmol/L	2.6 [2.0, 3.6]	2.2 [1.7, 2.9]
Urine osmolality, mOsm/kg	214 [135, 430]	192 [129, 298]
MRI characteristics, n (%)		
MRI performed	50 (64)	58 (72)
Pituitary stalk enlarged	5 (10)	10 (17)
Bright spot absent	23 (46)	26 (45)
Enlargement of the posterior pituitary	7 (14)	5 (9)
Allusion to adenohypophysis / hypophysitis	6 (12)	4 (7)
Other findings (e.g., adenoma)	16 (32)	25 (43)

Table S3 Characteristics of all included patients who underwent both tests and received a final diagnosis according to randomization order.
Data presented as frequency (percentage) and median [IQR].
n=number, IQR=interquartile range, AVP-Deficiency=arginine vasopressin deficiency, HRT=hormonal replacement therapy, MRI = magnetic resonance imaging

Supplementary Table S4

Arginine stimulation	**Outcome +**	**Outcome -**	**Total**
Test +	52	23	75
Test -	17	64	81
Total	69	87	156
Hypertonic saline stimulation	**Outcome +**	**Outcome -**	**Total**
Test +	63	1	64
Test -	6	88	94
Total	69	89	158

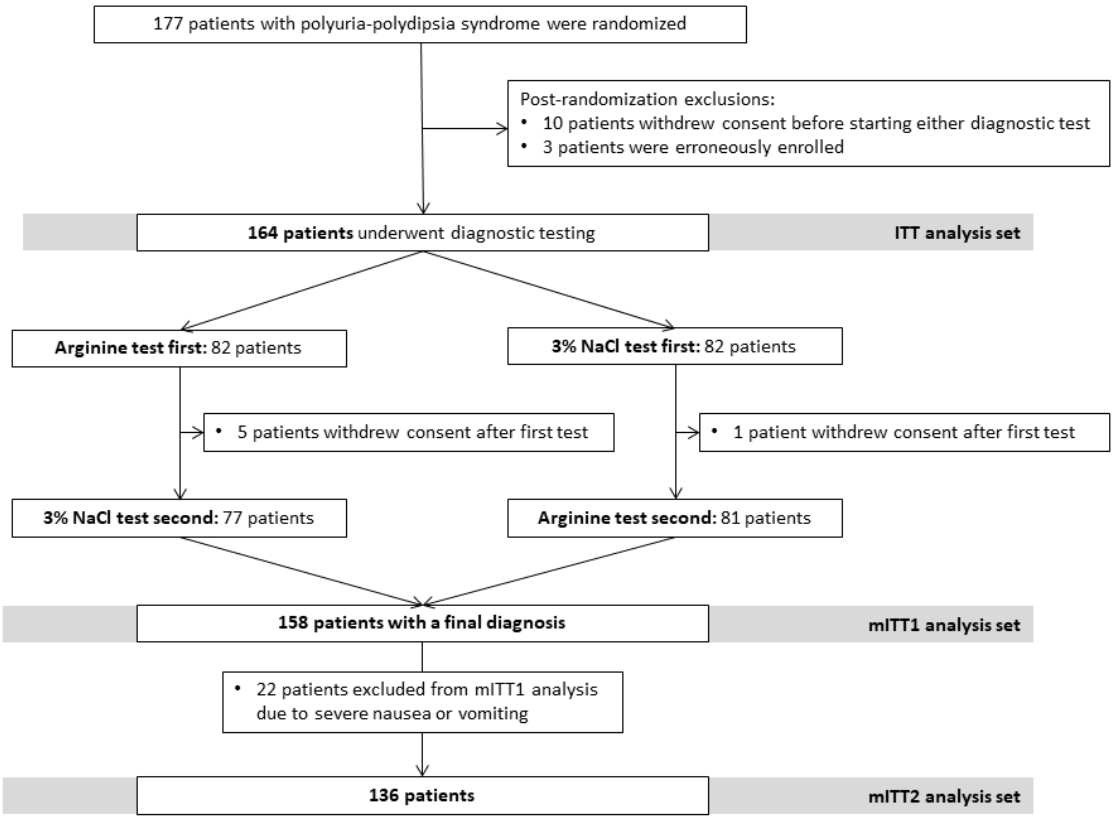
Supplementary Table S4: Diagnostic table for arginine and hypertonic saline stimulation.
Positives (+) refer to AVP-deficiency, *negatives (-)* refer to primary polydipsia;
Outcome corresponds to the final expert diagnosis;
Test corresponds to the test result of the according test.

Supplementary Table S5

Test	**cut-off**	**Accuracy**	**Sensitivity**	**Specificity**	**PPV**	**NPV**
Arginine stimulation	3.7 pmol/L, **60 minutes**	75.0 [67.7, 81.1]	73.9 [62.5, 82.8]	75.9 [65.9, 83.6]	70.8 [59.5, 80.1]	78.6 [68.7, 86.0]
Arginine stimulation	4.1 pmol/L, **90 minutes**	79.2 [72.1, 84.9]	79.1 [67.9, 87.1]	79.3 [69.6, 86.5]	74.6 [63.4, 83.3]	83.1 [73.7, 89.7]
Hypertonic saline stimulation	6.5 pmol/L	96.2 [92.0, 98.2]	95.7 [88.0, 98.5]	96.6 [90.6, 98.8]	95.7 [88.0, 98.5]	96.6 [90.6, 98.8]

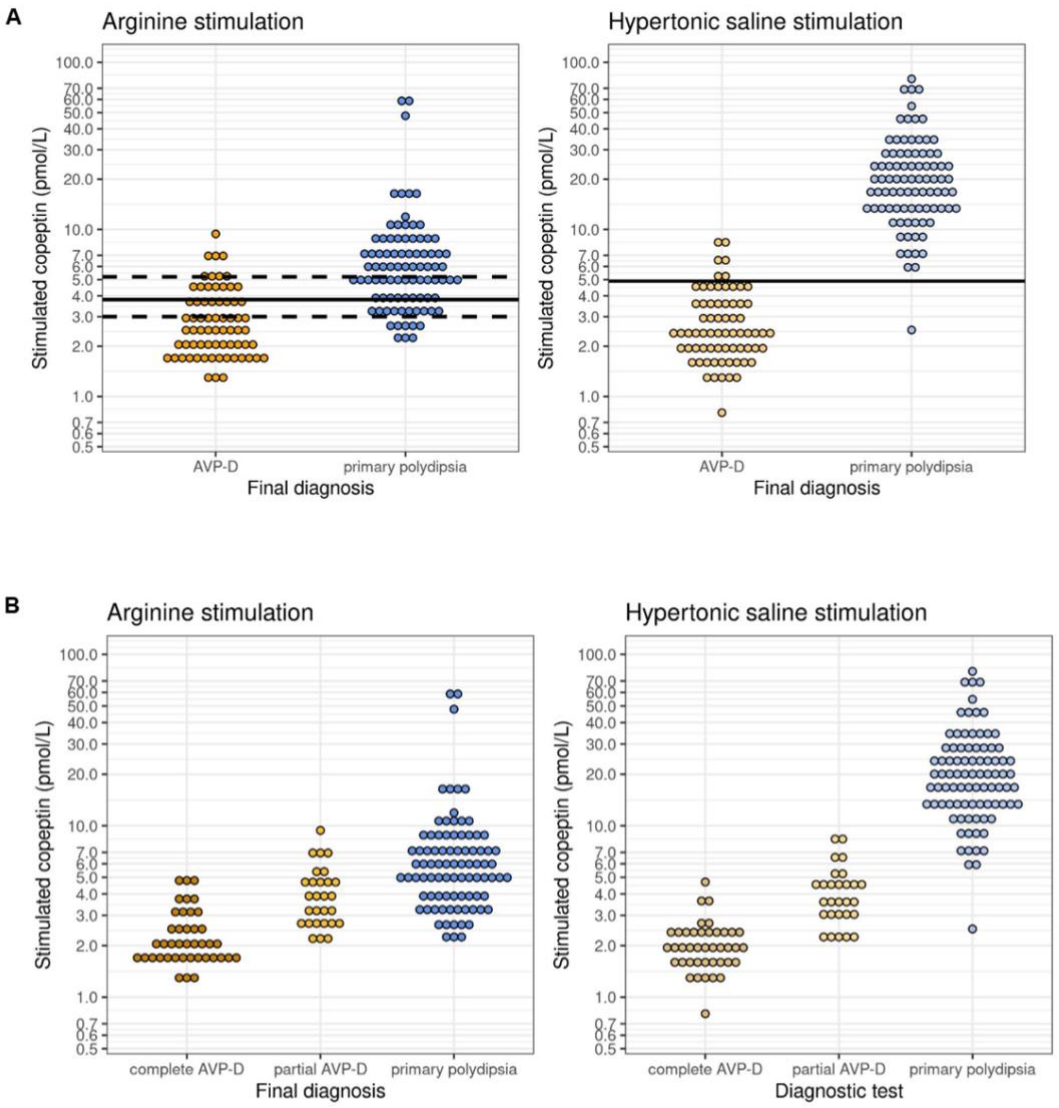
Supplementary Table S5: Diagnostic measures (% with 95% confidence interval) of different copeptin cut-offs.
PPV=positive predictive value, NPV=negative predictive value

Supplementary Figure S1



Supplementary Figure S1 Patient Flow Diagram
ITT=intention to treat, mITT=modified intention to treat

Supplementary Figure S2



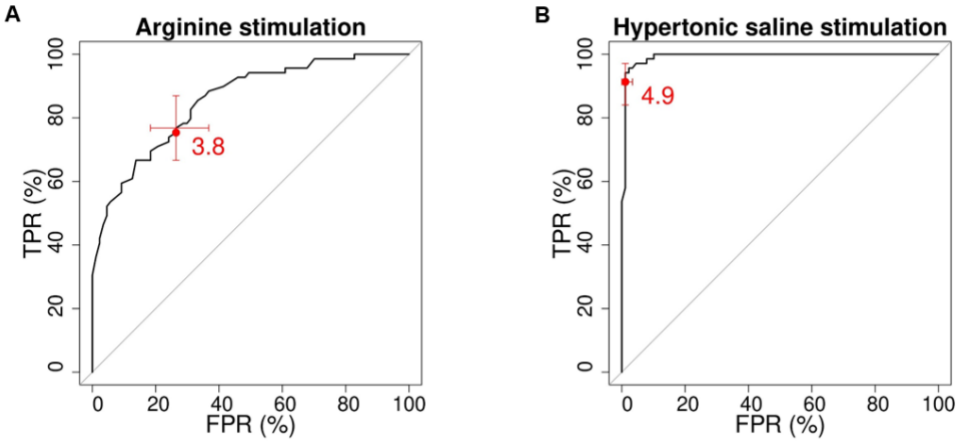
Supplementary Figure S2: Stimulated Copeptin Values after Each Test According to Final Diagnosis

This figure shows values differentiating between

A) AVP-deficiency (orange) and primary polydipsia (blue)
B) Complete AVP-deficiency (red), partial AVP-deficiency (yellow) and primary polydipsia (blue)

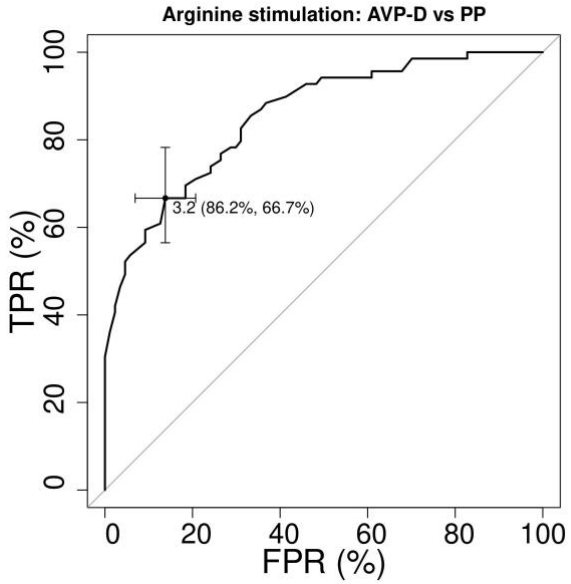
Horizontal black lines represent the cut-off for predefined copeptin values (solid) and newly calculated copeptin values (dashed). Each dot represents one patient. Y-axis is on log-scale for better visualization.

Supplementary Figure S3



Supplementary Figure S3: Receiver-Operator Characteristics (ROC) Curve for Arginine Stimulated Copeptin (A) and Hypertonic Saline Stimulated Copeptin (B).
Point estimates denote the cut-offs (red); error bars show the 95% confidence interval for TPR and FPR.
TPR=true positive rate=sensitivity; FPR=false positive rate=100-specificity.

Supplementary Figure S4



Supplementary Figure S4: Receiver-operator characteristics (ROC) curve for arginine stimulation to discriminate between patients with AVP-deficiency (AVP-D) from primary polydipsia (PP). The 'best' cut-off, maximizing the combination of sensitivity and specificity is indicated.

Point estimate denotes the cut-off; numbers in brackets are (sensitivity, specificity); error bars show the 95% confidence interval for TPR and FPR.

TPR=true positive rate=sensitivity; FPR=false positive rate=100-specificity.

References

1. Winzeler B, Cesana-Nigro N, Refardt J, et al. Arginine-stimulated copeptin measurements in the differential diagnosis of diabetes insipidus: a prospective diagnostic study. Lancet 2019;394:587-95.
2. Fenske W, Refardt J, Chifu I, et al. A Copeptin-Based Approach in the Diagnosis of Diabetes Insipidus. N Engl J Med 2018;379:428-39.
3. Tango T. Equivalence test and confidence interval for the difference in proportions for the paired-sample design. Stat Med 1998;17:891-908.
4. Fagerland MW, Lydersen S, Laake P. Recommended tests and confidence intervals for paired binomial proportions. Stat Med 2014;33:2850-75.
5. de Leon J, Dadvand M, Canuso C, Odom-White A, Stanilla J, Simpson GM. Polydipsia and water intoxication in a long-term psychiatric hospital. Biol Psychiatry 1996;40:28-34.
6. de Leon J. Polydipsia--a study in a long-term psychiatric unit. Eur Arch Psychiatry Clin Neurosci 2003;253:37-9.
7. Timper K, Fenske W, Kühn F, et al. Diagnostic Accuracy of Copeptin in the Differential Diagnosis of the Polyuria-polydipsia Syndrome: A Prospective Multicenter Study. J Clin Endocrinol Metab 2015;100:2268-74.
8. Fenske W, Quinkler M, Lorenz D, et al. Copeptin in the differential diagnosis of the polydipsia-polyuria syndrome--revisiting the direct and indirect water deprivation tests. J Clin Endocrinol Metab 2011;96:1506-15.
9. Maghnie M, Cosi G, Genovese E, et al. Central diabetes insipidus in children and young adults. N Engl J Med 2000;343:998-1007.
10. Atila C, Loughrey PB, Garrahy A, et al. Central diabetes insipidus from a patient's perspective: management, psychological co-morbidities, and renaming of the condition: results from an international web-based survey. Lancet Diabetes Endocrinol 2022.
11. Winzeler B, Sailer CO, Coynel D, et al. A randomized controlled trial of the GLP-1 receptor agonist dulaglutide in primary polydipsia. J Clin Invest 2021;131.
12. Hadjizacharia P, Beale EO, Inaba K, Chan LS, Demetriades D. Acute diabetes insipidus in severe head injury: a prospective study. J Am Coll Surg 2008;207:477-84.
13. Dilrukshi M, Vickars M, May C, et al. Management of cranial Diabetes Insipidus – clinical outcomes and patient perception of care. European Journal of Endocrinology 2022:EJE-22-0187.
14. Sjöström A, Bartuseviciene I, Höybye C. Simplified and improved fluid deprivation test for diagnosing diabetes insipidus. Eur J Endocrinol 2021;184:123-31.
15. Hawken ER, Crookall JM, Reddick D, Millson RC, Milev R, Delva N. Mortality over a 20-year period in patients with primary polydipsia associated with schizophrenia: a retrospective study. Schizophr Res 2009;107:128-33.
16. Iraqi HM, Pigarova E, Zacharieva S, et al. Idiopathic central diabetes insipidus in a large cohort of patients: the hypopituitarism ENEA rare observational (HEROS) study. Pituitary 2022;25:982-7.
17. Masri-Iraqi H, Hirsch D, Herzberg D, et al. CENTRAL DIABETES INSIPIDUS: CLINICAL CHARACTERISTICS AND LONG-TERM COURSE IN A LARGE COHORT OF ADULTS. Endocr Pract 2017;23:600-4.
18. Angelousi A, Mytareli C, Xekouki P, et al. Diabetes insipidus secondary to sellar/parasellar lesions. J Neuroendocrinol 2021;33:e12954.
19. Mu D, Ma Y, Cheng J, Qiu L, Chen S, Cheng X. Diagnostic Accuracy of Copeptin in the Differential Diagnosis of Patients With Diabetes Insipidus: A Systematic Review and Meta-analysis. Endocr Pract 2023.