ORIGINAL ARTICLE

Arginine or Hypertonic Saline–Stimulated Copeptin to Diagnose AVP Deficiency

Julie Refardt, M.D., Ph.D., Cihan Atila, M.D., Ph.D., Irina Chifu, M.D., Emanuele Ferrante, M.D., Zoran Erlic, M.D., Juliana B. Drummond, M.D., Ph.D., Rita Indirli, M.D., Roosmarijn C. Drexhage, M.D., Ph.D., Clara O. Sailer, M.D., Ph.D., Andrea Widmer, M.D., Susan Felder, M.D., Andrew S. Powlson, M.D., Nina Hutter, Deborah R. Vogt, Ph.D., Mark Gurnell, M.D., Ph.D., Beatriz S. Soares, M.D., Ph.D., Johannes Hofland, M.D., Ph.D., Felix Beuschlein, M.D., Martin Fassnacht, M.D., Bettina Winzeler, M.D., and Mirjam Christ-Crain, M.D., Ph.D.

ABSTRACT

BACKGROUND

Distinguishing between arginine vasopressin (AVP) deficiency and primary polydipsia is challenging. Hypertonic saline–stimulated copeptin has been used to diagnose AVP deficiency with high accuracy but requires close sodium monitoring. Arginine-stimulated copeptin has shown similar diagnostic accuracy but with a simpler test protocol. However, data are lacking from a head-to-head comparison between arginine-stimulated copeptin and hypertonic saline–stimulated copeptin in the diagnosis of AVP deficiency.

METHODS

In this international, noninferiority trial, we assigned adult patients with polydipsia and hypotonic polyuria or a known diagnosis of AVP deficiency to undergo diagnostic evaluation with hypertonic-saline stimulation on one day and with arginine stimulation on another day. Two endocrinologists independently made the final diagnosis of AVP deficiency or primary polydipsia with use of clinical information, treatment response, and the hypertonic-saline test results. The primary outcome was the overall diagnostic accuracy according to prespecified copeptin cutoff values of 3.8 pmol per liter after 60 minutes for arginine and 4.9 pmol per liter once the sodium level was more than 149 mmol per liter for hypertonic saline.

RESULTS

Of the 158 patients who underwent the two tests, 69 (44%) received the diagnosis of AVP deficiency and 89 (56%) received the diagnosis of primary polydipsia. The diagnostic accuracy was 74.4% (95% confidence interval [CI], 67.0 to 80.6) for arginine-stimulated copeptin and 95.6% (95% CI, 91.1 to 97.8) for hypertonic saline–stimulated copeptin (estimated difference, –21.2 percentage points; 95% CI, –28.7 to –14.3). Adverse events were generally mild with the two tests. A total of 72% of the patients preferred testing with arginine as compared with hypertonic saline. Arginine-stimulated copeptin at a value of 3.0 pmol per liter or less led to a diagnosis of AVP deficiency with a specificity of 90.9% (95% CI, 81.7 to 95.7), whereas levels of more than 5.2 pmol per liter led to a diagnosis of primary polydipsia with a specificity of 91.4% (95% CI, 83.7 to 95.6).

CONCLUSIONS

Among adult patients with polyuria polydipsia syndrome, AVP deficiency was more accurately diagnosed with hypertonic saline–stimulated copeptin than with arginine-stimulated copeptin. (Funded by the Swiss National Science Foundation; CARGOx ClinicalTrials.gov number, NCT03572166.)

The authors' affiliations are listed in the Appendix. Dr. Refardt can be contacted at julie.refardt@usb.ch or at the Department of Endocrinology, University Hospital Basel, Petersgraben 4, 4031 Basel, Switzerland.

This is the *New England Journal of Medicine* version of record, which includes all *Journal* editing and enhancements. The Author Accepted Manuscript, which is the author's version after external peer review and before publication in the *Journal*, is available at PubMed Central.

N Engl J Med 2023;389:1877-87.
DOI: 10.1056/NEJMoa2306263

CME at NEJM.org

A Quick Take is available at NEJM.org

Arginine vasopressin (AVP) deficiency (formerly known as central diabetes insipidus) and AVP resistance (formerly known as nephrogenic diabetes insipidus) must be differentiated from primary polydipsia, which is defined as excessive fluid intake despite initial adequate AVP secretion and renal response.[1,2] AVP deficiency is characterized by inadequate release of AVP, whereas AVP resistance results from renal insensitivity to AVP.[3,4] Complete and partial dysfunction have been described in both forms.[5] The differentiation of these conditions from primary polydipsia is critical, because treatments of the three conditions differ and potential misdiagnosis carries the risk of severe complications.[1]

Although the indirect water deprivation test was once considered to be 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 (a polypeptide that is released as part of normal AVP secretion) as a stable and quick osmosensitive surrogate marker for AVP,[8,9] the direct-test approach was rediscovered.[10] Whereas unstimulated copeptin levels (with a cutoff of >21.4 pmol per liter) can be used to diagnose AVP resistance,[11,12] a stimulated copeptin is required to differentiate between AVP deficiency and primary polydipsia. In a large multicenter trial,[6] hypertonic saline–stimulated copeptin was used to diagnose AVP deficiency with high accuracy (96.5%). A downside of that approach is the need for frequent sodium monitoring to avoid overstimulation and patient discomfort from induced hypernatremia.

An alternative test that uses arginine-stimulated copeptin has shown diagnostic accuracy (93%) similar to that of hypertonic saline with a simpler test protocol and an acceptable side-effect profile.[13] On the basis of these results, the use of arginine-stimulated copeptin would seem to be preferable to the use of hypertonic saline–stimulated copeptin as a standard test to differentiate between AVP deficiency and primary polydipsia, but data from a prospective head-to-head comparison are lacking.

In an international, noninferiority trial, we evaluated the diagnostic accuracy of hypertonic saline as compared with arginine in differentiating between AVP deficiency and primary polydipsia in adult patients with polyuria polydipsia syndrome. We hypothesized that the diagnostic accuracy of arginine-stimulated copeptin would be noninferior to the accuracy of hypertonic saline–stimulated copeptin.

Methods

Trial Design and Patients

We conducted the Use of Copeptin Measurement after Arginine Infusion for the Differential Diagnosis of Diabetes Insipidus (CARGOx) trial at seven tertiary medical centers in Europe and Brazil from September 2018 through September 2022, with a 3-month follow-up that was completed in December 2022. The local ethics committee at each center approved the protocol. Written informed consent was obtained from all the patients before any trial procedure.

We recruited adult patients (≥18 years of age) with polydipsia (self-reported consumption of >3 liters of fluid per day) and hypotonic polyuria (>50 ml of urine per kilogram of body weight in a 24-hour collection and urine osmolality of <800 mOsm per kilogram) or patients with a known diagnosis of AVP deficiency. Patients with AVP resistance or polyuria–polydipsia that was associated with other causes (diabetes mellitus, hypercalcemia, or hypokalemia) were excluded from the trial. Additional exclusion criteria were treatment for epilepsy, uncontrolled arterial hypertension, heart failure, liver cirrhosis, uncorrected adrenal or thyroid hormone deficiency, pregnancy or breast-feeding, or any relevant acute or terminal illness. Additional details regarding inclusion and exclusion criteria are provided in the protocol, available with the full text of this article at NEJM.org.

Procedures

Baseline Assessment

After obtaining a detailed medical history, we performed a standardized clinical and biochemical evaluation. Pituitary magnetic resonance imaging (MRI) was recommended for all trial patients. MRIs were assessed for general alterations of the pituitary (i.e., pituitary lesions or postoperative changes) and typical characteristics of AVP deficiency.[14,15]

The patients were randomly assigned to undergo first testing with either arginine or hypertonic-saline stimulation on two different days. Tests were performed in the morning after an overnight fast, with fluid intake allowed until

6 a.m. Desmopressin treatment was paused 24 hours before the tests or for a minimum of 12 hours in patients with severely symptomatic AVP deficiency. Treatment was restarted after completion of the assigned test. Patients who were receiving hydrocortisone received an individualized stress dose.

Arginine-Stimulation Test

An infusion of l-arginine hydrochloride (21%) at a dose of 0.5 g per kilogram of body weight (maximum, 40 g) diluted in 500 ml of normal saline (sodium chloride, 0.9%) was administered over a 30-minute period. Blood samples for copeptin measurement were obtained before the infusion and 60 and 90 minutes after the start of the infusion.

The diagnosis was made at the end of the trial according to prespecified copeptin cutoff values at 60 minutes.[13] A stimulated copeptin level of less than 2.4 pmol per liter indicated complete AVP deficiency, a level of 2.4 to 3.8 pmol per liter indicated partial AVP deficiency, and a level of more than 3.8 pmol per liter indicated primary polydipsia.

Hypertonic Saline–Stimulation Test

The patients received appropriate venous access in both arms, one for the infusion and one for blood sampling. After the infusion of a 250-ml bolus of hypertonic saline (sodium chloride, 3%), the infusion was continued at a rate of 0.15 ml per kilogram of body weight per minute. Blood samples were drawn every 30 minutes. Sodium levels were monitored by rapid venous blood gas analysis. The infusion was stopped once the sodium level in the blood gas analysis reached more than 149 mmol per liter, followed by immediate copeptin measurement. Once sampling was completed, patients received an oral water load (30 ml per kilogram of body weight) and a 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 prespecified cutoff levels for copeptin,[6] which was measured at a sodium level more than 149 mmol per liter. A stimulated copeptin level of less than 2.7 pmol per liter indicated complete AVP deficiency, a level of 2.7 to 4.9 pmol per liter indicated partial AVP deficiency, and a level of more than 4.9 pmol per liter indicated primary polydipsia.

Assessment of Test Burden and Adverse Events

We assessed the test burden and prespecified clinical symptoms (i.e., thirst, vertigo, headache, nausea, and malaise) in all the patients. We then rated the test burden and symptoms using a visual analogue scale (VAS), which ranged from 0 indicating no sensation or burden to 10 indicating a maximum sensation or burden.

Preliminary Diagnosis and Assessment of Treatment Response

After completing both tests, patients were discharged with a provisional diagnosis and treatment. At the 3-month follow-up visit, treatment response and clinical outcome were assessed and the preliminary diagnosis was reevaluated.

FINAL DIAGNOSIS

The final diagnosis was independently made by two endocrinologists after consideration of the patient's medical history and clinical symptoms, laboratory and imaging data, results of the hypertonic saline–stimulation test, and the therapeutic response at the 3-month follow-up. The experts were unaware of the results of the arginine-stimulation test, and their diagnoses were not bound to the results of the hypertonic saline–stimulation test. In the event of a discordant assessment, a third endocrine expert was consulted.

After the patients had been classified as having AVP deficiency or primary polydipsia, the distinction between partial or complete AVP deficiency was made mainly according to the prespecified cutoff values for copeptin on the hypertonic saline–stimulation test,[6] but the diagnosis could be overruled on the basis of clinical information.

LABORATORY MEASUREMENTS

Laboratory measurements were performed by automated biochemical analyses in trial center laboratories. Serum sodium levels were analyzed by means of indirect ion selective electrodes, and venous blood gas analysis was performed by means of direct ion selective electrodes.

Hypertonic saline–stimulated copeptin levels were measured immediately after a completed test by the investigator at each trial center, and arginine-stimulated copeptin levels were measured centrally at the end of the trial. All copeptin measurements were performed with the use of the BRAHMS Copeptin proAVP assay (Thermo Fisher Scientific). Details regarding all labora-

Table 1. Characteristics of the Patients.*

Characteristic	AVP Deficiency (N=69)			Primary Polydipsia (N=89)
	Complete (N=41)	Partial (N=28)	Complete and Partial (N=69)	
Demographic				
Median age (IQR) — yr	38 (31–47)	50 (39–58)	42 (32–54)	37 (28–50)
Female sex — no. (%)	24 (59)	14 (50)	38 (55)	68 (76)
Median body-mass index (IQR)†	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				
Median polydipsia (IQR) — 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)
Median polyuria (IQR) — 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)
Median emictions (IQR) — no./day	15 (8–20)	11 (8–14)	12 (8–15)	10 (9–15)
Nocturia				
Patients with condition — no. (%)	32 (78)	24 (86)	56 (81)	68 (76)
Median no. of times/night (IQR)	4 (3–5)	3 (2–4)	4 (3–5)	3 (2–3)
Fluid intake at night				
Patients with condition — no. (%)	32 (78)	19 (68)	51 (74)	60 (67)
Median no. of liters/night (IQR)	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 — no. (%)				
Anterior pituitary insufficiency	16 (39)	13 (46)	29 (42)	5 (6)
Corticotropin	14 (34)	11 (39)	24 (35)	2 (2)
Thyrotropic hormone	14 (34)	13 (46)	27 (39)	3 (3)
Growth hormone	4 (10)	4 (14)	8 (12)	1 (1)
Gonadotropins	13 (32)	10 (36)	23 (33)	3 (3)
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	1 (4)	1 (1)	1 (1)
Psychiatric disorder	3 (7)	5 (18)	8 (12)	24 (27)
Cardiovascular disease	3 (7)	3 (11)	6 (9)	4 (4)
Cerebrovascular disease	3 (7)	2 (7)	5 (7)	1 (1)
Other	28 (68)	20 (71)	48 (70)	53 (60)
Cause of AVP deficiency — no. (%)				
Postsurgical condition	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‡	0	1 (41)	1 (1)	NA
Hypophysitis	4 (10)	4 (14)	8 (12)	NA
Idiopathic	7 (17)	1 (4)	8 (12)	NA
Familial	3 (7)	0	3 (4)	NA
Laboratory data				
Median serum sodium (IQR) — mmol/liter	142 (140–143)	142 (140–143)	142 (140–143)	140 (138–141)
Median serum osmolality (IQR) — mOsm/kg	293 (289–296)	292 (290–295)	293 (290–296)	287 (283–291)
Median serum copeptin (IQR) — pmol/liter	1.8 (1.4–2.1)	2.7 (2.3–3.4)	2.2 (1.6–2.4)	2.6 (2.0–3.9)
Median urine osmolality (IQR) — mOsm/kg	137 (90–216)	230 (168–312)	181 (108–299)	222 (156–431)
Abnormal finding on MRI — no./total no. (%)	39/41 (95)	25/28 (89)	64/69 (93)	44/89 (49)

Table 1. (Continued.)

Characteristic	AVP Deficiency (N=69)			Primary Polydipsia (N=89)
	Complete (N=41)	Partial (N=28)	Complete and Partial (N=69)	
Pituitary stalk enlarged	8/39 (20)	5/25 (20)	13/64 (20)	2/44 (5)
Bright spot absent	27/39 (69)	16/25 (64)	43/64 (68)	6/44 (14)
Enlargement of posterior pituitary	7/39 (18)	3/25 (12)	10/64 (16)	2/44 (5)
Alterations to adenohypophysis or hypophysitis	4/39 (10)	6/25 (24)	10/64 (16)	0
Other findings	12/39 (31)	15/25 (60)	27/64 (42)	14/44 (32)

* Listed are the characteristics of all the trial patients who underwent diagnostic evaluation with both hypertonic-saline and arginine stimulation and who received a final diagnosis (modified intention-to-treat population 1). AVP denotes arginine vasopressin, IQR interquartile range, MRI magnetic resonance imaging, and NA not applicable.
† The body-mass index is the weight in kilograms divided by the square of the height in meters.
‡ Vascular causes included apoplexy and Sheehan's syndrome.

tory measures are provided in the Supplementary Appendix, available at NEJM.org.

OUTCOMES

The primary outcome was the overall diagnostic accuracy of the two tests in differentiating AVP deficiency from primary polydipsia on the basis of the prespecified cutoffs for copeptin. Overall diagnostic accuracy was calculated as the ratio of correct diagnoses (true positive values plus true negative values) to all the final diagnoses.

Secondary outcomes were an acceptable side-effect profile for the test, patients' preference between the two tests, and the diagnostic performance of previously derived and prespecified stimulated copeptin cutoffs. These cutoffs were evaluated as follows: first, in the analysis of AVP deficiency versus primary polydipsia, we used a cutoff value of 3.7 pmol per liter after 60 minutes and 4.1 pmol per liter after 90 minutes for arginine stimulation[13] and 6.5 pmol per liter for hypertonic-saline stimulation[6]; second, in the analysis of complete versus partial AVP deficiency, we used a cutoff of 2.4 pmol per liter after 60 minutes and 2.6 pmol per liter after 90 minutes for arginine stimulation[13] and 2.7 pmol per liter for hypertonic-saline stimulation.[6]

OVERSIGHT

The trial was funded by the Swiss National Science Foundation, which had no role in the design and conduct of the trial; in the collection, management, analysis, or interpretation of the data; or in the preparation, review, or approval of the manuscript. The first draft of the manuscript was written by the first author; all the authors submitted revisions and made the decision to submit the manuscript for publication. The authors vouch for the completeness and accuracy of the data and for the fidelity of the trial to the protocol. Confidentiality agreements regarding the data existed between the sponsor and all the authors until publication.

STATISTICAL ANALYSIS

We estimated that the enrollment of 139 patients would provide the trial with a power of at least 80% with a two-sided 5% type I error to show the noninferiority (at a margin of 10 percentage points) 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 withdrawal of 8%, we set the recruitment goal at 152 patients. Details regarding the sample estimation are provided in the Supplementary Appendix.

We calculated the overall diagnostic accuracy, sensitivity, specificity, and positive and negative predictive values with 95% Wilson confidence intervals for each examined copeptin cutoff value and estimated the area under the receiver-operator-characteristic curve with the bootstrap 95% confidence interval for each test procedure. For the testing of noninferiority, we calculated the difference in the diagnostic accuracy between arginine and hypertonic-saline stimulation with a 95% confidence interval by applying Tango's method for matched pairs.[16] We explored the

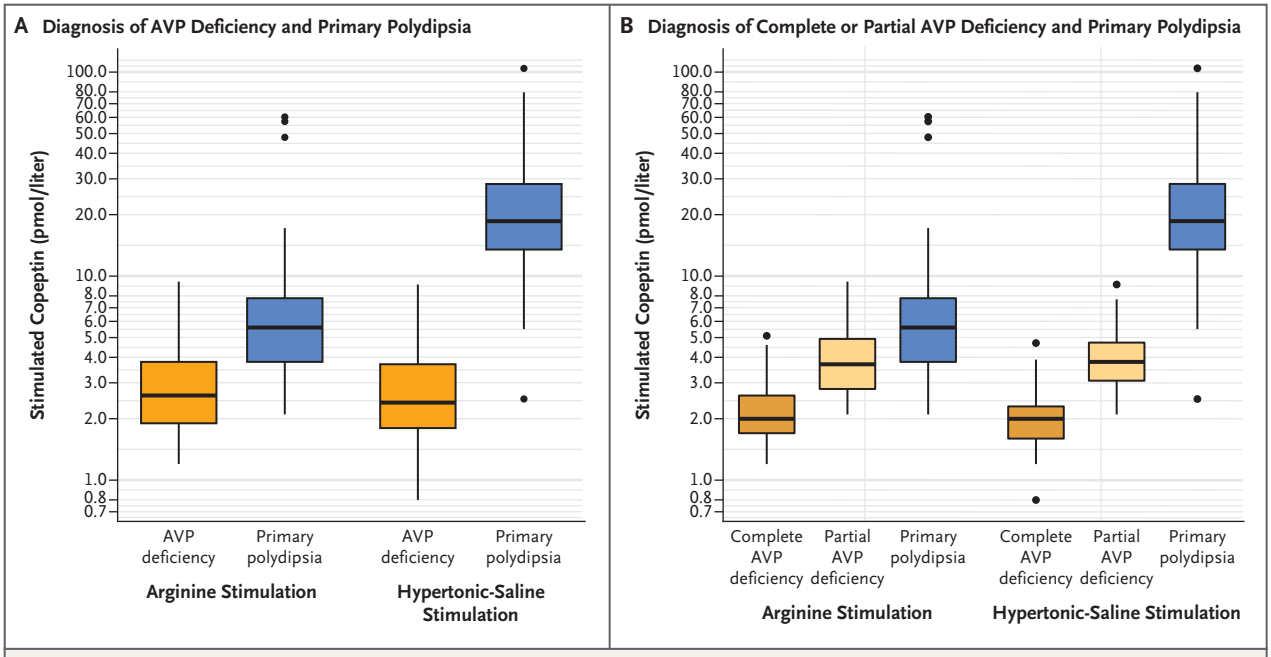

Figure 1. Copeptin Values after Arginine and Hypertonic-Saline Stimulation.

Shown is the diagnosis of arginine vasopressin (AVP) deficiency as compared with primary polydipsia (Panel A) and the diagnosis of complete or partial AVP deficiency as compared with primary polydipsia (Panel B). The results are shown according to the level of copeptin after arginine-stimulation testing and hypertonic-saline testing. The y axis is shown on a log scale for better visualization. The horizontal line in each box represents the median, the lower and upper boundaries of the boxes represent the interquartile range, the ends of the whiskers represent the minimum and maximum values within 1.5 times the interquartile range, and the individual data points represent outliers.

diagnostic potential of arginine-stimulated copeptin by deriving the best cutoffs using Youden's J statistic (jointly maximizing sensitivity and specificity).

All the diagnostic analyses were performed in two modified intention-to-treat (mITT) populations: mITT1, which included all the patients who had received a final diagnosis, and mITT2, which excluded patients with severe nausea or vomiting (postrandomization exclusion). Safety analyses were performed in the intention-to-treat population, which included all the patients who had started at least one diagnostic test. The widths of the confidence intervals have not been adjusted for multiplicity and may not be used in place of hypothesis testing. All the analyses were prespecified and conducted with the use of the statistical software package R (version 4.2.3).[17]

RESULTS

PATIENTS

A total of 177 patients were included in the trial. Of these patients, 13 were excluded after randomization (Fig. S1 in the Supplementary Appendix). Of the remaining patients, 164 underwent at least one test; 6 withdrew consent after the first diagnostic test (5 after arginine stimulation and 1 after hypertonic-saline stimulation). Data regarding prespecified clinical symptoms were available for 160 patients who were tested with arginine stimulation and 158 patients who were tested with hypertonic-saline stimulation. A total of 158 patients underwent both diagnostic tests, received a final diagnosis, and were evaluated for the primary outcome (mITT1). The prespecified mITT2 population excluded 22 patients with severe nausea (VAS, ≥8), vomiting, or both during the tests. The median interval between the two tests was 4 days (interquartile range, 1 to 8).

The trial population corresponded to the general published population of patients with AVP deficiency and primary polydipsia (Table S1). Of the 158 patients, 67% were women. In the final diagnosis, 69 (44%) were found to have AVP deficiency (41 [59%] with complete deficiency and 28 [41%] with partial deficiency), and 89

(56%) were found to have primary polydipsia (Table 1 and Table S2).

The main causes of AVP deficiency were postsurgical onset (in 30% of the patients), hypothalamic–pituitary lesions (in 26%), hypophysitis (in 12%), and idiopathic (in 12%). In addition, anterior pituitary deficiency was identified in 29 patients (42%).

Patients with complete AVP deficiency had higher quantities of polydipsia and polyuria than did patients with partial AVP deficiency or primary polydipsia (Table 1). Similar observations were made for baseline levels of copeptin and urine osmolality, which were lowest in patients with complete AVP deficiency. The characteristics of the patients who were assigned to be tested first with arginine stimulation (78 patients) or hypertonic-saline infusion (80 patients) were similar in the two subgroups (Table S3).

Pituitary MRI was performed in 108 patients (68%). Characteristics that were typical for AVP deficiency were observed in 67 patients (62%), among whom AVP deficiency was later diagnosed in 58 (Table 1).

PRIMARY OUTCOME

The overall diagnostic accuracy in differentiating patients with AVP deficiency from those with primary polydipsia was 74.4% (95% confidence interval [CI], 67.0 to 80.6) with arginine stimulation and 95.6% (95% CI, 91.1 to 97.8) with hypertonic-saline stimulation (estimated difference, –21.2 percentage points (95% CI, –28.7 to –14.3) (Fig. 1, Table 2, and Fig. S2). Thus, arginine stimulation did not meet the noninferiority margin, which was prespecified as an overall diagnostic accuracy that was no more than 10 percentage points lower than the diagnostic accuracy with hypertonic-saline stimulation. The area under the curve was 0.85 (95% CI, 0.80 to 0.91) for arginine-stimulated copeptin and 0.99 (95% CI, 0.98 to 1.00) for hypertonic saline–stimulated copeptin (Fig. S3).

Test performance was similar to the primary results after the exclusion of patients with severe nausea or vomiting (mITT2), with a diagnostic accuracy of 72.6% (95% CI, 64.5 to 79.4) for arginine stimulation and 96.3% (95% CI, 91.7 to 98.4) for hypertonic-saline stimulation.

Arginine-stimulated copeptin also had inferior performance as compared with hypertonic saline–stimulated copeptin in the differentiation

Table 2. Comparison of Arginine vs. Hypertonic-Saline Stimulation in the Diagnosis of AVP Deficiency vs. Primary Polydipsia.

Test and Diagnosis	Patients	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	Positive Predictive Value (95% CI)	Negative Predictive Value (95% CI)
	no.	*percentage*				
AVP deficiency vs. primary polydipsia						
Arginine-stimulated copeptin at 3.8 pmol/liter	156	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)
Hypertonic saline–stimulated copeptin at 4.9 pmol/liter	158	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)
Partial AVP deficiency vs. primary polydipsia						
Arginine-stimulated copeptin at 3.8 pmol/liter	115	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)
Hypertonic saline–stimulated copeptin at 4.9 pmol/liter	117	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)

Table 3. Performance of Arginine-Stimulated Copeptin Values in the Diagnosis of AVP Deficiency vs. Primary Polydipsia.*

Diagnosis of AVP Deficiency			Diagnosis of Primary Polydipsia		
Threshold Copeptin Value	Specificity (95% CI)	Sensitivity (95% CI)	Threshold Copeptin Value	Specificity (95% CI)	Sensitivity (95% CI)
pmol/liter	*percentage*		*pmol/liter*	*percentage*	
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)

* The copeptin cutoff values that had more than 90% specificity (on the basis of point estimates) for each diagnosis are marked in bold.

between partial AVP deficiency and primary polydipsia (Fig. 1, Table 2, and Fig. S2). Details regarding the patients who were misclassified by the two tests are described in the Supplementary Appendix, including in Table S4.

SECONDARY OUTCOMES

The use of other prespecified copeptin cutoffs resulted in overall diagnostic accuracies that were similar to the primary results. The diagnostic accuracy of the arginine-stimulation test was 75.0% (95% CI, 67.7 to 81.1) with a copeptin cutoff of 3.7 pmol per liter after 60 minutes and 79.2% (95% CI, 72.1 to 84.9) with a copeptin cutoff of 4.1 pmol per liter after 90 minutes. The diagnostic accuracy of the hypertonic-saline stimulation was 96.2% (95% CI, 92.0 to 98.2) for a copeptin cutoff of 6.5 pmol per liter (Table S5).

For hypertonic saline stimulation, the prespecified copeptin cutoff of 2.7 pmol per liter differentiated between complete and partial AVP deficiency, with a diagnostic accuracy of 88.4% (95% CI, 78.8 to 94.0) with a sensitivity of 92.7% (95% CI, 80.6 to 97.5) and a specificity of 82.1% (95% CI, 64.4 to 92.1).

Exploratory analyses of data-derived best copeptin cutoff values did not reveal any material difference in performance (Fig. S4). However, an arginine-stimulated copeptin level of 3.0 pmol per liter or less led to a diagnosis of AVP deficiency with a specificity of 90.9% (95% CI, 81.7 to 95.7) and a sensitivity of 59.5% (95% CI, 49.1 to 69.1), whereas levels of more than 5.2 pmol per liter led to a diagnosis of primary polydipsia with a specificity of 91.4% (95% CI, 83.7 to 95.6) and a sensitivity of 56.4% (95% CI, 44.6 to 67.4) (Table 3 and Fig. S2).

The application of these two cutoffs to our cohort (i.e., 156 patients in the ITT population for whom copeptin measures were available) allowed for a correct test result in 91 of 156 patients (58.3%; 95% CI, 50.5 to 65.8). The same cutoffs led to an inconclusive test result in 48 of 156 patients (30.8%; 95% CI, 24.1 to 38.4) and an incorrect test result in 17 of 156 patients (10.9%; 95% CI, 6.9 to 16.8).

Table 4. Adverse Events.

Variable	Arginine-Stimulation Test		Hypertonic Saline–Stimulation Test	
	Patients	VAS Score*	Patients	VAS Score*
Prespecified clinical symptoms — no. (%)†				
Thirst	158 (99)	8.0 (7.0–9.0)	155 (98)	9.0 (8.0–10.0)
Vertigo	42 (26)	3.5 (2.0–5.0)	75 (47)	5.0 (3.0–6.5)
Headache	59 (37)	3.0 (2.0–5.5)	94 (59)	4.0 (3.0–7.0)
Nausea	40 (25)	3.5 (1.0–6.0)	50 (32)	3.5 (2.0–7.0)
Malaise	52 (32)	3.5 (2.0–5.5)	81 (51)	5.0 (3.0–7.0)
Overall symptom burden		2.0 (0–3.0)		4.0 (2.0–7.0)
Adverse events — no. (%)‡				
Neuromuscular symptoms§	6 (4)		23 (14)	
Emesis	11 (7)		9 (6)	
Symptomatic hypoglycemia	1 (1)		0	
Dyspnea or coughing	0		3 (2)	
Rash or urticaria	1 (1)		1 (1)	
Weakness	1 (1)		2 (1)	
Diarrhea	0		1 (1)	
Back pain	0		1 (1)	
Presyncope after venous cannulation	1 (1)		0	
Patients' assessment — no./total no. (%)				
Preference in choice of test¶	103/143 (72)		17/143 (12)	

* Scores on the visual-analogue scale (VAS) are reported as the median and interquartile range. Scores range from 0 to 10, with 0 indicating no symptoms and 10 indicating severe symptoms.
† Data regarding prespecified clinical symptoms were available for 160 patients who were tested with arginine stimulation and 158 patients who were tested with hypertonic-saline stimulation.
‡ Data regarding adverse events were available for 163 patients who were tested with arginine stimulation and 159 patients who were tested with hypertonic-saline stimulation.
§ Neuromuscular symptoms included agitation, blurred vision, muscle spasms, paresthesia, shivering, and tremor.
¶ Data regarding patients' preference between the two tests were available for 143 patients; of those patients, 23 indicated no preference.

SAFETY OUTCOMES

In general, the patients reported no unacceptable side effects with either test (Table 4). Nearly all the patients reported severe thirst (median VAS, 8.0; interquartile range, 7.0 to 9.0) at the end of the arginine stimulation test, followed by mild headache (in 37% [median VAS, 3.0; interquartile range, 2.0 to 5.5]) and malaise (in 32% [median VAS, 3.5; interquartile range, 2.0 to 5.5]). Severe thirst was also the main adverse effect of hypertonic-saline stimulation (in 98% [median VAS, 9.0; interquartile range, 8.0 to 10.0]), followed by mild headache (in 59% [median VAS, 4.0; interquartile range, 3.0 to 7.0]) and malaise (in 51% [median VAS, 5.0; interquartile range, 3.0 to 7.0]).

The overall intensity of adverse effects was low with both tests but occurred with more frequency and higher intensity during hypertonic-saline stimulation. No severe adverse events occurred during either test. No adverse events were noted in the 6 patients who had withdrawn consent after the first test. Overall, the majority of patients (72%) preferred the arginine test to the hypertonic-saline test.

DISCUSSION

In this trial, we found that arginine-stimulated copeptin was inferior to hypertonic saline–stimulated copeptin in the diagnosis of AVP deficiency. Furthermore, arginine-stimulated copeptin showed a greater overlap between diagnoses of AVP deficiency and primary polydipsia. Hypertonic saline–stimulated copeptin thus was shown to be the test with higher diagnostic accuracy and confirmed safety. Nevertheless, arginine

stimulation was preferred by patients, and arginine-stimulated copeptin levels of 3.0 pmol or less per liter and more than 5.2 pmol per liter showed high specificity in correctly diagnosing AVP deficiency or primary polydipsia in more than half the patients.

The diagnostic performance of arginine stimulation in this cohort was lower than the previously described 93%.[13,18] The previous accuracy was derived from the smaller monocentric CARGO trial,[13] which involved 96 patients, of whom 40% had AVP deficiency and 60% had primary polydipsia — similar to the distribution in the current trial. According to the mITT2 analysis, severe nausea or emesis — which are nonosmotic copeptin stimuli[19,20] — were responsible for copeptin overstimulation in 3 patients. Several factors may explain the worse performance of arginine and the weaker-than-expected copeptin stimulation than that in the CARGO trial. First, symptom severity among patients with primary polydipsia was more accentuated in the current cohort. Although the distribution of polydipsia and polyuria was similar to that in the CARGO trial, patients with primary polydipsia in the current cohort had lower baseline values for serum and urinary osmolality. It is possible that the diagnostic accuracy of arginine stimulation could be improved by raising serum osmolality by overnight water deprivation. Second, the 40-g upper limit in the arginine dose in the current cohort may have led to weaker stimulative potency in the 3 patients with obesity. Third, arginine stimulation had a stronger comparator in the current trial. In the CARGO trial, the water deprivation test (known diagnostic accuracy, 70 to 77%[6,7]) was part of the expert diagnosis, whereas in the current trial such diagnosis was based on the hypertonic-saline test.

This last finding highlights the second important result of the current trial. The high diagnostic accuracy of the hypertonic-saline test was validated at 95.6%.[6] The hypertonic saline–stimulated copeptin also differentiated reliably between partial and complete AVP deficiency.

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

However, hypertonic-saline testing has some limitations. First, it can be performed only in patients in whom appropriate venous accesses can be placed and in settings in which constant surveillance and rapid sodium measurements are available. In addition, limited safety data are available in patients older than 65 years of age, and several coexisting illnesses that were exclusion criteria in the current trial prevented patients from receiving this diagnostic evaluation.

For decades, arginine stimulation has been performed for evaluation of the anterior pituitary.[21,22] Most clinicians are familiar with its protocol, which can be performed in the outpatient setting. Arginine stimulation is shorter than hypertonic-saline stimulation, and in the current trial, it led to fewer side effects and was preferred by patients. Accordingly, the use of arginine stimulation can be recommended as an initial diagnostic test. In this regard, arginine stimulation is also preferable to the water deprivation test.

Although arginine stimulation did not result in a single hoped-for copeptin cutoff value, it showed high specificity in diagnosing AVP deficiency and primary polydipsia according to copeptin cutoffs of 3.0 pmol or less and more than 5.2 pmol per liter. However, patients who have copeptin levels between these cutoff values or who have severe nausea or vomiting during the arginine stimulation may be advised to undergo hypertonic-saline stimulation for further evaluation.

The importance of a reliable diagnostic test was emphasized by overlapping clinical and laboratory characteristics of patients with partial AVP deficiency and those with primary polydipsia, who showed no difference in the amount of polydipsia and polyuria nor in urine osmolality. Pituitary MRI was performed in two thirds of the patients. In such patients with a high pretest probability of having AVP deficiency, findings that were typical for AVP deficiency were seen in 58 of 67 patients. Conversely, several patients with AVP deficiency had no abnormalities and several patients with primary polydipsia had false positive results. Accordingly, MRI findings will always need to be evaluated in the clinical context.

The main limitation of our trial is the absence of a clear diagnostic standard for AVP deficiency. Although the diagnoses were based on careful review of all patient data, they also included the outcome of the hypertonic-saline stimulation. To overcome this incorporation bias, the treatment response at 3 months was integrated into the

final diagnosis. Nevertheless, the diagnostic value of hypertonic saline–stimulated copeptin may be overestimated. The strength of the trial is the randomized, international design and large sample size of well-characterized patients with AVP deficiency and primary polydipsia.

For the diagnosis of AVP deficiency, arginine-stimulated copeptin was inferior to hypertonic saline–stimulated copeptin, although arginine-stimulated copeptin was preferred by the trial patients.

Supported by grants (P2BSP3-18172 from the Swiss National Science Foundation and from the Goldschmidt Jacobson Foundation) to Dr. Refardt. Dr. Atila received the Young Talents in Clinical Research grant from the Swiss Academy of Medical Sciences and the G&J Bangerter–Rhyner Foundation. Dr. Drummond was supported by a grant (FAPEMIG-APQ-01521-21) from of the Minas Gerais Research Support Foundation. Dr. Gurnell was supported by a grant (NIHR203312) from the NIHR Cambridge Biomedical Research Centre. Dr. Christ-Crain was supported by a grant (SNF-162608) from the Swiss National Science Foundation. Laboratory measurement of copeptin was funded by Thermo Fisher Scientific.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

A data sharing statement provided by the authors is available with the full text of this article at NEJM.org.

We thank all the patients for their participation, the medical and laboratory staff at all trial sites for their contributions, and all the trial nurses and staff members for their invaluable support.

APPENDIX

The authors' affiliations are as follows: the Departments of Endocrinology, Diabetology, and Metabolism (J.R., C.A., C.O.S., A.W., S.F., N.H., B.W., M.C.-C.) and Clinical Research (J.R., C.A., C.O.S., A.W., S.F., N.H., D.R.V., B.W., M.C.-C.), University Hospital Basel, University of Basel, Basel, and the Department of Endocrinology, Diabetology, and Clinical Nutrition, University Hospital Zurich and University of Zurich (Z.E., F.B.), and the LOOP Zurich–Medical Research Center (F.B.), Zurich — all in Switzerland; the Department of Internal Medicine, Section of Endocrinology, Erasmus Medical Center, Rotterdam, the Netherlands (J.R., R.C.D., J.H.); the Division of Endocrinology and Diabetes, Department of Internal Medicine I, University Hospital, University of Würzburg (I.C., M.F.), and Central Laboratory, University Hospital Würzburg (M.F.), Würzburg, and Medizinische Klinik und Poliklinik IV, Klinikum der Universität, Ludwig Maximilians Universität München, Munich (F.B.) — all in Germany; the Endocrinology Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico (E.F., R.I.), and the Department of Clinical Sciences and Community Health, University of Milan (R.I.) — both in Milan; the Department of Internal Medicine, Medical School of the Federal University of Minas Gerais, Belo Horizonte, Brazil (J.B.D., B.S.S.); and Wellcome–MRC Institute of Metabolic Science, University of Cambridge and Addenbrooke's Hospital, Cambridge Biomedical Campus (A.S.P., M.G.) and Cambridge NIHR Biomedical Research Centre, Cambridge University Hospitals (M.G.) — both in Cambridge, United Kingdom.

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(5):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, Hörning 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. Vienna: R Foundation for Statistical Computing, 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(5):R671-R678.
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 MCS, de Herder WW. The science behind the relations among cancer, height, growth patterns, and growth hormone axis. Endocr Relat Cancer 2023;30(4):e220400.
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(7):e220504.