Disorders of Water Metabolism

Joseph G. Verbalis, MD

CONTENTS

INTRODUCTION

BODY FLUID COMPARTMENTS

TOTAL AND EFFECTIVE OSMOLALITY

WATER METABOLISM

SODIUM METABOLISM

HYPEROSMOLALITY AND HYPERNATREMIA

HYPOOSMOLALITY AND HYponATREMIA

REFERENCES

INTRODUCTION

Disorders of body fluids are among the most commonly encountered problems in the practice of clinical medicine. This is in large part because many different disease states can potentially disrupt the finely balanced mechanisms that control the intake and output of water and solute. Since body water is the primary determinant of the osmolality of the extracellular fluid (ECF), disorders of water metabolism can be broadly divided into hypopsmolar disorders, in which there is an excess of body water relative to body solute, and hyperosmolar disorders, in which there is a deficiency of body water relative to body solute. Because sodium is the main constituent of plasma osmolality (P_{osm}), these disorders are typically characterized by hyponatremia and hypernatremia, respectively. Before discussing these disorders, this chapter will first review the regulatory mechanisms underlying water and sodium metabolism, the two major determinants of body fluid homeostasis.

BODY FLUID COMPARTMENTS

Water constitutes approx 55–65% of body weight, varying somewhat with age, sex, and amount of body fat, and, therefore, constitutes the largest single constituent of the body. Total body water (TBW) is distributed between the intracellular fluid (ICF) and the ECF compartments. Estimates of the relative
sizes of these two important pools differ significantly depending on the tracer used to measure the ECF volume, but most studies in animals and man have suggested that 55–65% (or just under two-thirds) of TBW resides in the ICF, and 35–45% (or slightly more than one-third) is in the ECF (1). Approximately three-quarters of the ECF compartment is interstitial fluid, and one-fourth is intravascular fluid (blood volume). Figure 1 summarizes the estimated body fluid spaces of an average weight adult. The solute composition of the ICF and ECF differ considerably, since membrane-bound Na⁺/K⁺ pumps maintain Na⁺ in a primarily extracellular location and K⁺ in a primarily intracellular location. Nonetheless, it is important to remember that the osmotic pressure, which is a function of the concentrations of all the solutes in a fluid compartment, must always be equivalent in the ICF and ECF, because most biological membranes are semipermeable (i.e., freely permeable to water, but not to aqueous solutes). Thus, water will flow across membranes into a compartment with a higher solute concentration until a steady state is reached, in which the osmotic pressures have equalized on both sides of the cell membrane. An important consequence of this thermodynamic law is that the volume of distribution of body Na⁺ and K⁺ is actually the TBW rather than just the ECF or ICF volume, respectively (2). For example, any increase in ECF [Na⁺] will cause water to shift from the ICF to the ECF until the ICF and ECF osmotic pressures are equal, thereby in effect distributing the Na⁺ across both extracellular and intracellular water.

TOTAL AND EFFECTIVE OSMOLALITY

Osmolality is defined as the concentration of all of the solutes in a given weight of water. \( P_{\text{osm}} \) can be measured directly (via determination of freezing point depression or vapor pressure, since each of these are colligative properties of the number of free solute particles in a given volume of plasma) or estimated as:

\[
P_{\text{osm}} \text{ (mOsm/kg H}_2\text{O)} = 2 \times \text{ plasma [Na}^+\text{]} \text{ (mEq/L)} + \frac{\text{glucose (mg/dL)}}{18} + \frac{\text{BUN (mg/dL)}}{2.8}
\]

Both methods produce comparable results under most conditions, as will simply doubling the plasma sodium concentration ([Na⁺]), since sodium and its accompanying anions are by far the predominant solutes present in plasma. However, the total osmolality of plasma is not always equivalent to the “effective” osmolality (sometimes referred to as the “tonicity” of the plasma), because the latter is a function of the relative solute permeability properties of the membranes separating the two compartments. Solutes that are impermeable to cell membranes (Na⁺, mannitol) are restricted to the ECF compartment and are effective solutes, since they create osmotic pressure gradients across cell membranes, leading to osmotic movement of water from the ICF to the ECF compartments. Solutes that are permeable to cell membranes (urea, ethanol, methanol) are
ineffective solutes, since they do not create osmotic pressure gradients across cell membranes and, therefore, are not associated with such water shifts (3). Glucose is a unique solute, since at normal physiologic plasma concentrations, it is taken up by cells via active transport mechanisms and, therefore, acts as an ineffective solute, but under conditions of impaired cellular uptake (e.g., insulin deficiency), it becomes an effective extracellular solute (4).

The importance of this distinction between total and effective osmolality lies with the fact that only the effective solutes in plasma are determinants of whether clinically significant hyperosmolality or hypooosmolality is present. An example of this is uremia: a patient with a urea concentration that has increased by 30 mEq/L will have a corresponding 30 mOsm/kg H₂O elevation in Posm, but the effective osmolality will remain normal, since the increased urea is proportionally distrib-
uted across both the ECF and ICF. In contrast, a patient whose plasma \([\text{Na}^+]\) has increased by 15 mEq/L will also have a 30 mOsm/kg \(\text{H}_2\text{O}\) elevation of \(\text{Posm}\), since the increased cation must be balanced by an equivalent increase in plasma anions. In this case, however, the effective osmolality will also be elevated by 30 mOsm/kg \(\text{H}_2\text{O}\), since the \([\text{Na}^+]\) and accompanying anions will largely remain restricted to the ECF due to the relative impermeability of cell membranes to \([\text{Na}^+]\) and other univalent ions. Thus, elevations of solutes such as urea, unlike elevations in plasma \([\text{Na}^+]\), do not cause cellular dehydration and, consequently, do not activate mechanisms that defend body fluid homeostasis by acting to increase body water stores.

**WATER METABOLISM**

Water metabolism represents a balance between the intake and excretion of water. Each side of this balance equation can be considered to consist of a “regulated” and an “unregulated” component, the magnitudes of which can vary quite markedly under different physiological and pathophysiological conditions. The unregulated component of water intake consists of the intrinsic water content of ingested foods, the consumption of beverages primarily for reasons of palatability or desired secondary effects (e.g., caffeine), or for social or habitual reasons (e.g., alcoholic beverages), whereas the regulated component of water intake consists of fluids consumed in response to a perceived sensation of thirst. Similarly, the unregulated component of water excretion occurs via insensible water losses from a variety of sources (cutaneous losses from sweating, evaporative losses in exhaled air, gastrointestinal losses), as well as the obligate amount of water that the kidneys must excrete to eliminate solutes generated by body metabolism. The regulated component of water excretion is comprised of the renal excretion of free water in excess of the obligate amount necessary to excrete metabolic solutes \(^{(5)}\). In effect, the regulated components are those that act to maintain water balance by compensating for whatever perturbations result from unregulated water losses or gains. Within this framework, it is clear that the two major mechanisms responsible for regulating water metabolism are thirst and pituitary secretion of the hormone vasopressin.

**Thirst**

Thirst is the body’s defense mechanism to increase water consumption in response to perceived deficits of body fluids. Thirst can be stimulated in animals and man either by intracellular dehydration, caused by increases in the effective osmolality of the ECF, or by intravascular hypovolemia, caused by losses of ECF. Substantial evidence to date has supported mediation of the former by osmoreceptors located in the anterior hypothalamus of the brain, whereas the
latter appears to be stimulated primarily via activation of low- and/or high-pressure baroreceptors, with a likely contribution from circulating angiotensin II during more severe degrees of intravascular hypovolemia and hypotension (6,7). Controlled studies in animals have consistently reported thresholds for osmotically induced drinking, ranging from 1–4% increases in $P_{\text{osm}}$ above basal levels, and analogous studies in humans using quantitative estimates of subjective symptoms of thirst have confirmed that increases in $P_{\text{osm}}$ of similar magnitudes are necessary to produce an unequivocal sensation described as thirst (8,9).

Conversely, the threshold for producing hypovolemic, or extracellular, thirst is significantly greater in both animals and humans. Studies in several species have shown that sustained decreases in plasma volume or blood pressure of at least 4–8%, and in some species 10–15%, are necessary to consistently stimulate drinking. In humans, it has been difficult to demonstrate any effects of mild to moderate hypovolemia to stimulate thirst independently of osmotic changes occurring with dehydration. This blunted sensitivity to changes in ECF volume or blood pressure in humans probably represents an adaptation that occurred as a result of the erect posture of primates, which predisposes them to wider fluctuations in blood and atrial filling pressures as a result of orthostatic pooling of blood in the lower body; stimulation of thirst (and vasopressin secretion) by such transient postural changes in blood pressure might lead to overdrinking and inappropriate antidiuresis in situations where the ECF volume was actually normal but only transiently maldistributed. Consistent with a blunted response to baroreceptor activation, recent studies have also shown that systemic infusion of angiotensin II to pharmacological levels is a much less potent stimulus to thirst in humans than in animals (10). Nonetheless, this response is not completely absent in humans, as demonstrated by rare cases of polydipsia in patients with pathological causes of hyperreninemia.

Although osmotic changes clearly are more effective stimulants of thirst than are volume changes in humans, it is not clear whether relatively small changes in $P_{\text{osm}}$ are responsible for day-to-day fluid intakes. Most humans consume the majority of their ingested water as a result of the unregulated components of fluid intake discussed previously, and generally ingest volumes in excess of what can be considered to be actual “need” (11). Consistent with this observation is the fact that, under most conditions, $P_{\text{osm}}$ in man remain within 1–2% of basal levels, and these relatively small changes in $P_{\text{osm}}$ are generally below the threshold levels that have been found to stimulate thirst in most individuals. This suggests that despite the obvious vital importance of thirst during pathological situations of hyperosmolality and hypovolemia, under normal physiological conditions, water balance in man is accomplished more by regulated free water excretion than by regulated water intake (5).
Arginine Vasopressin Secretion

The prime determinant of free water excretion in animals and man is the regulation of urinary flow by circulating levels of arginine vasopressin (AVP) in plasma. Before AVP was biochemically characterized, early studies of antidiuresis used the term “antidiuretic hormone” (ADH) to describe this substance. Now that its structure and function as the only naturally-occurring antidiuretic substance are known, it is more appropriate to refer to it by its real name. AVP is a 9-amino acid peptide that is synthesized in specialized (magnocellular) neural cells located in two discrete areas of the hypothalamus, the supraoptic (SON) and paraventricular (PVN) nuclei. The synthesized peptide is enzymatically cleaved from its prohormone and is transported to the posterior pituitary where it is stored within neurosecretory granules until specific stimuli cause secretion of AVP into the bloodstream (12). Antidiuresis then occurs via interaction of the circulating hormone with AVP V2 receptors in the kidney, which results in increased water permeability of the collecting duct through the insertion of a water channel called aquaporin-2 into the apical membranes of collecting tubule principal cells (13). The importance of AVP for maintaining water balance is underscored by the fact that the normal pituitary stores of this hormone are very large, allowing more than a week’s supply of hormone for maximal antidiuresis under conditions of sustained dehydration. Knowledge of the different conditions that stimulate pituitary AVP release in man is, therefore, essential for understanding water metabolism.

Osmotic Regulation

The primary renal response to AVP is an increase in water permeability of the kidney collecting tubules. Although an increase in solute reabsorption (primarily urea) occurs as well, the total solute reabsorption is proportionally much less than water. Consequently, a decrease in urine flow and an increase in Uosm occur as secondary responses to the increased net water reabsorption. With refinement of radioimmunoassays for AVP, the unique sensitivity of this hormone to small changes in osmolality, as well as the corresponding sensitivity of the kidney to small changes in plasma AVP levels, have become apparent (14). Although some debate still exists with regard to the exact pattern of osmotically stimulated AVP secretion, most studies to date have supported the concept of a discrete osmotic threshold for AVP secretion above which a linear relationship between P_{osm} and AVP levels occurs (Fig. 2). The slope of the regression line relating AVP to P_{osm} can vary significantly across individual human subjects, in part because of genetic factors (12). In general, each 1 mOsm/kg H2O increase in P_{osm} causes an increase in plasma AVP level from 0.4 to 0.8 pg/mL. The renal response to circulating AVP is similarly linear, with urinary concentration that is directly proportional to AVP levels from 0.5 to 4–5 pg/mL, after which urinary osmolality (Uosm) is maximal and cannot increase further despite additional increases in AVP levels.
Thus, changes of 1% or less in Posm are sufficient to cause significant increases in plasma AVP levels with proportional increases in urine concentration, and maximal antidiuresis is achieved after increases in Posm of only 5–10 mOsm/kg H2O (2–4%) above the threshold for AVP secretion. However, even this analysis underestimates the sensitivity of this system to regulate free water excretion for the following reason. U osm is directly proportional to plasma AVP levels as a consequence of the fall in urine flow induced by the AVP, but urine volume is inversely related to Uosm (Fig. 3). Thus, an increase in plasma AVP concentration from 0.5–2 pg/mL has a much greater relative effect to decrease urine flow than does a subsequent increase in AVP concentration from 2–5 pg/mL, thereby further magnifying the physiological effects of small initial changes in plasma AVP levels (15). The net result of these relations is a finely tuned regulatory system that adjusts the rate of free water excretion accurately to the ambient P osm via changes in pituitary AVP secretion. Furthermore, the rapid response of pituitary AVP secretion to changes in P osm coupled with the short half-life (10–20 minutes) of AVP in human plasma enables this regulatory system to adjust renal water excretion to changes in P osm on a minute-to-minute basis.

VOLEMIC REGULATION

As in the case of thirst, hypovolemia also is a stimulus for AVP secretion in man; an appropriate physiological response to volume depletion should include urinary concentration and renal water conservation. But similar to thirst, AVP
secretion is much less sensitive to small changes in blood volume and blood pressure than to changes in osmolality (12); some have even suggested that the AVP response to decreases in blood volume is absent in man, though this most likely is simply a manifestation of the significantly higher threshold for AVP secretion to volemic stimuli. Such marked differences in AVP responses represent additional corroborative evidence that osmolality represents a more sensitive regulatory system for water balance than does blood or ECF volume.

Fig. 3. Schematic representation of normal physiological relationships among $P_{\text{osm}}$, plasma AVP concentrations, $U_{\text{osm}}$, and urine volume in man. Note particularly the inverse nature of the relation between $U_{\text{osm}}$ and urine volume, resulting in disproportionate effects of small changes in plasma AVP concentrations on urine volume at lower AVP levels (modified with permission from ref. 15).
Nonetheless, modest changes in blood volume and pressure influence AVP secretion indirectly, even though they are weak stimuli by themselves. This occurs via shifting the sensitivity of AVP secretion to osmotic stimuli, so that a given increase in osmolality will cause a greater secretion of AVP during hypovolemic conditions than during euvolemic states (Fig. 4). Although this effect has been demonstrated in human as well as in animal studies, it has only been shown convincingly with substantial degrees of hypovolemia, and the magnitude of this effect during mild degrees of volume depletion remains conjectural. Consequently, it is reasonable to conclude that the major effect of moderate degrees of hypovolemia on both AVP secretion and thirst is to modulate the gain of the osmoregulatory responses, with direct effects on thirst and AVP secretion occurring only during more severe degrees of hypovolemia (e.g., 10% reductions in blood volume).

Other Stimuli

Several other nonosmotic stimuli to AVP secretion have been described in man. Most prominent among these is nausea. The sensation of nausea, with or without vomiting, is by far the most potent stimulus to AVP secretion known in
man. While 20% increases in osmolality will typically elevate plasma AVP levels to the range of 5–20 pg/mL, and 20% decreases in blood pressure to 10–100 pg/mL, nausea has been described to cause AVP elevations in excess of 200–400 pg/mL (16). The reason for this profound stimulation is not known (although it has been speculated that the AVP response assists evacuation of stomach contents via contraction of gastric smooth muscle, AVP is not necessary for vomiting to occur), but it is probably responsible for the intense vasoconstriction, which produces the pallor often associated with this state. Hypoglycemia also stimulates AVP release in man, but to relatively low levels that are not consistent among individuals. As will be discussed in the clinical disorders, a variety of drugs also stimulate AVP secretion, including nicotine (17). However, despite the importance of these stimuli during pathological conditions, none of them is a significant determinant of physiological regulation of AVP secretion in man.

Integration of Thirst and AVP Secretion

A synthesis of what is presently known about the regulation of thirst and AVP secretion in man leads to a relatively simple but elegant system to maintain water balance. Under normal physiological conditions, the sensitivity of the osmoregulatory system for AVP secretion accounts for maintenance of $P_{\text{osm}}$ within narrow limits by adjusting renal water excretion to small changes in osmolality. Stimulated thirst does not represent a major regulatory mechanism under these conditions, and unregulated fluid ingestion supplies adequate water in excess of true need, which is then excreted in relation to osmoregulated pituitary AVP secretion. However, when unregulated water intake cannot adequately supply body needs in the presence of plasma AVP levels sufficient to produce maximal antidiuresis, then $P_{\text{osm}}$ rises to levels that stimulate thirst and produce water intake proportional to the elevation of osmolality above this threshold. In such a system, thirst essentially represents a backup mechanism called into play when pituitary and renal mechanisms prove insufficient to maintain $P_{\text{osm}}$ within a few percent of basal levels. This arrangement has the advantage of freeing man from frequent episodes of thirst that would require a diversion of activities toward behavior oriented to seeking water when water deficiency is sufficiently mild to be compensated for by renal water conservation, but would stimulate water ingestion once water deficiency reaches potentially harmful levels. Stimulation of AVP secretion at $P_{\text{osm}}$ below the threshold for subjective thirst acts to maintain an excess of body water sufficient to eliminate the need to drink whenever slight elevations in $P_{\text{osm}}$ occur. This system of differential effective thresholds for thirst and AVP secretion nicely complements many studies that have demonstrated excess unregulated, or need-free, drinking in both man and animals (6).

Therefore, in summary, during normal day-to-day conditions, body water homeostasis appears to be maintained primarily by *ad libitum*, or unregulated, fluid intake in association with AVP-regulated changes in urine flow, most of
which occurs before the threshold is reached for osmotically stimulated, or regulated, thirst. But when these mechanisms become inadequate to maintain body fluid homeostasis, then thirst-induced regulated fluid intake becomes the predominant defense mechanism for the prevention of dehydration.

SODIUM METABOLISM

Maintenance of sodium homeostasis requires a simple balance between intake and excretion of Na+. As in the case of water metabolism, it is possible to define regulated and unregulated components of both Na+ intake and Na+ excretion. Unlike water intake, however, there is little evidence in humans to support a significant role for regulated Na+ intake, with the possible exception of some pathological conditions. Consequently, there is an even greater dependence on mechanisms for regulated renal excretion of sodium than is the case for excretion of water (18).

Whether for this reason or not, the mechanisms for renal excretion of sodium are more numerous and substantially more complex than the relatively simple, albeit quite efficient, system for AVP-regulated excretion of water.

Salt Appetite

The only solute for which any specific appetite has been clearly demonstrated in man is sodium (as with animals, this is generally expressed as an appetite for the chloride salt of sodium, so it is usually called NaCl, or salt, appetite). Because of the importance of Na+ for ensuring maintenance of the ECF volume, which in turn directly supports blood volume and pressure, its uniqueness insofar as merit a specific mechanism for regulated intake seems appropriate. However, despite abundant evidence in many different species demonstrating a salt appetite that is proportionately related to Na+ losses (19), there is only one pathological condition in which a specific stimulated sodium appetite has been unequivocally observed in humans, namely Addison’s disease, which is caused by adrenal insufficiency. Almost since the initial discovery of this disorder, salt craving has remained one of the well-known manifestations of Addison’s disease (20). A robust salt appetite also occurs prominently in adrenalectomized animals, and appears to be related in part to the high plasma levels of adrenocorticotropin (ACTH) produced as a result of the loss of cortisol feedback on the pituitary. However, despite the presence of Na+ deficiency in most patients with untreated Addison’s disease, only 15–20% of such patients manifest salt-seeking behavior (21).

Even more striking is the apparent absence of salt appetite during a variety of other disorders causing severe Na+ and ECF volume depletion in humans (patients with hemorrhagic blood loss, diuretic-induced hypovolemia, or hypotension of any etiology become thirsty when intravascular deficits are marked, but almost never express a pronounced desire for salty foods or fluids). Yet, as with thirst, the possibility of subclinical activation of neural mechanisms stimulating
salt intake without a conscious subjective sensation of salt “hunger” must be entertained. However, this possibility cannot be supported either, because many such patients actually become hyponatremic as a result of continued ingestion of only water or osmotically dilute fluids in response to their volume depletion (18). It is also interesting to note that athletes must be instructed to ingest sodium as NaCl tablets or electrolyte solutions during periods of sodium losses from profuse sweating since they fail to develop a salt appetite, which would be protective under these circumstances. As a corollary to the infrequency of stimulated salt appetite in man, there is also no evidence to support inhibition of sodium intake under conditions of Na+ and ECF excess, as demonstrated by the difficulty in maintaining even moderate degrees of sodium restriction in patients with edema-forming diseases such as congestive heart failure.

Renal Sodium Excretion

Although specific mechanisms exist for regulated renal excretion of all major electrolytes, none is as numerous or as complex as those controlling Na+ excretion, which is not surprising in view of the fact that maintenance of ECF volume is crucial to normal health and function. The most important of these mechanisms are discussed briefly below, but given their complexity, the reader is referred to more complete reviews of this subject (22,23).

Glomerular Filtration Rate

Glomerular filtration rate (GFR) is one of two classical mechanisms known to regulate renal Na+ excretion. Multiple factors influence GFR, including the glomerular plasma flow, the glomerular capillary surface area, the hydrostatic pressure gradient between the glomerular capillaries and Bowman’s capsule, and the oncotic pressure produced by the proteins in glomerular capillaries. Because the amount of Na+ filtered through the kidney is huge (approx 25,000 mmol/d in healthy adults), relatively small changes in GFR can potentially have large effects on filtered Na+. However, changes in filtered load of Na+ are compensated for by concomitant changes in proximal tubular sodium reabsorption via a process known as tubuloglomerular feedback (24). As the filtered Na+ load increases, Na+ absorption in the proximal tubule also increases, largely compensating for the increased filtered load. Although the mechanism(s) responsible for tubuloglomerular feedback are not completely understood, one important factor appears to be changes in peritubular capillary forces, which is analogous to the Starling forces in systemic capillaries. An increase in filtered fluid at the glomerulus decreases the hydrostatic pressure and increases the oncotic pressure of the nonfiltered fluid delivered to the peritubular capillaries, thereby increasing the pressure gradient for reabsorbing the Na+, which is actively transported from the proximal tubular epithelial cells into the extracellular fluid surrounding the proximal tubule. Although this mechanism dampens the effects of alterations in...
GFR on renal Na⁺ excretion and prevents large changes in urine Na⁺ excretion in response to minor changes in GFR. Nonetheless, many experimental results indicate that sustained alterations of GFR can significantly modulate renal Na⁺ excretion.

**Aldosterone**

The second major factor long known to influence renal Na⁺ excretion is adrenal aldosterone secretion, which increases Na⁺ resorption in the distal nephron by inducing the synthesis and activity of ion channels that affect sodium reabsorption and sodium–potassium exchange in tubular epithelial cells, particularly the epithelial sodium channel (ENaC) (25). The importance of this hormone for Na⁺ homeostasis is best illustrated by the well-known renal Na⁺ wasting of patients with primary adrenal insufficiency. Multiple factors stimulate adrenal mineralocorticoid secretion. Most prominent of these is angiotensin II, which is formed as the end result of renin secretion from the juxtaglomerular apparatus in response to renal hypoperfusion. High plasma K⁺ concentrations also stimulate aldosterone secretion, thereby increasing urinary K⁺ excretion at the expense of Na⁺ retention. More recently two inhibitors of aldosterone secretion have been described: atrial natriuretic peptide (ANP) and hyperosmolality; both of these stimuli appear to be sufficiently potent to completely block stimulated aldosterone secretion (26). Although aldosterone clearly plays an important role in sodium homeostasis, its effects to stimulate Na⁺ resorption in the distal tubule can be overridden by other natriuretic factors. This is evident in the phenomenon of renal “escape” from mineralocorticoids, in which experimental animals and man reestablish sodium balance after an initial period of Na⁺ retention and ECF volume expansion. Potential mechanisms responsible for this phenomenon are discussed below.

**Intrarenal Hemodynamic and Peritubular Factors**

Although GFR and aldosterone effects can account for much of the observed variation in renal Na⁺ excretion, it has long been known that they cannot completely explain the natriuresis that occurs in the absence of measurable changes in GFR or aldosterone secretion during isotonic saline volume expansion. This led to the postulation of the existence of a “third factor” or factors regulating Na⁺ excretion. Intrarenal hemodynamic factors are now known to be important in this regard, particularly changes in renal perfusion pressure. This is illustrated by aldosterone escape described above, which appears to be mediated primarily by increased renal perfusion pressure with subsequent increased fractional sodium excretion (27). In effect, this represents a “safety-valve” mechanism; when renal artery pressure rises as a result of volume expansion, the increase in filtered load of Na⁺ is sufficient to overwhelm the aldosterone-mediated distal sodium resorption. This phenomenon has been called a pressure diuresis and natriuresis. Note
that the term escape is somewhat of a misnomer, since aldosterone effects are still present, but a new steady-state of volume expansion has been reached in which no additional sodium retention occurs due to activation of compensatory mechanisms for sodium excretion. Although sodium balance is reestablished, a substantial degree of volume expansion persists nonetheless, thus confirming the presence of continued systemic mineralocorticoid effects.

OTHER FACTORS

Several factors in addition to those discussed above have also been found to influence renal sodium excretion. These include angiotensin II, arginine vasopressin, atrial natriuretic peptide, dopamine, renal sympathetic nerve activity, and renal prostaglandins. However, none of these have yet been clearly demonstrated to play a major role in regulating renal sodium excretion in man.

HYPEROSMOLALITY AND HYPERNATREMIA

Pathogenesis

Hyperosmolality indicates a deficiency of water relative to solute in the ECF. Because water moves freely between the ICF and ECF, this also indicates a deficiency of TBW relative to total body solute. Although hypernatremia can be caused by an excess of body sodium, the vast majority of cases are due to losses of body water in excess of body solutes, caused by either insufficient water intake or excessive water excretion. Consequently, most of the disorders causing hyperosmolality are those associated with inadequate water intake and/or deficient AVP secretion (Table 1). The best known of these is diabetes insipidus (DI), in which AVP secretion, or its renal effects, is impaired without an abnormality of thirst. Much less common are disorders of osmoreceptor function, resulting in abnormalities of both AVP secretion and thirst. Although hyperosmolality from inadequate water intake is seen frequently in clinical practice, this is usually not due to an underlying defect in thirst, but rather results from a generalized incapacity to obtain and/or ingest fluids, often stemming from a depressed sensorium. An example is hyperosmolar coma caused by renal water losses from hyperglycemia-induced diuresis in elderly patients who eventually are unable to drink enough fluid to keep up with their unrelenting osmotic diuresis.

Differential Diagnosis

Evaluation of hyperosmolar patients should include a careful history, clinical assessment of ECF volume, a thorough neurological evaluation, serum electrolytes, glucose, blood urea nitrogen (BUN), and creatinine, calculated and/or directly measured $P_{\text{osm}}$, simultaneous urine electrolytes and osmolality, and urine glucose. Hypernatremia is always synonymous with hyperosmolality, since $Na^+$ is the main constituent of $P_{\text{osm}}$, but hyperosmolality can exist without
Chapter 2/Water Metabolism Disorders

Table 1
Pathogenesis of Hyperosmolar Disorders

Water depletion (decreases in total body water in excess of body solute):

1. Insufficient water intake
   - Unavailability of water.
   - Hypodipsia (osmoreceptor dysfunction, age).
   - Neurological deficits (cognitive dysfunction, motor impairments).

2. Hypotonic fluid loss
   - A. Renal: Diabetes Insipidus
     - Insufficient AVP secretion (central DI, osmoreceptor dysfunction).
     - Insufficient AVP effect (nephrogenic DI).
   - B. Renal: Other Fluid Loss
     - Osmotic diuresis (hyperglycemia, mannitol).
     - Diuretic drugs (furosemide, ethacrynic acid, thiazides).
     - Postobstructive diuresis
     - Diuretic phase of acute tubular necrosis.
   - C. Nonrenal fluid loss
     - Gastrointestinal (vomiting, diarrhea, nasogastric suction).
     - Cutaneous (sweating, burns).
     - Pulmonary (hyperventilation).
     - Peritoneal dialysis.

Solute excess (increases in total body solute in excess of body water):

1. Sodium
   - Excess Na⁺ administration (NaCl, NaHCO₃).
   - Sea water drowning.

2. Other
   - Hyperalimentation (intravenous, parenteral).

*Most hypotonic fluid losses will not produce hyperosmolality unless insufficient free water is ingested or infused to replace the ongoing losses, so these disorders also usually involve some component of insufficient water intake.

Hypernatremia when there is an excess of non-sodium solute. This occurs most often with marked elevations of plasma glucose, as in patients with nonketotic hyperglycemic hyperosmolar coma. As for cases of artificial hypernatremia caused by elevated plasma lipids or protein, misdiagnosis can be avoided by direct measurement of P_\text{osm}^\text{urea} or by correcting the serum [Na⁺] by 1.6 mEq/L for each 100 mg/dL increase in plasma glucose concentration above 100 mg/dL (29), though more recent studies have indicated a more complex relation between hyperglycemia and serum [Na⁺] and suggested that a more accurate correction factor is closer to 2.4 mEq/L (30). Evaluation of the patient’s ECF volume status is important as a guide to fluid replacement therapy, but is not as useful for
differential diagnosis, since most hyperosmolar patients will manifest some degree of hypovolemia. Rather, assessment of urinary concentrating ability provides the most useful data with regard to the type of disorder present. Using this approach, disorders of hyperosmolality can be categorized as those in which renal water conservation mechanisms are intact, but are unable to compensate for inadequately replaced losses of hypotonic fluids from other sources, or those in which renal concentrating defects are a contributing factor to the deficiency of body water.

An appropriately concentrated urine in a hyperosmolar patient usually eliminates the possibility of a primary renal cause of the disorder in most cases. Maximum urine concentrating ability varies between individuals and decreases with age, but in general $U_{\text{osm}}$ above 800 mOsm/kg H$_2$O are considered sufficient to verify normal AVP secretion and renal response. In such cases, potential causes of nonrenal fluid losses should be investigated, particularly gastrointestinal and cutaneous losses (although subsequent ingestion of free water can produce hyperosmolality in such patients as a result of AVP-induced water retention). In the absence of disorders causing fluid losses, primary disorders of thirst should be considered, especially in the elderly who have a decreased sensation of thirst and ingest lesser amounts of fluids in response to induced dehydration (31). One situation in which a normally concentrated urine may not completely eliminate the possibility of an underlying renal concentrating defect is in patients with mild partial central DI, who can sometimes achieve maximally concentrated urine during extreme dehydration through a combination of severely limited GFR and stimulated AVP secretion at high $P_{\text{osm}}$, as will be discussed below. However, as $P_{\text{osm}}$ is corrected, these patients will demonstrate inappropriate dilution of their urine before reaching normal levels of $P_{\text{osm}}$.

An inappropriately low $U_{\text{osm}}$ (e.g., less than 800 mOsm/kg H$_2$O in a hyperosmolar patient) signifies the presence of a renal concentrating defect. The urine should always be checked for glucose, since a solute diuresis will limit urine concentrating ability and $U_{\text{osm}}$ can approach isotonicity at high rates of urine excretion. In the absence of glucosuria or any other cause of osmotic diuresis, inadequate urine concentration in a hyperosmolar patient generally indicates the presence of DI and further testing is then indicated to ascertain the etiology.

**DIABETES INSIPIDUS**

DI can result from either inadequate AVP secretion (central or neurogenic DI) or inadequate renal response to AVP (nephrogenic DI) (Table 2). Central DI is caused by a variety of acquired or congenital anatomic lesions that disrupt the hypothalamic-posterior pituitary axis, including pituitary surgery, tumors, trauma, hemorrhage, thrombosis, infarction, or granulomatous disease (12). Severe nephrogenic DI is most commonly congenital, due to defects in the gene
for the AVP V2 receptor (X-linked recessive pattern of inheritance) or in the gene for the aquaporin-2 water channel (autosomal recessive pattern of inheritance) (52), but relief of chronic urinary obstruction or therapy with drugs, such as lithium, can cause an acquired form sufficient to warrant treatment. Short-lived
nephrogenic DI can result from hypokalemia or hypercalcemia, but the mild concentrating defect generally does not by itself cause hyperoncity and responds to correction of the underlying disorder. Regardless of the etiology of the DI, the end result is a free water diuresis due to an inability to concentrate urine appropriately. Because renal mechanisms for sodium conservation are unimpaired, there is no accompanying sodium deficiency. Although untreated DI can lead to both hyperoncity and volume depletion, until the water losses become severe, volume depletion is minimized by osmotic shifts of water from the ICF to the more osmotically concentrated ECF. This phenomenon is not as evident following increases in ECF [Na⁺], since such osmotic shifts result in a slower increase in the serum [Na⁺] than would otherwise occur. However, when nonsodium solutes such as mannitol are infused, this effect is more obvious due to the progressive dilutional decrease in serum [Na⁺] caused by translocation of intracellular water to the ECF compartment.

Because patients with DI do not have impaired urine Na⁺ conservation, the ECF volume is generally not markedly decreased, and regulatory mechanisms for maintenance of osmotic homeostasis are primarily activated: stimulation of thirst and AVP secretion (to whatever degree the neurohypophysis is still able to secrete AVP). In cases where AVP secretion is totally absent (complete DI), patients are dependent entirely on water intake for maintenance of water balance. However, in cases where some residual capacity to secrete AVP remains (partial DI), P_{\text{max}} can eventually reach levels that allow moderate degrees of urinary concentration (recall from Fig. 3 that even small concentrations of AVP can have substantial effects to limit urine volume). As the P_{\text{max}} increases, some patients with partial DI can secrete enough AVP to achieve near maximal U_{osm} (Fig. 5). However, this should not cause confusion about the diagnosis of DI, since in such patients the U_{osm} will still be inappropriately low at P_{\text{max}} within normal ranges, and they will respond to exogenous AVP administration with a further rise in U_{osm}.

Distinguishing between central and nephrogenic DI in a patient who is already hyperonosmolar is straightforward and consists simply of evaluating the response to administered AVP (5 U subcutaneously [sc]) or, preferably, the AVP V₂ receptor agonist desmopressin (1-deamino-8-D-arginine vasopressin [dDAVP]; 1 µg sc or intravenously [IV]). A significant increase in U_{osm} within 1 to 2 h after injection indicates insufficient endogenous AVP secretion and, therefore, central DI, whereas an absent response indicates renal resistance to AVP effects and, therefore, nephrogenic DI (NDI) (15). Although conceptually simple, interpretational difficulties often arise because the water diuresis produced by AVP deficiency produces a “wash-out” of the renal medullary concentrating gradient, so that increases in U_{osm} in response to administered AVP or dDAVP are not as great as would be expected (more recent experimental results suggest that down-regulation of collecting tubule aquaporin-2 water channels as a result of
AVP deficiency also contributes to the blunted response to subsequent acute AVP or dDAVP administration (33). Generally, increases of $U_{\text{osm}}$ of 50% reliably indicate central DI, and responses of 10% indicate nephrogenic DI, but responses between 10–50% are less certain (34). For this reason, plasma AVP levels should be measured to aid in this distinction: hyperosmolar patients with nephrogenic DI will have clearly elevated AVP levels, while those with central DI will have absent (complete) or blunted (partial) AVP responses relative to their $P_{\text{osm}}$. Since it will not be known beforehand which patients will have diagnostic vs indeterminate responses to AVP or dDAVP, a plasma AVP level should be drawn prior to AVP or dDAVP administration in all patients (35). One drawback to using the AVP levels for diagnosis is the relatively long turnaround time (4–10 d in most laboratories) for results. An alternative in such cases is to continue dDAVP treatment for 1 to 2 d as a clinical trial; if central DI is present, the medullary tonicity will gradually reestablish itself, and as it does, more pronounced responses to successive administered dDAVP doses will occur, thereby confirming the diagnosis.

Since patients with DI have intact thirst mechanisms, most often they do not present with hyperosmolality, but rather with a normal $P_{\text{osm}}$ and $[\text{Na}^+]$ and symptoms of polyuria and polydipsia. In these cases it is most appropriate to perform a water deprivation test (Table 3). This entails following the patient’s serum $[\text{Na}^+]$, urine volume, and $U_{\text{osm}}$ in the absence of fluid intake until the serum $[\text{Na}^+]$
is 146 mEq/L or the U_{osm} reaches a plateau (generally defined as 3 successive urines with less than 10% differences in osmolality from the preceding sample) and the patient has lost at least 2% of body weight. At this point, a plasma AVP level is drawn and the patient is given AVP or dDAVP (as discussed above for hyperosmolar patients). The same criteria are used to evaluate the etiology of the DI following this test, but one additional entity, primary polydipsia, must be considered in the differential diagnosis of normonatremic polyuria and polydipsia (Table 2). Primary polydipsia is usually a result of psychiatric disease. Such patients ingest large amounts of fluids for a variety of reasons, but generally not because of physiological sensations of thirst; this is referred to as psychogenic polydipsia. A smaller subset of patients with primary polydipsia have a true disorder of thirst regulation, usually manifested by a downward resetting of the osmotic threshold for stimulated thirst; this is sometimes called dipsogenic dia-

Table 3

<table>
<thead>
<tr>
<th>Procedure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Initiation of the deprivation period depends on the severity of the DI; in routine cases, the patient should be made to have nothing by mouth (NPO) after dinner, while in cases with more severe polyuria and polydipsia, this may be too long a period without fluids and the water deprivation should be begun early in the morning of the test (e.g., 6 AM).</td>
</tr>
<tr>
<td>2. Stop the test when body weight decreases by 3%, the patient develops orthostatic blood pressure changes, the U_{osm} reaches a plateau (i.e., less than 10% change over 3 consecutive measurements), or the serum [Na⁺] is &gt;145 mmol/L.</td>
</tr>
<tr>
<td>3. Obtain a plasma AVP level at the end of the test when the P_{osm} is elevated, preferably above 300 mOsm/kg H₂O.</td>
</tr>
<tr>
<td>4. If the serum [Na⁺] is &lt;146 mmol/L or the P_{osm} is &lt;300 mOsm/kg H₂O, then consider infusion of hypertonic saline (3% NaCl at a rate of 0.1 mL/kg/min for 1 to 2 h) to reach these endpoints.</td>
</tr>
<tr>
<td>5. Administer AVP (5 U) or dDAVP (1 µg) sc and continue following U_{osm} and volume for an additional 2 h.</td>
</tr>
</tbody>
</table>

**Interpretation**

1. An unequivocal urine concentration after AVP/dDAVP (>50% increase) indicates neurogenic DI and an unequivocal absence of urine concentration (<10%) strongly suggests NDI or primary polydipsia (PP). |
2. Differentiating between NDI and PP, as well as for cases in which the increase in U_{osm} after AVP administration is more equivocal (e.g., 10–50%) is best done using the plasma AVP levels obtained at the end of the dehydration period and/or hypertonic saline infusion and the relation between pAVP levels and U_{osm} under basal conditions. |
betes insipidus (36). Regardless of the cause of the excessive fluid intake, because the ensuing water diuresis can wash out the medullary concentration gradient and down-regulate kidney aquaporin-2 water channels, such patients may concentrate their urine subnormally in response to water deprivation and therefore, resemble partial central DI. In contrast to central DI, however, patients with primary polydipsia will generally concentrate their urine <10% in response to administered AVP or dDAVP and will have plasma AVP levels appropriate to their $P_{\text{osm}}$. With use of the water deprivation test combined with plasma AVP determinations, greater than 95% of all cases of polyuria and polydipsia can be diagnosed appropriately; diagnoses in the remaining patients will generally become evident over time based on their responses to therapeutic clinical trials.

**Osmoreceptor Dysfunction**

There is an extensive literature in animals indicating that the primary osmoreceptors that control AVP secretion and thirst are located in the anterior hypothalamus. Lesions of this region in animals cause hyperosmolality through a combination of impaired thirst and osmotically stimulated AVP secretion (37, 38). Initial reports in humans described this syndrome as “essential hypernatremia,” and subsequent studies used the term “adipsic hypernatremia” in recognition of the profound thirst deficits found in most of the patients. Rather than focus on semantic issues, it makes more sense to group all of these syndromes as disorders of osmoreceptor function. Four major patterns of osmoreceptor dysfunction have been described as characterized by defects in thirst and/or AVP secretory responses: (i) upward resetting of the osmostat for both thirst and AVP secretion (normal AVP and thirst responses but at an abnormally high $P_{\text{osm}}$); (ii) partial osmoreceptor destruction (blunted AVP and thirst responses at all $P_{\text{osm}}$s); (iii) total osmoreceptor destruction (absent AVP secretion and thirst regardless of $P_{\text{osm}}$); and (iv) selective dysfunction of thirst osmoregulation with intact AVP secretion (39). Most of the cases reported to date have represented various degrees of osmoreceptor destruction associated with different brain lesions. As opposed to lesions causing central DI, these lesions usually occur more anteriorly in the hypothalamus, consistent with the anterior hypothalamic location of the primary osmoreceptor cells (12). Whether some of these patients also have an inability to suppress as well as stimulate AVP secretion, thereby leading to hypovolemia and hypotension in some situations, remains an interesting but incompletely evaluated possibility. For all cases of osmoreceptor dysfunction, it is important to remember that afferent pathways from the brainstem to the hypothalamus remain intact; therefore, these patients will usually have normal AVP and renal concentrating responses to baroreceptor-mediated stimuli, such as hypovolemia and hypotension (39). This often causes confusion, since at some times these patients appear to have DI, and yet at other times they can concentrate their urine quite normally.
Clinical Manifestations

The clinical manifestations of any disease entity are important aids to the differential diagnosis, and therefore, it is essential that the clinician be aware of the clinical manifestations of hyperosmolar patients. These can be divided into the signs and symptoms produced by dehydration, which are largely cardiovascular, those caused by the hyperosmolality itself, which are largely neurological and reflect brain dehydration as a result of osmotic water shifts out of the central nervous system, and those that are secondary to excessive renal water losses in patients with DI. Cardiovascular manifestations of hypertonic dehydration include hypotension, azotemia, acute tubular necrosis secondary to renal hypoperfusion or rhabdomyolysis, and shock. Neurological manifestations range from nonspecific symptoms, such as irritability and decreased sensorium, to more severe manifestations, such as chorea, seizures, coma, focal neurological deficits, and cerebral infarction. The severity of symptoms can be roughly correlated with the degree of hyperosmolality, but individual variability is marked and for any single patient, the level of serum [Na+] at which symptoms will appear cannot be predicted. Similar to hypoosmolar syndromes, the length of time over which hyperosmolality develops can markedly affect clinical symptomatology. Rapid development of severe hyperosmolality is frequently associated with marked neurologic symptoms, whereas gradual development over several days or weeks generally causes milder symptoms. In this case, the brain counteracts osmotic shrinkage by increasing intracellular content of solutes. These include electrolytes such as potassium and a variety of organic osmolytes which previously had been called “idiogenic osmoles” (for the most part these are the same organic osmolytes that are lost from the brain during adaptation to hypoosmolality) [40]. The net effect of this process is to protect the brain against excessive shrinkage during sustained hypertonicity. However, once the brain has adapted by increasing its solute content, rapid correction of the hyperosmolality can cause brain edema, since it takes a finite time (24–48 h in animal studies) to dissipate the accumulated solutes, and until this process has been completed, the brain will accumulate excess water as P osm is normalized. This effect is most often seen in dehydrated pediatric patients, who can develop seizures with rapid rehydration, but has been described only rarely in adults, including the most severely hyperosmolar patients with nonketotic hyperglycemic hyperosmolar coma.

The characteristic symptoms of DI are the polyuria and polydipsia that result from the underlying impairment of urinary concentrating mechanisms. Interestingly, patients with DI typically describe a craving for cold water, which seems to quench their thirst better. Patients with central DI also typically describe a precipitous onset of their polyuria and polydipsia, which simply reflects the fact that urinary concentration can be maintained fairly well until the number of
AVP-producing neurons in the hypothalamus decreases to 10–15% of normal, after which plasma AVP levels decrease to the range where urine output increases dramatically (see Fig. 3) (15).

HYPOOSMOLALITY AND HYPONATREMIA

Pathogenesis

Hypoosmolality indicates excess water relative to solute in the ECF; because water moves freely between ECF and ICF, this also indicates an excess of TBW relative to total body solute. Imbalances between body water and solute can be generated either by depletion of body solute more than body water, or by dilution of body solute from increases in body water more than body solute (Table 4) (3). This represents an oversimplification, because most hypoosmolar states include components of both solute depletion and water retention (e.g., isotonic solute losses, as occurs during an acute hemorrhage, do not produce hypoosmolality until subsequent retention of ingested or infused hypotonic fluids causes a secondary dilution of the remaining ECF solute). Nonetheless, this concept has proven to be useful because it provides a simple framework for understanding the diagnosis and therapy of hypoosmolar disorders.

Differential Diagnosis

Evaluation of hypoosmolar patients should include a careful history (especially concerning medications), clinical assessment of ECF volume, thorough neurological evaluation, serum electrolytes, glucose, uric acid, BUN, and creatinine, calculated and/or directly measured P_{osm}, and simultaneous urine electrolytes and osmolality (28). Hyponatremia and hypoosmolality are usually synonymous, with two exceptions. First, pseudohyponatremia can be produced by marked elevation of serum lipids and/or proteins; although the [Na+]_{L/P} plasma water is unchanged, the [Na+]_{L/P} plasma is decreased because of the increased nonaqueous portion of the plasma occupied by lipid or protein. However, the directly measured P_{osm} is not affected by increased lipids or proteins. Second, high concentrations of effective solutes other than Na⁺, e.g., glucose, cause relative decreases in serum [Na⁺] despite an unchanged P_{osm}. Misdiagnosis can be avoided again by direct measurement of P_{osm}, or in the case of hyperglycemia by correcting the serum [Na⁺] by 1.6 mEq/L for each 100 mg/dL increase in the plasma glucose concentration above 100 mg/dL, which provides an estimate of the contribution of the glucose to the P_{osm}. Definitive identification of the etiology of the hypoosmolality is not always possible at the time of presentation, but categorization according to the patient’s ECF volume status will allow determination of an appropriate initial therapy in the majority of cases (Fig. 6).
Table 4
Pathogenesis of Hypoosmolar Disorders

Solute depletion
(primary decreases in total body solute plus secondary water retention^a)

1. Renal solute loss
   - Diuretic use.
   - Solute diuresis (glucose, mannitol).
   - Salt wasting nephropathy.
   - Mineralocorticoid deficiency.

2. Nonrenal solute loss
   - Gastrointestinal (diarrhea, vomiting, pancreatitis, bowel obstruction).
   - Cutaneous (sweating, burns).
   - Blood loss.

Solute dilution
(primary increases in total body water ± secondary solute depletion^a)

1. Impaired renal free water excretion
   A. Increased proximal nephron reabsorption
      - Congestive heart failure.
      - Cirrhosis.
      - Nephrotic syndrome.
   B. Impaired distal nephron dilution
      - SIADH.
      - Glucocorticoid deficiency.

2. Excess water intake
   - Primary polydipsia.

^aVirtually all disorders of solute depletion are accompanied by some degree of secondary retention of water by the kidneys in response to the resulting intravascular hypovolemia; this mechanism can lead to hypoosmolality even when the solute depletion occurs via hypotonic or isotonic body fluid losses. Disorders of water retention can cause hypoosmolality in the absence of any solute losses, but often some secondary solute losses occur in response to the resulting intravascular hypervolemia, and this can then further aggravate the dilutional hypoosmolality.

DECREASED ECF VOLUME

Clinically detectable hypovolemia indicates some degree of solute depletion. Elevation of BUN is a useful laboratory correlate of decreased ECF volume. Even isotonic or hypotonic fluid losses can cause hypoosmolality if water or hypotonic fluids are subsequently ingested or infused. A low urine [Na+] (UNa) suggests a nonrenal cause of solute depletion, whereas a high UNa suggests renal causes of solute depletion (Table 4). Diuretic use is the most common cause of hypovolemic hypoosmolality; thiazides are more commonly associated with severe hyponatremia than are loop diuretics such as furosemide (41). Although
in the center emphasizes that the presence of CNS dysfunction due to hyponatremia should always be assessed immediately, so that appropriate therapy can be started as soon as possible in symptomatic patients while the outlined diagnostic evaluation is proceeding. Abbreviations: N, no; Y, yes; ECF, extracellular fluid volume; NSS, normal (isotonic) saline; Rx, treat; 1°, primary; 2°, secondary; P-osmol, P-osmol, d/c, discontinue; SIADH, syndrome of inappropriate antidiuretic hormone secretion; numbers referring to osmolality are in mOsm/kg H2O, numbers referring to [Na+] are in mEq/L (modified with permission from ref. 28).

Fig. 6. Schematic summary of the evaluation of hypotonic patients. This seemingly is a simple example of solute depletion, the pathophysiological mechanisms underlying the hyponatremia are complex and include multiple components. Many such patients do not present with clinical evidence of hypovolemia, in part because ingested water has been retained in response to nonosmotically stimulated vasopressin (AVP) secretion, which occurs in all disorders of solute depletion as an attempt to maintain volume homeostasis. In addition, urine [Na+] may be high or low depending on when the last diuretic dose was taken. Consequently, any suspicion of diuretic use mandates careful consideration of this diagnosis regardless of clinical or laboratory findings. Most other etiologies of solute losses causing hypovolemic hyponatremia will be clinically apparent, although some salt-wasting nephropathies (chronic interstitial nephropathy, polycystic kidney disease, obstructive uropathy, or Bartter’s syndrome) or mineralocorticoid deficiency (Addison’s disease) may be challenging to diagnose during early phases of these diseases.
NORMAL ECF VOLUME

Virtually any disorder causing hypoosmolality can present with a volume status that appears normal by standard methods of clinical evaluation. Because clinical assessment of volume status is not very sensitive, the presence of normal or low BUN and uric acid concentrations are helpful laboratory correlates of relatively normal ECF volume. In these cases, a low UNa (<30 mEq/L) suggests depletional hypoosmolality secondary to ECF losses with subsequent volume replacement by water or other hypotonic fluids (42); as discussed earlier, such patients may appear euvolemic by the usual clinical parameters used to assess hydration status. Hypoosmolar disorders caused primarily by dilution (Table 4) are less likely with a low UNa, although this can occur in hypothyroidism or in the syndrome of inappropriate antidiuretic hormone secretion (SIADH) with superimposed volume depletion. A high UNa (>30 mEq/L) generally indicates a dilutional hypoosmolality such as SIADH, which is the most common cause of euvolemic hypoosmolality. The clinical criteria necessary for a diagnosis of SIADH remain as defined by Bartter and Schwartz (Table 5) (43). First, ECF hypoosmolality must be present and hyponatremia secondary to pseudo-hyponatremia or hyperglycemia excluded. Second, U_osm must be inappropriate for plasma hypoosmolality; this simply requires that the urine be less than maximally dilute (i.e., U_osm >100 mOsm/kg H2O). Furthermore, U_osm need not be inappropriately elevated at all levels of P_osm, but simply at some level below 275 mOsm/kg H2O. In patients with a reset osmostat, AVP secretion is suppressed at some lower level of P_osm, resulting in maximal urinary dilution and free water excretion at P_osms below this level. Third, clinical euvolemia must be present; this does not mean that patients with SIADH cannot become hypovolemic for other reasons, but in such cases, it is impossible to make a diagnosis of SIADH until the patient is made euvolemic. The fourth criterion, an elevated UNa, has caused much confusion; its importance lies in differentiating hypoosmolality caused by decreased relative intravascular volume, in which renal Na+ conservation occurs, from dilutional disorders, in which urinary Na+ excretion is normal or increased due to ECF volume expansion. The continued excretion of ingested Na+ by such patients reflects the importance of mechanisms for volume homeostasis, which in this case override osmotic homeostatic mechanisms that would favor Na+ conservation. However, UNa can also be high in renal causes of solute depletion, such as diuretic use or Addison’s disease, and conversely, patients with SIADH can have a low urinary Na+ excretion if they subsequently become hypovolemic or solute depleted. Consequently, although elevated urinary Na+ excretion is the rule in most patients with SIADH, its presence does not confirm this diagnosis nor does its absence rule it out. Finally, SIADH is a diagnosis of exclusion, and other potential causes of hypoosmolality must always be excluded (Fig. 6). Glucocorticoid deficiency and SIADH can be especially difficult to distinguish,
since hypocortisolism can cause elevated plasma AVP levels and impair maximal urinary dilution (44); no patient should be diagnosed as having SIADH without an evaluation of adrenal function, preferably via a rapid ACTH stimulation test. Although additional testing is generally not necessary to establish a diagnosis of SIADH, abnormal excretion (<80%) of a standard water load (20 mL/kg body weight) within 4 h can be helpful in confirming the diagnosis in difficult cases (45). However, water loading should be avoided in patients with more severe hypoosmolality (serum [Na⁺] <125 mEq/L), since abnormal retention of the ingested water can cause a significant (4–6 mEq/L) further decrease in plasma [Na⁺]. Many different disorders have been associated with SIADH; these can be divided into four major etiologic groups: tumors, central nervous system (CNS) disorders, drug effects, and pulmonary diseases (Table 6).

**INCREASED ECF VOLUME**

Clinically detectable hypervolemia indicates whole body sodium excess, and hypoosmolality in these patients suggests a relatively decreased intravascular volume and/or pressure leading to water retention as a result of elevated plasma AVP levels and decreased distal delivery of glomerular filtrate (46,47). Such patients usually have a low U₉₅, because of secondary hyperaldosteronism, but under certain conditions, the U₉₅ may be elevated (e.g., diuretic therapy).

### Table 5

**Criteria for the Diagnosis of SIADH**

**Essential**

1. Decreased effective osmolality of the ECF (Pₑₒᵢᵣ <275 mOsm/kg H₂O).
2. Inappropriate urinary concentration (Uₒᵢᵣ > 100 mOsm/kg H₂O with normal renal function) at some level of hypoosmolality.
3. Clinical euvoolemia, as defined by the absence of signs of hypovolemia (orthostasis, tachycardia, decreased skin turgor, dry mucous membranes) or hypervolemia (subcutaneous edema, ascites).
4. Elevated urinary sodium excretion while on a normal salt and water intake.
5. Absence of other potential causes of euvolemic hypoosmolality: hypothyroidism, hypocortisolism (Addison’s disease or pituitary ACTH insufficiency) and diuretic use.

**Supplemental**

6. Abnormal water load test (inability to excrete at least 80% of a 20 mL/kg water load in 4 h and/or failure to dilute Uₒᵢᵣ to <100 mOsm/kg H₂O).
7. Plasma AVP level inappropriately elevated relative to Pₑₒᵢᵣ.
8. No significant correction of serum [Na⁺] with volume expansion but improvement after fluid restriction.

since hypocortisolism can cause elevated plasma AVP levels and impair maximal urinary dilution (44); no patient should be diagnosed as having SIADH without an evaluation of adrenal function, preferably via a rapid ACTH stimulation test. Although additional testing is generally not necessary to establish a diagnosis of SIADH, abnormal excretion (<80%) of a standard water load (20 mL/kg body weight) within 4 h can be helpful in confirming the diagnosis in difficult cases (45). However, water loading should be avoided in patients with more severe hypoosmolality (serum [Na⁺] <125 mEq/L), since abnormal retention of the ingested water can cause a significant (4–6 mEq/L) further decrease in plasma [Na⁺]. Many different disorders have been associated with SIADH; these can be divided into four major etiologic groups: tumors, central nervous system (CNS) disorders, drug effects, and pulmonary diseases (Table 6).
Verbalis

Hyponatremia generally does not occur until advanced stages of congestive heart failure, cirrhosis, or nephrotic syndrome, so diagnosis is usually not difficult. Renal failure can also cause retention of both sodium and water, but in this case, the factor limiting excretion of excess body fluid is not decreased effective circulating volume but rather decreased glomerular filtration.

Although primary polydipsia can sometimes cause hypoosmolality, especially if renal free water excretion is impaired, these patients rarely, if ever, manifest signs of hypervolemia, since water retention alone without sodium excess does not cause significant volume expansion.

**Clinical Manifestations**

The clinical manifestations of hyponatremia are largely neurological, and primarily reflect brain edema resulting from osmotic water shifts into the brain (5,48). These range from nonspecific symptoms, such as headache and confusion, to more severe manifestations, such as decreased sensorium, coma, seizures, and death. Significant CNS symptoms generally do not occur until serum [Na+] falls below 125 mEq/L, and the severity of symptoms can be roughly

**Table 6**

**Common Etiologies of SIADH**

<table>
<thead>
<tr>
<th>Tumors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pulmonary/mediastinal (bronchogenic carcinoma, mesothelioma, thymoma).</td>
</tr>
<tr>
<td>Nonchess (duodenal carcinoma, pancreatic carcinoma, ureteral/prostate carcinoma, uterine carcinoma, nasopharyngeal carcinoma, leukemia).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CNS disorders</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass lesions (tumors, brain abscesses, subdural hematoma).</td>
</tr>
<tr>
<td>Inflammatory diseases (encephalitis, meningitis, systemic lupus).</td>
</tr>
<tr>
<td>Degenerative/demyelinating diseases (Guillain-Barré, spinal cord lesions).</td>
</tr>
<tr>
<td>Miscellaneous (subarachnoid hemorrhage, head trauma, acute psychosis, delirium tremens, pituitary stalk section).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Drug induced</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulated AVP release (nicotine, phenothiazines, tricyclics).</td>
</tr>
<tr>
<td>Direct renal effects and/or potentiation of AVP effects (dDAVP, oxytocin, prostaglandin synthesis inhibitors).</td>
</tr>
<tr>
<td>Mixed or uncertain actions (chlorpropamide, clofibrate; carbamazepine, cyclophosphamide, vincristine).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pulmonary diseases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infections (tuberculosis, aspergillosis, pneumonia, empyema).</td>
</tr>
<tr>
<td>Mechanical/ventilatory (acute respiratory failure, chronic obstructive pulmonary disease (COPD), positive pressure ventilation).</td>
</tr>
</tbody>
</table>

**Clinical Manifestations**

The clinical manifestations of hyponatremia are largely neurological, and primarily reflect brain edema resulting from osmotic water shifts into the brain (5,48). These range from nonspecific symptoms, such as headache and confusion, to more severe manifestations, such as decreased sensorium, coma, seizures, and death. Significant CNS symptoms generally do not occur until serum [Na+] falls below 125 mEq/L, and the severity of symptoms can be roughly
correlated with the degree of hypoosmolality (49). Individual variability is marked, and for any patient the level of serum [Na+] at which symptoms will appear cannot be predicted. Several factors other than the severity of the hypoosmolality also affect the degree of neurological dysfunction. The most important is the time-course over which hypoosmolality develops. Rapid development of severe hypoosmolality frequently causes marked neurologic symptoms, whereas gradual development over several days or weeks is often associated with relatively mild symptomatology despite profound degrees of hypoosmolality. This is because the brain counteracts osmotic swelling by extruding extracellular and intracellular solutes (including potassium and organic osmolytes) (50). Since this is a time-dependent process, rapid development of hypoosmolality can result in brain edema before this adaptation occurs, but with slower development of the same degree of hypoosmolality brain cells can lose solute sufficiently rapidly to prevent cell swelling, brain edema, and neurological dysfunction (3). Underlying neurological disease also affects the level of hypoosmolality at which CNS symptoms appear; moderate hypoosmolality is of little concern in an otherwise healthy patient, but can cause morbidity in a patient with an underlying seizure disorder. Non-neurological metabolic disorders (hypoxia, hypercapnia, acidosis, hypercalcemia, etc.) can similarly affect the level of P_{osm} at which CNS symptoms occur (51). Recent clinical studies have suggested that menstruating females and young children may be particularly susceptible to the development of neurological morbidity and mortality during hyponatremia, especially in the acute postoperative setting (52,53). The true clinical incidence, as well as the underlying mechanisms responsible for these sometimes catastrophic cases, remains to be determined.

REFERENCES
