Current concepts

Osmosis, osmometry, and osmoregulation

R C C Lord

Summary

The maintenance of adequate body fluid volume and the correct distribution of this fluid between the body compartments is a critical part of homeostasis. The process of osmosis plays an important role in movement of fluid within the body and the use of osmometry is an important part of the management of many patients. In addition to the application of osmometry to the measurement of body fluids, most commonly plasma and urine, osmotic action plays a part in some therapeutic actions of drugs and its strength needs to be quantified in fluids administered to patients. Unfortunately confusion often exists in the various terms that are used in the field of osmometry. This review aims to explain the different terms used, the laboratory methodology involved in osmometry, and the clinical application and interpretation of the results obtained.

Keywords: homeostasis; osmolality; osmolarity; colligative properties

Osmosis

If two aqueous solutions with different concentrations of particles are separated from each other by a semipermeable membrane, then water will move across the membrane from the solution with the lower concentration to the solution with the higher concentration. The movement of the water will depend on the difference in the concentration of the particles and the nature of permeability of the membrane. This movement of water is termed osmosis and the pressure which would need to be exerted to halt its movement is called the osmotic pressure. Consider an aqueous solution of sucrose contained within a sac made of a strictly semipermeable membrane attached to narrow diameter glass tube. If this is placed into a beaker of water, water will move across the membrane into the sac of sucrose. The sucrose solution will thus rise up the glass tube. At equilibrium the gravitational pressure of this column of solution equals the osmotic pressure and so will prevent further net movement of water from the beaker.

It is important to realise that the osmotic pressure is determined by the total number of particles in solution, regardless of molecular nature. The total number of particles will thus depend on the degree of dissociation of solutes. If the sucrose in the preceding experiment were replaced by an aqueous solution of sodium chloride of the same molarity, the solution would reach a height in the glass tube almost twice as high as the sucrose. This is because sodium chloride dissociates into two ions per molecule. In reality, this dissociation is actually incomplete and in addition there is association between particles in solution. To take account of this, a term called the osmotic coefficient is used to correct for the deviation from the ‘ideal’ behaviour of the system.

Osmometry

Osmometry is a technique for measuring the concentration of particles in a solution, ie, the osmolar concentration. Osmolar concentration can be expressed in two ways:

- osmolality expressed as mmol/kg of solvent
- osmolarity expressed as mmol/l of solution

Osmolality is a thermodynamically more precise expression because solution concentrations expressed on a weight basis are temperature independent while those based on volume will vary with temperature in a manner dependent on the thermal expansion of the solution.

If a solute is dissolved in a solvent then the following properties of the solvent change:

- osmotic pressure increases
- vapour pressure decreases
- boiling point increases
- freezing point decreases.

These are known as colligative properties and are all directly related to the total number of solute particles per mass of solvent, ie, the osmolality. Theoretically, any of the four colligative properties could be used as a basis for the measurement of osmolality. The most commonly used method in the case of physiological fluids is freezing point depression.

FREEZING POINT DEPRESSION OSMOMETER

The sample is supercooled within a cooling fluid or using solid state cooling. A rapid stir mechanism is used to initiate crystallisation. The temperature will then rise due to latent heat of crystallisation. A thermistor (temperature-dependent resistor) reading is noted. Comparison with standards allows calculation of the osmolality.
VAPOUR PRESSURE OSMOMETER
A decrease in the vapour pressure of the solution results in a decrease in the \textit{dew point}. The dew point is the temperature at which the saturation vapour pressure is equal to the actual vapour pressure of the contained water vapour.

A filter paper soaked in the sample is placed in a chamber which is then sealed. Once thermal and vapour equilibrium has been reached, a thermocouple (a device that measures the potential difference between two dissimilar metals) is cooled below the dew point. Water condenses on the thermocouple, releasing heat of condensation. The temperature of the thermocouple is thus raised until it reaches the dew point, at which temperature no more water condenses. The thermocouple reading is noted and compared with standards to calculate the osmolality.

Accurate measurement requires the sample chamber and thermocouple to be very clean. However, this method is less precise than that of freezing point depression (coefficients of variation are more than double, 2.5% compared to 1%), and the method cannot be used in the presence of volatile solutes such as ethanol. 

Tonicity
Although the terms tonicity and osmolality are often used interchangeably, there is a clear distinction. Osmolality is a physical property dependent on the total number of solute particles present in a solution whereas tonicity is a physiological process dependent upon the selectively permeable characteristics of a membrane. For example, solutes such as urea and ethanol permeate cells freely and therefore will have no effect on tonicity but will increase the measured osmolality.

Specific gravity
Specific gravity is defined as the density of a solution relative to the density of water. Before osmolality determinations became feasible as a routine laboratory procedure, the specific gravity of urine was used as an indirect index of its osmolality. The specific gravity of urine was used as a guide to the adequacy of the renal concentrating mechanism.

In most circumstances the specific gravity bears a constant relationship to osmolality. However, specific gravity is dependent on the mass concentration of urinary solutes while osmolality depends on molecular concentration. The relationship between the two is thus influenced by the composition of the solutes in the urine. A specific gravity of 1.010 corresponds to an osmolality of 320 mOsmol/kg, but this may vary over a range of 126 to 520 mOsmol/kg, depending on dietary intake. In particular, correction should be made for urinary glucose and protein. Radio-opaque contrast media in the urine can result in a grossly elevated specific gravity. Refrigeration of the sample may cause a moderate increase in specific gravity.

Specific gravity can be measured using a variety of techniques:

**HYDROMETER**
Hydrometers are devices for measuring the density of liquids by the buoyancy of a plummet with a calibrated stem. The urinometer is a hydrometer adapted to measure the specific gravity of urine. The urinometer is calibrated in specific gravity units using standard solutions. Prior to placement of the urinometer in the urine it needs to be cleaned and then rinsed in a small amount of the urine before being inserted into the urine with a spinning motion to ensure that it is floating freely. The fluid level is read at the bottom of the meniscus at eye level with the urinometer in a vertical position. Corrections for temperature, glucose concentration and protein concentration are required. Each 10 g/l of glucose increases the specific gravity by 0.004 units and each 4 g/l of protein increases it by 0.001 units. Automated methods based on the hydrometer principle are also available.

**FALLING DROP**
A drop of urine is dropped into a column of an organic solvent the specific gravity of which is known. Measurement of the speed at which the drop of urine falls in the column of organic solvent can be used to calculate the specific gravity. This method has been automated and can be quite precise.

**VIBRATING CAPILLARY**
This technique is based on the principle that the frequency of vibration of sound is related to the density of the medium through which it has to travel. Measurement of shifts in harmonic oscillation on addition of a urine sample to the meter can be used to calculate the specific gravity of the urine.
REFRACTOMETRY

Refractometers detect the deviation of light by a solution as a measure of specific gravity.\(^1\)

REAGENT STRIP

The strip consists of a polymer with repeating carboxylic acid groups.\(^4\) The dissociation of these groups is influenced by the ionic strength of the medium, so that when the strip is dipped into the urine, there is a release of protons that decreases the pH of the strip. The change in pH is detected with a coloured indicator (e.g., bromthymol blue) and the colour noted against a chart calibrated with urines of known specific gravities. The measurement of specific gravity using the strip is rapid and convenient but unpredictable in alkaline urine.\(^13\)\(^16\)

Given the problems associated with the measurement of specific gravity it is recommended that osmolality should be measured whenever possible.

Colloid osmotic pressure

Colloid is a term used to describe solute particles with a molecular weight greater than 30000. Colloid osmotic pressure, or onotic pressure, describes an equilibrium pressure measurement when two solutions, one of which contains colloid, are separated by a semipermeable membrane.\(^17\)\(^18\) Interest in its measurement has come from studies in critical care medicine in the prediction of intercompartmental body water movements, in particular as a useful prognostic indicator of pulmonary oedema, and of mortality in the critically ill.\(^19\)\(^21\)

Measurement of the colloid osmotic pressure of plasma or serum can be carried out in an osmometer consisting of a sample cell separated from a reference cell by a synthetic membrane, simulating the vascular membrane. The patient's serum is introduced into the sample cell. The reference cell contains isotonic saline to produce approximately the same e\(_{\text{V}}\)ct as would interstitial fluid. There is a net migration of water molecules and diffusible solute ions from the reference cell into the sample cell. The resultant negative pressure in the reference cell is sensed by a pressure transducer and the signal converted into a pressure reading once equilibrium has been attained.

Physiology of water balance

In normal humans the osmolality of body fluids is tightly regulated.\(^22\) Normal serum osmolality lies between 285 and 290 mOsmol/kg. This narrow range is maintained in the face of a wide variation in fluid input and loss. This is accomplished through the increased or decreased excretion of water by the kidneys, regulated by the action of vasopressin (antidiuretic hormone), and the development of the sensation of thirst to prevent excessive hypertonicity.

Vasopressin is derived from a precursor molecule synthesised in the supra-optic and paraventricular nuclei of the hypothalamus. The precursor is packaged into neurosecretory granules and then transported along axons to the posterior pituitary gland. Along the way the precursor is cleaved into three peptides, one of which is vasopressin.

The main stimulus to the secretion of vasopressin is an increase in osmolality detected by osmoreceptors lying close to the nuclei in the hypothalamus. The osmoreceptors are extremely sensitive to changes in osmolality, a 1% change in osmolality will result in a measurable release of vasopressin.\(^23\) At a plasma osmolality of 280 mOsmol/kg vasopressin secretion is suppressed. Above this there is a linear increase in vasopressin secretion. Volume status is less important in the regulation of vasopressin release, although the presence of marked hypovolaemia can result in substantial release of vasopressin.

Vasopressin acts in the kidneys by binding to V2 receptors in the collecting ducts. The binding of vasopressin to its receptors promotes the insertion of vesicles containing aquaporin 2 (a water channel protein) into the luminal membrane, thus increasing the permeability of the membrane to water.\(^24\) Efflux of water at the basolateral membrane appears to be mediated by aquaporin 3. Hence an increase in osmolality sensed in the brain results in increased vasopressin release, which causes increased water reabsorption by the kidneys and a lowering of plasma osmolality.
Clinical applications

SERUM/PLASMA OSMOLALITY

Hyperosmolar state
In hyperosmolar non-ketotic coma, glycaemic decompensation, usually over a period of several days, results in hyperglycaemia. When the plasma glucose exceeds the renal threshold glycosuria occurs. This produces an osmotic diuresis, and if fluid intake is not maintained the fluid loss will result in a hyperosmolar state. Movement of water out of the brain tissue due to an osmotic gradient may contribute to eventual coma. This condition is associated with considerable mortality.

Calculated osmolarity
The osmolarity of plasma is almost entirely due to sodium and its counter anions. Potassium, glucose and urea also make a small contribution. It is therefore possible to calculate what the osmolarity of plasma is from the measurement of these constituents. 25–27 In clinical practice, probably the easiest to remember is:

Calculated osmolarity (mOsm/l) = 2(Na + K) + urea + glucose (mmol/l)

Osmolal gap
The usefulness of deriving a calculated osmolarity lies in the concept of the osmolal gap. 28 29 This is obtained by subtracting the calculated osmolarity from the measured osmolality:

Osmolal gap = Measured osmolality − Calculated osmolarity

This is obtained by subtracting the calculated osmolarity from the measured osmolality:

Osmolal gap = Measured osmolality − Calculated osmolarity

A ‘normal’ gap is less than 2 units if the equation given above for calculated osmolarity is used. The presence of a positive osmolal gap can occur in several situations. A spuriously low sodium measurement can arise in the presence of hyperlipidaemia or hyperproteinaemia if specimen dilution is used prior to analysis (for example, using flame photometry or an indirect ion-selective electrode technique) because of lipid or protein making up a large part of the plasma volume. This pseudohyponatraemia will give rise to a falsely low calculated osmolarity and hence an osmolal gap.

The finding of an osmolal gap is useful in toxicology for the detection of low molecular weight substances in the plasma. The most common example of this is ethanol, but other substances such as methanol, ethylene glycol and isopropanol need also to be considered (box 1). It must be emphasised that the presence of an osmolal gap should only be used as a clinical guide and an analytical procedure should than be used to formally identify the substance responsible. 31 The osmolal gap may also arise as a result of the production of endogenous substances such as ketoacids and organic acids, occurring in some disease states, although the elevation is usually mild. 32

URINE OSMOLALITY
The collection and analysis of urine for biochemical analysis can be a very useful technique in the evaluation of the renal handling of water. In healthy adults the urine osmolality varies from about 50 to around 1400 mOsmol/kg depending on the state of hydration of the subject. Thus to interpret a urine osmolality value one needs to also measure the plasma osmolality and see whether the urine osmolality is appropriate.

Diabetes insipidus
Diabetes insipidus is the inability to conserve water due to:

• failure of secretion of vasopression in response to an osmotic stimulus (cranial)
• failure of the kidneys to respond to circulating vasopression (nephrogenic)

Measurement of both the plasma and urine osmolality reveals that the urine osmolality is inappropriately low. Further investigation is then carried out to confirm the diagnosis. 33 34

Water deprivation test
The patient is denied fluid for a period of time, usually 8 hours. The normal response to fluid deprivation is the concentration of urine. If the urine osmolality fails to rise above 300 mOsmol/kg then the patient probably has diabetes insipidus. The patient is then given desmopressin (DDAVP), a synthetic analogue of vasopressin. If the urine osmolality remains below 300

<table>
<thead>
<tr>
<th>Causes of an increased serum osmolal gap</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exogenous causes</strong></td>
</tr>
<tr>
<td>• ethanol</td>
</tr>
<tr>
<td>• methanol</td>
</tr>
<tr>
<td>• ethylene glycol</td>
</tr>
<tr>
<td>• isopropanol</td>
</tr>
<tr>
<td>• acetone</td>
</tr>
<tr>
<td>• mannitol</td>
</tr>
<tr>
<td>• glycine</td>
</tr>
<tr>
<td><strong>Endogenous causes</strong></td>
</tr>
<tr>
<td>• diabetic ketoacidosis</td>
</tr>
<tr>
<td>• alcoholic ketoacidosis</td>
</tr>
<tr>
<td>• chronic renal failure</td>
</tr>
<tr>
<td><strong>Artifactual</strong></td>
</tr>
<tr>
<td>• hyperlipidaemia*</td>
</tr>
<tr>
<td>• hyperproteinaemia*</td>
</tr>
<tr>
<td>* when measured with a technique</td>
</tr>
<tr>
<td>requiring sample dilution</td>
</tr>
</tbody>
</table>

Box 1
mOsmol/kg then the type of disorder is nephrogenic, whereas if the osmolality rises to greater than 750 mOsmol/kg the type is cranial. In practice, the results are often found not to be as clear cut as this, since polyuria may result in the washout of renal interstitial solute thus decreasing the kidneys’ concentrating ability.

**Hypertonic saline infusion test**

The serum osmolality needs to be above 295 mOsmol/kg for an adequate osmotic stimulus. If the serum osmolality does not exceed 295 mOsmol/kg on fluid deprivation, then careful infusion of 5% saline over 2 hours can be carried out as an osmotic stimulus. At 2 hours plasma is taken for osmolality and vasopressin measurement and urine for osmolality. Plotting the results on a nomogram is carried out to decide whether the disorder is of the cranial or nephrogenic type.35

**Syndrome of inappropriate antidiuresis (SIADH)**

This syndrome is defined as the continued, and therefore inappropriate, secretion of antidiuretic hormone when the plasma osmolality is low and other stimuli to its secretion, such as hypovolaemia and drugs, have been excluded and renal and adrenal function are normal.36–38

The urine osmolality is thus inappropriately elevated (an osmolality which is greater than the maximally dilute urine of 50–100 mOsmol/kg). Tumours are a common cause (box 2). The management of this hyponatraemic state remains controversial, in particular the rate at which the biochemical correction needs to be made.39 40 Treatment includes fluid restriction, 3% saline with frusemide and demeclocycline.

**Free water clearance**

Free water clearance is used to express quantitatively the renal tubular handling of water. It represents the difference between the actual volume of urine produced per unit time and the volume required to excrete that amount of solute iso-osmotically with plasma. It can be calculated using the equation:

\[
\text{Free water clearance} = R \left(1 - \frac{U_{\text{osm}}}{P_{\text{osm}}}\right)
\]

where \(R\) = urine flow rate, \(U_{\text{osm}}\) = urine osmolality, and \(P_{\text{osm}}\) = plasma osmolality.

The concept is seldom used in routine clinical settings but can be a sensitive method to study water metabolism in nephrology research.41

**FAECES**

Patients with Munchausen’s syndrome sometimes add water to the stool to simulate diarrhoea. This will result in a low stool osmolality. However, the intestine cannot secrete free water, so that stool should never be hypo-osmolar.42

The faecal osmotic gap has been proposed as a simple means of differentiating between secretory (small gap) and osmotic (large gap) diarrhoea.43 It is sometimes requested as a preliminary step in the differential diagnosis of chronic diarrhoea of unknown cause.

Although the concept is relatively simple, the following points need to be borne in mind in the collection of the faecal sample, analysis, calculation, and interpretation of results.44 Stool osmolality (which should normally approximate the corresponding plasma osmolality) increases with standing at room temperature (over 120 mOsmol/kg in 24 h) due to the metabolism of carbohydrate by faecal bacteria.45 This is reduced, but not prevented, by refrigeration. The electrolytes, sodium and potassium, remain the same. Faecal samples should therefore be analysed as soon as possible for osmolality measurement. Severe dehydration can also lead to a high stool osmolality. Thus, an elevated stool osmolality in the absence of an increased plasma osmolality indicates improper stool storage.46

The measurement of faecal osmolality should be carried out using the method of freezing point depression. The dew point technique underestimates the result by as much as 50 mOsmol/kg, since volatile substances are not detected. Furthermore, since the degree of underestimation is variable, this method introduces an unpredictable error.47

Two methods for the calculation of the osmotic gap in faeces have been described48:

- Faecal osmotic gap = faecal osmolality \(-\) 2(Na + K)\(_f\)
- Faecal osmotic gap = plasma osmolality \(-\) 2(Na + K)\(_p\)
The intestinal contents are iso-osmolar with plasma as they leave the ileum. Due to the rapid transit through the colon in diarrhoea, little change in osmolality of the faeces can take place, thus the osmolality of faecal fluid as it exits the rectum will be close to $\sim 290 \text{ mOsmol/kg}$. The authors recommend that, given the difficulty in obtaining a reliable faecal osmolality measurement, the second equation should be used, as it gave better discrimination between the two types of diarrhoea and, in osmotic diarrhoea, a better correlation with severity of diarrhoea.

In interpreting the results there is the question as to what cut-off value for the osmotic gap should be used to distinguish between secretory and osmotic diarrhoea. In the literature, values in the range 40–100 units have been used.

Although the distinction between secretory and osmotic diarrhoea is excellent in concept, it fails to take into account that many diarrhoeal disorders have more than one pathophysiological mechanism. In coeliac sprue at least three different mechanisms coexist to account for the observed diarrhoea — two are secretory and one osmotic in nature. Sodium sulphate produces diarrhoea by an osmotic action. The poorly absorbed sulphate anion is the main driving force for the diarrhoea. However, cations need to remain in the gut lumen to maintain electrical balance and hence the osmotic gap is small. Given that many studies have found a considerable overlap between the two groups, the use of the faecal osmotic gap to distinguish between osmotic and secretory diarrhoea should be viewed with caution. In patients with osmotic diarrhoea, fasting should eliminate the underlying offending agent and the diarrhoea. This clinical clue to the nature of the diarrhoea is a useful simple initial step in the investigation of the patient. If laxative abuse is suspected then it is probably more useful to carry out a toxicology screen on urine.

 Patients with cystic fibrosis have greater sodium and chloride concentrations in their sweat than usual. This phenomenon has been used as the basis of the ‘sweat test’ to diagnose cystic fibrosis. Some authors have proposed the use of sweat osmolality as an alternative to electrolyte assay. The chlorides of sodium and potassium account for about 80% of the osmolality of sweat, the remainder being made up of mostly urea and lactic acid. Sweat osmolality correlates well with sweat sodium measurement. The small sample volume required for the osmolality measurement means that the process of sweat collection need only take 10 minutes, making it an attractive procedure for use with small children.

 Osmotic studies have been carried out on other body fluids including gastric juice, saliva, bile, cerebrospinal fluid, synovial fluid, pleural fluid, ascitic fluid and pericardial fluid, but measurement provides little useful clinical information.

 Osmotic strength largely determines the physiologic acceptability of many solutions used for therapeutic and nutritional purposes. Direct therapeutic effects of osmotic action include intravenous mannitol, used to draw water away from the brain in cerebral oedema, and lactulose, a semi-synthetic disaccharide used as a laxative. The high osmolality of solutions used in parenteral nutrition necessitates their infusion into a large central vein. The greater blood flow compared to peripheral veins is necessary to dilute the solutions and so avoid the development of thrombophlebitis.

 Parenteral nutrition using a peripheral vein has been advocated for patients who only require short-term parenteral nutrition, because it avoids many of the complications of central venous catheterization. It also offers cost-saving benefits. Some guidelines suggest that the osmolality of peripherally infused solutions should be kept below 1000 mOsmol/kg (a figure that means that some patients would be unable to meet their nutritional requirements). However, a recent study using fine-bore polyurethane catheters to administer the solutions into a peripheral vein showed that 1700 mOsmol/kg feeding could be used without increasing the risk of thrombophlebitis.

 Conclusion

Measurement of osmolality is simple and rapid with the instruments available today. It is most frequently of use for analysing plasma and urine in assessing the state of water balance within a patient. In addition, calculation of an osmolality from other measurements and the concept of an osmotic gap are useful in the detection of substances, in plasma and sometimes in faeces. The clinical utility of its measurement in other body fluids is limited.