

Insulin: understanding its action in health and disease

P. Sonksen^{1*} and J. Sonksen²

¹GKT Medical School, King's College London, St Thomas' Hospital, London SE1 7EH, UK. ²Russells Hall Hospital, Dudley, West Midlands DY1 2HQ, UK

*Corresponding author

Br J Anaesth 2000; 85: 69–79

Keywords: metabolism, diabetes; complications; metabolism, insulin; hormones; anaesthesia; surgery

The results of pancreas extirpation and pancreas grafting are best explained by supposing that the islet tissue produce an Autacoid which passes into the blood stream and effects carbohydrate metabolism and carbohydrate storage in such a manner that there is no undue accumulation of glucose in the blood. Provisionally it will be convenient to refer to this hypothetical substance as insuline.

Sir Edward Schafer, *The Endocrine Organs* (1916)

Historical background

Sir Edward Schafer, who was Professor of Physiology in Edinburgh, appears to have named insulin and described its actions. He did so in a book, *The Endocrine Organs*, based on a lecture series he gave in California in 1913. In this book, published in 1916,⁵ he gave the hypothetical substance a name that stuck; what is more, with remarkable vision, he described its likely formation from activation of an inert precursor:

It must however be stated that it has yet to be determined whether the active substance is produced as such in the pancreas or whether it exists there as pro-insuline which becomes elsewhere converted into the active autacoid.

Insulin was discovered 8 yr later by Banting and Best in 1921. The first patient was treated a year later in 1922 and pro-insulin was discovered (and re-named) more than 50 yr later by George Steiner of the University of Chicago in 1967.

Schafer deliberately avoided using the word 'hormone' and used his preferred terms 'autacoid' and 'chalone'. This was as a result of long-standing academic rivalry with his contemporaries Professors Baylis and Starling at University College, London. They had previously described secretin as the first hormone to be isolated and characterized. They had

coined the term 'hormones' to describe the class of substance produced in one part of the body and acting elsewhere. Schafer preferred his own terms, which were based on terms used at the time to describe actions of drugs, 'autacoid' being a substance with excitatory action and 'chalone' one with inhibitory action.

Schafer went on to describe how 'insuline' had both excitatory and inhibitory actions. His description of how he thought the hypothetical substance 'insuline' acted in the body is remarkable because the passage of time has shown him to be correct almost word for word. Things have been confused, however, by a 20 yr 'black age' of endocrinology (between approximately 1960 and 1980), where leading scientists—through extrapolating beyond their new discoveries—confused scientific thinking and teaching. They formulated new hypotheses based for the first time on hard scientific evidence but they got it badly wrong through extrapolating (incorrectly) from *in vitro* experimental data in rat tissues to *in vivo* metabolism in humans.

The effects of this 'black age' are still with us because these incorrect hypotheses have, with the passage of time, been turned into dogma and become cast into 'tablets of stone' in undergraduate textbooks. They are also carried forward into postgraduate teaching. For example, even in well respected texts it is still common to find statements such as 'The basic action of insulin is to facilitate glucose entry into cells, primarily skeletal muscle and hepatocytes.'¹

Purpose of this review

The main thrust of this article is to set the record straight, provide up-to-date information about current understanding and confirm that Schafer, through careful thought and despite lack of experimental data, was correct. We also hope to show how this 'new' concept of insulin action affects the modern management of diabetes during surgery and in the intensive care unit.

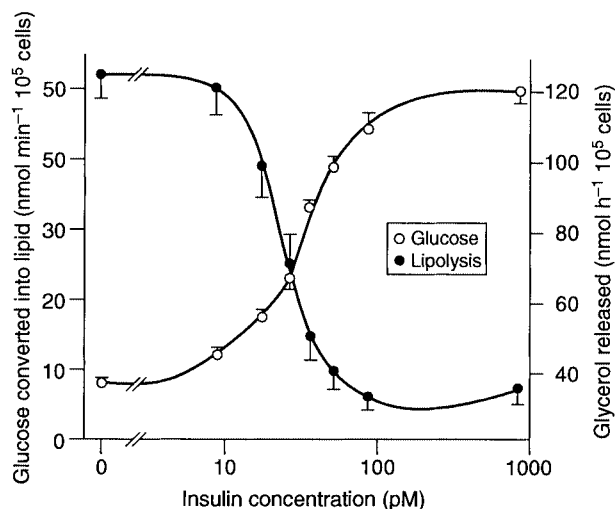


Fig 1 Insulin action on isolated rat adipocytes. Insulin simultaneously stimulates lipogenesis from glucose and inhibits lipolysis. Both mechanisms are mediated by the same receptor. The data illustrate what Schafer referred to as the 'autacoid' (excitatory) action, i.e. stimulation of lipogenesis, and the 'chalone' (inhibitory) action, i.e. inhibition of lipolysis, both of which are integral to the hormone and mediated by the same receptor. It is the 'chalone' action which has most relevance to human physiology and the use of insulin therapeutically.

Insulin as an autacoid and chalone

As Schafer hypothesized, insulin has two classes of action: (i) excitatory (autacoid), for example stimulating glucose uptake and lipid synthesis; and (ii) inhibitory (chalone), for example inhibiting lipolysis, proteolysis, glycogenolysis, gluconeogenesis and ketogenesis. These dual actions are illustrated in Fig. 1, where insulin's well recognized action in stimulating lipogenesis from glucose (in rat adipocytes) is illustrated.⁶ What perhaps is less well known is insulin's simultaneous inhibitory action on lipolysis. It is quite clear from Fig. 1 that both these actions occur simultaneously over the same concentration range. They are both mediated through the same cell membrane receptor. These actions of insulin *in vitro* were discovered in the late 1950s when it was also shown that insulin stimulated glucose uptake by rat muscle. It was extrapolation of this last observation in rat muscle to explain the pathophysiology of diabetes that was erroneous. The consequence of this error was the (fallacious) concept of insulin being 'required' for glucose entry into cells rather than just accelerating glucose uptake. The hyperglycaemia of diabetes was interpreted as a 'damming back' of glucose in the blood stream as a consequence of a lack of insulin. This became established teaching and, although the concept was shown to be erroneous in the mid-1970s, the teaching has not changed. Consequently, therapy has been based on a flawed concept.

It is now well established that what Schafer called insulin's 'chalone' (or inhibitory) actions are the physiologically more important. Indeed, its autacoid (excitatory)

action has recently been shown to be, on the whole, physiologically unimportant. Insulin's excitatory actions are effectively abolished in the genetically engineered muscle insulin receptor knock-out (MIRKO) mouse because they have a complete lack of the insulin-sensitive glucose transport protein, Glut 4, in muscle, so insulin has no stimulatory effect on glucose uptake. Fascinatingly, this is of little or no consequence to the animal that does not have fasting hyperglycaemia or diabetes.

The nature of fasting hyperglycaemia

The key question is: is the fasting hyperglycaemia of diabetes due to overproduction of glucose by the liver or underutilization of glucose by peripheral tissues?

When this question was posed to participants of the Royal College of Anaesthetists' Basic Science Course in 1998 and 1999, a show of hands revealed virtually unanimous support for the second alternative, with only one or two hands in favour of the correct answer (overproduction of glucose by the liver). This is a typical response, seen from the overwhelming majority of both undergraduate and post-graduate audiences in the UK and the rest of the world. It is important to have the right understanding if insulin is to be used appropriately and to maximum benefit. This may perhaps seem a trivial or pedantic point but it is not. Correct understanding of what goes wrong with metabolism in the face of insulin deficiency is essential in the understanding and execution of logical and safe forms of treatment.

Fasting normoglycaemia

Consider a person with a normal fasting blood glucose concentration of 5 mmol litre⁻¹. The stable fasting blood glucose concentration conceals an underlying dynamic steady state where the glucose that is continuously being cleared from the circulation for metabolism is exactly matched by new glucose being added by the liver. This dynamic steady state is normally controlled very tightly by insulin. Glucose input into the circulation (rate of appearance, or R_a) is normally approximately 12 $\mu\text{mol kg}^{-1} \text{min}^{-1}$ (2 mg kg⁻¹ min⁻¹), and in the fasting state is exclusively from the liver, exactly matches glucose disposal in peripheral tissues (rate of disappearance, or R_d). The main contributors to glucose disappearance (or clearance) in the fasting, resting state are the brain (~50%), lean tissues (~30%), adipose tissue (~10%) and red blood cells (~10%). This is illustrated in Fig. 2.

Fasting hyperglycaemia

Next consider a hypothetical patient with type 2 diabetes presenting for elective surgery. Fasting blood glucose in this case is, let us say, 12 mmol litre⁻¹ (which remains quite steady with short-term fasting). This again reflects a balance between hepatic glucose production on one hand and

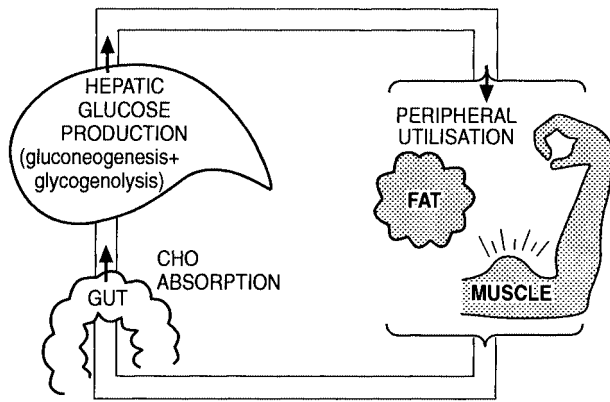


Fig 2 A cartoon illustrating the dynamic steady state that exists between glucose production and utilization. The brain is not shown. Blood glucose concentration is tightly controlled during fasting and feeding through mechanisms balancing exactly the supply and demand of energy substrate. Insulin is the prime controller of this process. When fasting hyperglycaemia develops, as in diabetes mellitus, it could result from either over-production of glucose, through excess glycogenolysis and gluconeogenesis, or a failure of adequate glucose uptake in the peripheral tissues (including the brain). Most people think that hyperglycaemia reflects a reduction in glucose uptake as a result of insulin deficiency. This is not true.

peripheral glucose utilization on the other. In this case, however, because blood glucose concentration is above the renal threshold for glucose, the rate of glucose disposal (R_d) is a combination of the rate of tissue utilization (R_{ut}) and the rate of urinary loss (R_u). The two always balance to produce a steady state because glycosuria (R_u) increases rapidly and linearly with the increase in blood glucose concentration once the renal threshold has been crossed. Thus, R_d increases until it matches R_a , when a new steady state is achieved. For this reason it is the extent of the hepatic overproduction of glucose (R_a) that sets the degree of fasting hyperglycaemia; the higher the fasting hyperglycaemia, the higher R_a and the more severe the diabetes.

It is now easy to measure R_a and R_d experimentally and thus answer the question posed at the beginning of this section. Initially, this was done by injecting radioactive (^{14}C - or ^3H -labelled) glucose and measuring the rate at which it was diluted with unlabelled glucose from the liver (the concentration of glucose remains constant but the concentration of ^{14}C - or ^3H -labelled glucose falls exponentially with time). It is quite easy to calculate accurately the actual rates of glucose production and utilization from this isotope dilution curve. It is now possible to make these measurements without the need for radioactive tracers, using instead glucose labelled with a natural stable isotope of carbon (^{13}C), thus avoiding any radiation hazard and even allowing research on diabetes in pregnancy.

In the fasting state, virtually all glucose R_a comes from the liver, from a combination of glycogenolysis and gluconeogenesis. When hepatic glycogen stores are plentiful, most glucose comes from glycogenolysis, but as the

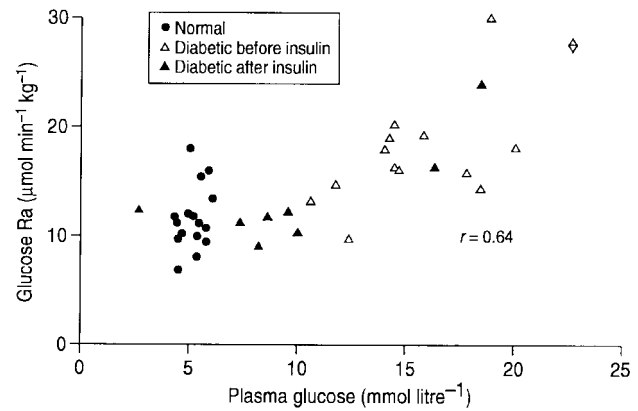


Fig 3 Fasting plasma glucose and isotope-determined glucose turnover rate in a group of newly diagnosed patients with diabetes studied before and after starting insulin compared with non-diabetic controls. In the newly diagnosed patients, glucose turnover (rate of production and rate of utilization) is elevated in proportion to fasting blood glucose. This shows that the fasting hyperglycaemia of diabetes results from hepatic over-production of glucose alone, since peripheral glucose utilization is increased despite the lack of insulin. Insulin treatment reduces glucose concentration through inhibiting hepatic glucose production. Under these conditions glucose utilization decreases, thus insulin administration *reduces* glucose utilization. This indicates that the plasma glucose concentration, rather than plasma insulin, is the prime determinant of glucose uptake. There are clearly sufficient glucose transporters present, even in the newly diagnosed diabetic state, to ensure adequate glucose metabolism. Thus insulin regulates glucose production more than glucose utilization.

hepatic glycogen stores become depleted, more comes from gluconeogenesis. After 24 h of starvation, hepatic glycogen stores are completely exhausted and all glucose production comes from gluconeogenesis (muscle glycogen cannot contribute to blood glucose as there is no glucose-6-phosphatase in muscle and glucose transporters cannot transport glucose 6-phosphate out of muscle. The kidney may, on occasions, make a minor contribution to gluconeogenesis.)

Experimental observations in people with newly diagnosed diabetes or in insulin-withdrawn patients with both type 1 and type 2 diabetes have shown that there is a strongly positive relationship between the fasting blood glucose concentration and the rate of hepatic glucose production (R_a). This was first shown in the early 1970s in animals and in the mid-1970s in humans.³ It is illustrated in Fig. 3.

The normal fasting–feeding cycle (‘energy metabolism’)

The role of glucose

Normally, fasting blood glucose concentration is very tightly controlled by insulin secretion at a mean of about 5 mmol litre⁻¹. There is a much wider range of glucose production rates (R_a) and glucose utilization rates (R_d)

between normal individuals than there is in fasting blood glucose concentrations (see Fig. 3). This is because of the feedback control loop between blood glucose concentration and insulin secretion. The rates of fasting hepatic glucose production and glucose utilization (R_a and R_d) in each individual remain exactly matched and relatively constant over days, months and even years and are probably largely genetically determined.

The fact that fasting blood glucose concentration remains constant over minutes and hours reflects an exact balance between input and output of glucose from the circulation ($R_a = R_d$). In non-diabetics, all of R_d is going to metabolism, so the hormonal control mechanisms are geared to releasing sufficient glucose to match the body's need. This is a very basic and robust process, so that a healthy individual can survive a remarkably long time without eating, so long as they have access to water. Hypoglycaemia is not a feature of prolonged starvation; there are powerful 'survival' values in being able to maintain brain function during prolonged famine. Since hepatic glycogen stores are depleted within 24 h of fasting, blood glucose concentrations are maintained thereafter entirely through gluconeogenesis. Gluconeogenesis is mainly dependent on protein breakdown (a small amount comes from the glycerol released during lipolysis) and it thus results in protein wasting. It is the effects of protein malnutrition that lead to the eventual lack of ability to cough properly and keep the airways clear, in turn leading to pneumonia and death during prolonged starvation; hypoglycaemia does not occur. During fasting, mobilization of fat and the consumption of fat and ketones by some tissues as 'alternative substrates' ensures that protein loss is minimized and used to provide glucose for tissues, such as the brain, that cannot survive without it.

With normal feeding, digestion of carbohydrates leads to a rise in blood glucose, which normally triggers a brisk insulin response of sufficient magnitude to control blood glucose concentration very tightly in the 4–10 mmol litre⁻¹ range. The *average* peak plasma insulin response to a high carbohydrate meal in a group of normal healthy slim individuals rises 10-fold from a mean fasting value of 5 mU litre⁻¹ to a peak of about 50 mU litre⁻¹. The *range* of insulin response in this group is enormous, however, with peak values ranging between 15 and 250 mU litre⁻¹ while glucose values are no different between individuals. This reflects a genetically determined variation in insulin sensitivity between individuals.

Some glucose is stored in muscle as glycogen. This compartment of stored glucose is entirely for the 'local use' of muscle and is integral to the 'flight or fight' survival mechanism that has resulted in humans being the dominant creatures on earth. Muscle lacks a certain phosphorylase enzyme that removes the phosphate group from glucose 6-phosphate and as a result, free intracellular glucose is not produced by glycogenolysis in muscle (unlike the liver) and muscle cannot therefore 'export' glucose into the circulation. Muscle glycogen allows a sudden burst of fuel (through

anaerobic metabolism of glycogen to glucose to lactate) which allows muscle to undertake an immediate sudden burst of extreme activity on locally available energy stores before blood flow builds up to carry the lactate away and bring other energy fuels to sustain more prolonged exercise. Thus the fuel used by muscle progresses from local glycogen (sprint) to glucose and fatty acids brought via the circulation (middle distance) to predominantly fat (marathon).

The roles of fat and protein

The other major fuel required for 'energy metabolism' is fat. Stored fat is in the form of triglyceride and contains enormous amounts of stored energy. Fat contains 9 kcal g⁻¹ stored energy while glycogen contains only 2 kcal g⁻¹ because there is water of hydration in glycogen in both muscle and the liver. This means that glycogen stores are heavy and of limited value to mobile creatures like us. They are there for short-term needs but, when food is scarce, a transition occurs to substitution of fat in all tissues that can metabolize it. Hepatic glycogen stores are soon lost during fasting, and gluconeogenesis—mainly from protein—is then required to maintain glucose supplies to the central nervous system. Since protein stores are also limited and critical to survival, the substitution of fat wherever possible is of great value. The enormous amount of energy stored in fat is liberated through lipolysis where triglyceride is hydrolysed to its constituents, glycerol and non-esterified fatty acids (NEFA; also known as free fatty acids (FFA)). Fat is able to substitute for glucose as an energy substrate in many but not all tissues. In muscle, there is no real preference for glucose or NEFA and the tissue can and does readily change from one to the other. The brain and other neural tissues have an obligatory requirement for glucose and cannot survive without it; however, they can substitute ketones (but not NEFA) for some, but not all, of their energy needs.

Protein is used mainly for structural processes but is continuously being synthesized and broken down. When food is short and glycogen stores are depleted, protein becomes the major source of gluconeogenic precursors via the 'glucogenic' amino acids released during proteolysis. Proteolysis is inhibited by insulin (another of Schafer's 'chalone' actions) and, in states of relative insulin lack such as prolonged fasting, insulin concentrations fall low enough to allow release of sufficient gluconeogenic amino acids to maintain blood glucose in the normal range. The price for this is a gradual but progressive loss of structural protein. As stated earlier, it is this that determines survival during prolonged fasts.

What goes wrong in diabetes?

Fasting hyperglycaemia is the hallmark of diabetes. In the very early stages, so-called impaired glucose tolerance (IGT), fasting blood glucose concentration is normal and only the response to a glucose challenge is impaired. In

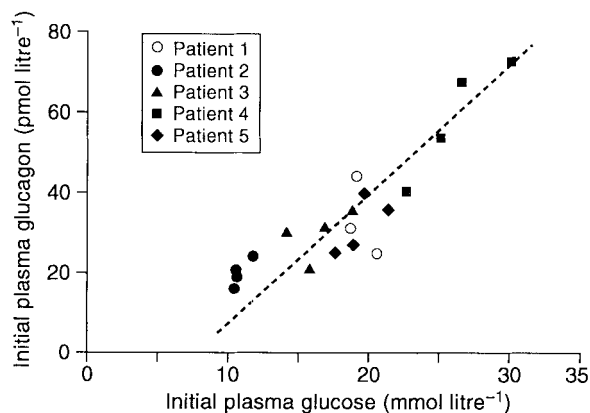


Fig 4 Fasting plasma glucose concentration is determined not only by fasting plasma insulin but also by glucagon. These data were obtained from a group of five type 1 diabetic patients who withheld insulin for 24 h on four separate occasions. There was a very close correlation between fasting blood glucose and plasma glucagon; also, on each occasion the same individual achieved a remarkably similar fasting plasma glucose and glucagon concentration.

symptomatic diabetes, fasting hyperglycaemia is always present and random blood glucose concentrations are always in excess of 11 mmol litre⁻¹, often considerably so. For the classical symptoms of thirst, polyuria and weight loss to occur, there must be heavy glycosuria. By and large, the severity of the metabolic disturbance is reflected accurately in the degree of elevation of the fasting blood glucose concentration (which reflects the magnitude of the hepatic overproduction of glucose). This is illustrated in Fig. 3.

Even when the fasting blood glucose concentration is as high as 20 mmol litre⁻¹, it remains constant over several hours of fasting. During this time, a dynamic steady state exists when glucose production matches glucose disappearance from the bloodstream exactly. The magnitude of the hyperglycaemia is determined by the absolute rate of hepatic glucose production (Fig. 3) and this in turn is determined by the extent of, on one hand, the insulin deficiency and, on the other hand, the magnitude of the glucagon rise in response to insulin deficiency. This is another key point. In patients with diabetes mellitus secondary to pancreatectomy, where glucagon concentrations are low, fasting hyperglycaemia is modest. In type 1 diabetes, glucagon production is unimpaired and, in the face of insulin deficiency, plasma concentrations rise rapidly to high concentrations. In a group of insulin-withdrawn type 1 patients, where no insulin had been given for 24 h, fasting plasma glucose (and hence hepatic glucose production) correlated remarkably closely with fasting plasma glucagon² as shown in Fig. 4.

Glucose uptake into cells is usually normal and often high in untreated diabetes

Since there is a dynamic steady state during short-term fasting, glucose production exactly matches glucose clear-

ance from the plasma. In the face of hyperglycaemia, glucose clearance represents both tissue glucose metabolism and urinary glucose loss. By collecting urine and measuring the losses directly, tissue glucose metabolism can be calculated readily. Such calculations show that, in the face of hyperglycaemia, tissue glucose uptake is usually increased above normal even when insulin deficiency is severe. This cannot be reconciled with the concept that insulin is required for glucose uptake by insulin-sensitive tissues. Indeed it proves beyond question that insulin is *not* required. We now know the detailed mechanisms involved and can explain this. Glucose uptake by all cells is by means of a specific transport protein (glucose transporter) of which at least six isomers (Glut 1 to Glut 6) are known. Glucose is a highly polar substance, being freely soluble in water but insoluble in fat. It cannot enter cells except through the specific transport system utilizing Glut 1–6. Glut 4 is the transport protein present in muscle and adipose tissue, which is known to be ‘insulin sensitive’. This means that, in addition to the transporters resident in the cell membrane at any given moment, there is a pool of glucose transporter molecules in the cytoplasm of the cell which can be recruited in response to a rise in plasma insulin, to join those already in the cell membrane in the fasting state. We now know, from experiments like those illustrated in Fig. 3, that even in the fasting state or in a state of absolute insulin deficiency, there are sufficient glucose transporters already in place in the cell membrane to allow glucose uptake to exceed that of a normal individual when the gradient of glucose concentration across the cell membrane is sufficiently high.

This ‘mass action’ effect accounts for the observations which show unequivocally that tissue glucose uptake can exceed normal even in the face of severe insulin deficiency such as occurs in uncontrolled diabetes mellitus.

Ketosis inhibits glucose uptake in many tissues and thus can ‘spare’ structural protein

Glucose uptake is not, however, increased in all diabetic patients with hyperglycaemia. When significant ketosis occurs, this can inhibit glucose metabolism in cells, because the ketone bodies (3-hydroxybutyrate, acetoacetate and acetone) are freely soluble in both water and fat and diffuse across the cell membrane in direct proportion to their concentration in plasma. They are metabolized intracellularly to acetoacetate, which enters the Krebs’ cycle and becomes oxidized to carbon dioxide and water. When ketone concentrations in the plasma are high (above about 15 mmol litre⁻¹), their metabolism can account for most of the energy needs of the Krebs’ cycle. This results in a physiological ‘damming back’ of glycolysis from glucose. As a result, the concentration of the intermediate compounds in the glycolytic pathway increases until the tissue concentration of glucose 6-phosphate is so high that further phosphorylation of glucose that has been transported into

BIOCHEMICAL CONSEQUENCES OF INSULIN DEFICIENCY

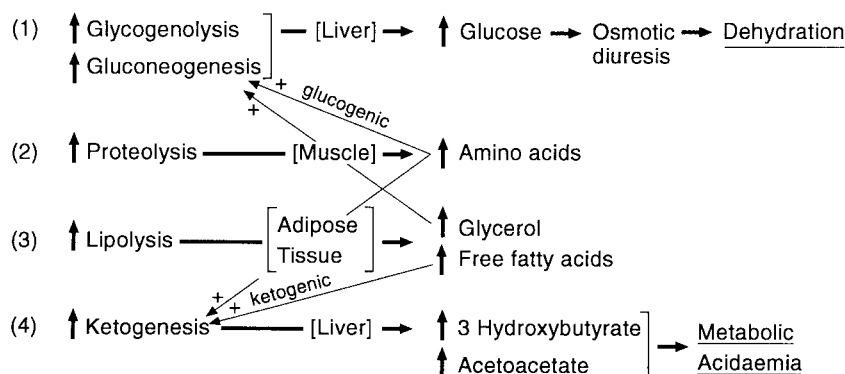


Fig 5 The biochemical consequences of insulin deficiency are summarized in this illustration. Insulin normally has a tonic inhibitory control (what Schafer termed 'chalone' action) on the supply of energy substrates arising mainly from metabolism in the liver, muscle and adipose tissue. Normally insulin ensures sufficient release of appropriate substrate to allow humans to cope with the eventualities of life in a highly controlled manner. Once insulin deficiency develops this control is lost, and energy substrates are over-produced and flood the system. The metabolic consequences result from the excess of substrates not (as is often misconceived) by a lack of energy substrate getting to the tissues.

the cell becomes inhibited. Even when this stage is reached, glucose transport into the cell continues but, rather than being phosphorylated and passing down the glycolytic pathway and oxidized in the Krebs' cycle, glucose is transported back out of the cell ('futile cycling'). The glucose transporter is specific for glucose and cannot transport glucose 6-phosphate.

Thus glucose transport and uptake into cells are dependent on the degree of ketosis. In the face of extreme ketosis, such as that not uncommonly seen in type 1 diabetes, where concentrations rise to as high as 25 mmol litre⁻¹, glucose uptake can be reduced to almost zero in tissues, such as muscle, that can metabolize ketones as readily as glucose. Under these conditions, glucose uptake in brain and other neural tissues (which can metabolize limited amounts of ketones) is reduced but ketones can never substitute completely for glucose. Red blood cells, on the other hand, do not have nuclei or mitochondria, cannot metabolize ketones at all and rely entirely on glucose conversion to lactate for their energy needs.

Ketoacidosis in diabetes is an uncontrolled process where serious lack of insulin 'takes the brake off' lipolysis and ketogenesis. The result is severe ketosis, which can readily become ketoacidosis and can easily be fatal once ketoacids reach concentrations above 20 mmol litre⁻¹ and compensatory mechanisms become exhausted.

Controlled ketosis occurs with prolonged fasting in non-diabetics. Insulin levels fall sufficiently to release enough NEFA and to activate the hepatic enzymes of ketogenesis sufficiently to allow ketoacids to reach concentrations of <7 mmol litre⁻¹. This is a 70-fold increase over their normal fasting concentration (around 0.1 mmol litre⁻¹) and sufficient to supply energy for those tissues able to substitute ketones for glucose. In muscle this is as high as 100%, in red blood cells it is 0 and in the central nervous system it can be

as high as about 50%. Thus the need for glucose is reduced and structural protein is preserved.

Mechanisms leading to the development of diabetic ketoacidosis and hyperosmolar non-ketotic diabetic coma

Schafer's concept of insulin's main action being as a chalone has been shown to be correct. Thus insulin normally 'keeps the brakes on' a number of key metabolic processes including the supply of all 'fuels' needed for metabolism from glycogenolysis, gluconeogenesis, lipolysis, ketogenesis and proteolysis. In the presence of severe insulin deficiency, such as occurs in diabetic ketoacidosis, the 'brake' is removed and the net result is a sort of metabolic mayhem. In keeping with the basis of 'Murphy's law' (when one thing goes wrong, so does every thing else); insulin deficiency itself leads to over-secretion of the 'anti-insulin' hormones glucagon, cortisol, growth hormone and catecholamines. This in turn aggravates the metabolic effects of insulin lack. This is summarized in Fig. 5.

Glycogenolysis and gluconeogenesis are activated as a result of lack of insulin in the liver; the rates of these processes are further enhanced by over-production of hormones such as glucagon and cortisol, which stimulate these processes, particularly in the face of insulin deficiency. The net result is that the liver pours glucose into the circulation much more quickly than the tissues can metabolize it. This leads to a progressive increase in blood glucose concentration. When this exceeds the renal threshold, glycosuria develops and fasting blood glucose eventually reaches a dynamic steady state where over-production of glucose by the liver is matched by increase clearance through metabolism in the tissues (R_{ut}) and

excretion in the urine (R_u). Under these steady-state conditions, $R_a = R_d = R_{ut} + R_u$.

Proteolysis is enhanced as a result of lack of insulin in muscle, providing a supply of glucogenic and ketogenic amino acids. The glucogenic amino acids (alanine, arginine, etc.) pass to the liver to fuel the furnace of gluconeogenesis. In addition, lack of insulin at the liver activates the enzymes of gluconeogenesis. The 'ketogenic' amino acids (leucine, isoleucine and valine) released from proteolysis contribute to the substrates for hepatic ketogenesis.

As a result of lack of insulin in adipose tissue, lipolysis proceeds at an accelerated rate and this is further enhanced by over-secretion of growth hormone and catecholamines. Stored triglyceride is hydrolysed to glycerol and NEFA. The glycerol passes to the liver and contributes significantly to gluconeogenesis while the NEFA are transported to the liver and provide the major substrate for ketogenesis. Plasma concentrations of NEFA rarely increase above $1.5 \text{ mmol litre}^{-1}$ since NEFA are insoluble in water (highly non-polar) and rely on albumin binding for transport in the extracellular fluids. At $1.5 \text{ mmol litre}^{-1}$, all protein binding is saturated. NEFA are metabolized very quickly with a plasma half-life of 1–2 min, thus flux through the blood can be very high without the concentration rising above $1.5 \text{ mmol litre}^{-1}$.

As a result of insulin lack, the liver is also 'primed' for ketogenesis as insulin deficiency activates the appropriate enzymes of ketogenesis. With increased flux of NEFA substrate arriving at the liver from accelerated lipolysis, the net result is that the liver pours out ketoacids (3-hydroxybutyric and acetoacetic acids) into the blood stream, where R_a exceeds R_d , so their concentrations in the blood increase progressively. In this case, although ketones are excreted in the urine, there is no 'renal threshold' and unlike glucose the urinary losses are not very great and do not lead to the development of a new steady-state blood ketone concentration. The only significant loss of ketones from the circulation is, therefore, from metabolism. Since ketones are freely soluble in water and lipid, they enter all cells at high concentrations and are metabolized in preference to glucose in cells capable of doing so. If unchecked, ketone concentrations increase progressively and lead to the development of a progressively serious metabolic acidosis. If unchecked, ketoacids accumulate and exceed the body's compensatory mechanisms (buffering with bicarbonate and hyperventilation) and H^+ concentration rises to the point where blood pH drops to around 6.7–6.8 (when total acids reach about $25 \text{ mmol litre}^{-1}$). Beyond this, the condition is rapidly fatal. Of considerable diagnostic importance is the fact that the ratio of the concentrations of 3-hydroxybutyric acid:acetoacetic acid is determined by the H^+ concentration such that the acetoacetic acid concentration decreases as blood pH decreases. Since the conventional bedside tests for ketones (Ketostix or Acetest tablets) react only with acetoacetic acid and require $>1\text{--}2 \text{ mmol litre}^{-1}$ acetoacetic acid to trigger a positive test, it is possible to have a patient with severe ketoacidosis with a plasma 3-hydroxybutyrate concentra-

tion as high as $20 \text{ mmol litre}^{-1}$ and a urine ketone test showing only a trace of ketones (from a blood aceto-acetate concentration of $2 \text{ mmol litre}^{-1}$). This may often be ignored as irrelevant and the patient incorrectly assumed to have lactic acidosis; this can be a fatal mistake as the treatment of lactic acidosis is quite different from that of ketoacidosis.

Thus, there are two parallel metabolic processes that go wrong in the face of insulin deficiency: (i) glucose over-production and (ii) ketone over-production. Usually, both go wrong together and this is what one sees when insulin is withdrawn from a longstanding type 1 diabetic patient. Hyperglycaemia and ketoacidosis develop simultaneously and in parallel. Occasionally one pathway dominates. The commonest clinical scenario is hyperosmolar non-ketotic diabetic coma (HONK). This happens when insulin supplies are not quite zero and there is sufficient present to inhibit lipolysis but not gluconeogenesis. There is often an aggravating factor present: the patient may turn to a 'high-energy' drink to quench the thirst arising from the osmotic diuresis of the underlying type 2 diabetes. In the early stages, the patients may feel unwell and lacking in energy and, as a result of television advertising, try high-glucose drinks in the hope that they will make them feel more energetic. This just augments the already elevated rate of glucose entry into the circulation, aggravates osmotic diuresis and accelerates the process of decompensation. The very high glucose concentrations may themselves also help suppress lipolysis and ketogenesis by encouraging re-esterification of NEFA released from fat.

A rarer occurrence is 'normoglycaemic ketotacidosis'. This arises usually in an established diabetic patient who becomes unwell and cuts down food intake, monitors his/her own blood glucose and takes extra insulin to prevent hyperglycaemia. Occasionally the drive of fat-mobilizing hormones, such as catecholamines and growth hormone, break through the inhibition of insulin on lipolysis and ketosis results. However, this is very uncommon.

The development of electrolyte disturbances

During the hyperglycaemic phase when the hyperglycaemia-induced osmotic diuresis leads to thirst and polyuria, the choice of fluid that people make to slake their thirst varies enormously and often has a profound effect on outcome. We have already covered the scenario where a patient turns to a sweet beverage; in the case of another person who chooses a hypo-osmolar fluid such as water, the outcome may be quite different. In this case the fluid lost through the osmotic diuresis is nearly matched with the water intake and a tenuous 'compensated osmotic diuresis' exists and can be maintained for a long time. Despite this, electrolytes, particularly sodium, are continually lost with the diuresis. In the early stages of diuresis, sodium is lost at around $150 \text{ mmol litre}^{-1}$ of urine (and potassium at around $5 \text{ mmol litre}^{-1}$ of urine). If this is replaced by water, reasonable osmolar balance is maintained. The net result,

however, is progressive depletion of sodium. This activates the renin–angiotensin system and aldosterone secretion increases, leading to increased sodium resorption from the tubules. Sodium losses decline as aldosterone concentrations increase. Aldosterone is very effective in retaining sodium even in the face of an osmotic diuresis. This hypersecretion of aldosterone has a knock-on effect on potassium stores since the high aldosterone concentrations reduce potassium resorption and urinary potassium excretion increases. The longer the phase of ‘compensated’ osmotic diuresis, where osmolar balance is maintained within reasonable limits, the greater the degree of overall sodium and potassium depletion. There is always also a ‘free water deficit’ as water, in the relative absence of sodium, is carried out with the diuresis.

Eventually, if untreated and if significant ketoacidosis does not develop, the sodium depletion reaches the stage where dehydration is so great that hypotension occurs. The patient then becomes oliguric and uraemic in the face of a substantial osmolar load. This is a very precarious situation characterized by a prune-like appearance with marked hypotension and raised serum creatinine and urea. If not treated promptly with parenteral rehydration, it is rapidly fatal. At this stage, the patient is depleted of, at the very least, 6 litres of water, 600 mmol of sodium and 500 mmol of potassium. Although the patient is so depleted of potassium, plasma potassium may be dangerously high as a result of potassium leaking out of cells into the extracellular fluid but no longer being excreted by diuresis. The most severe cases of fluid and electrolyte depletion are often seen in those who do not develop ketoacidosis. The depressed mental state correlates most closely with the degree of hyperosmolality and dehydration. Therapy should be targeted at correcting the electrolyte disturbance; insulin administration is of secondary importance.

If ketoacidosis does develop, the problems of metabolic acidosis may appear to dominate the clinical picture. It is important to remember, however, that severe fluid and electrolyte depletion is inevitable and that correction of this in a timely and appropriate manner is the hallmark of successful treatment.

Insulin replacement and the control of metabolism in diabetes

Current dogma would have us believe that administration of insulin to somebody with severely deranged diabetes suddenly and miraculously allows the cells in the body to breathe again and be restored to their former healthy state. This is, as we have seen, untrue, so it is amazing how long this dogma has persisted. A major aim of this review has been to set the record straight and rehabilitate Sir Edward Schafer’s correct perception of the role of insulin in regulating energy metabolism. As has been outlined in the previous section, the major problems of lack of insulin are

consequences of: (i) the osmotic diuresis resulting from the over-production of glucose or (ii) the metabolic acidosis resulting from the over-production of ketoacids. These occur as a result of a failure of insulin to regulate energy substrate release from stores in an orderly manner (chalone action of insulin). The consequent metabolic mayhem results from the unrestrained release of all energy stores and flooding of the extracellular fluid with too much of what is normally a ‘good thing’.

How then does insulin lower the blood glucose in diabetes?

Again this question is simple to answer and, by now, the answer will not be surprising. The use of tracer glucose infusions has shown not only that hyperglycaemia in the face of insulin deficiency is the result of over-production of glucose by the liver but also that insulin infusion lowers blood glucose by inhibiting hepatic glucose production. Indeed, rather than stimulating glucose uptake in tissues such as muscle, insulin in fact *reduces* glucose uptake. This is because the main factor driving glucose uptake is the ‘mass action’ effect of hyperglycaemia and the concentration gradient between the extracellular and intracellular glucose concentrations. Glucose transporters are *not* rate limiting under these conditions, even in the face of severe insulin deficiency.

Some of the experimental data behind our understanding of these mechanisms are illustrated in the results² shown in Fig. 6. In these studies a group of established type 1 diabetic patients volunteered to be studied on three occasions in the ‘insulin-withdrawn’ state 24 h after their last dose of insulin. The order of events was randomized. Patients received either an infusion of saline, low-dose insulin (2.6 u h^{-1}) or high-dose insulin (10.6 u h^{-1}) for 2 h. Tracer glucose was infused for 2 h before the insulin was started and the rates of glucose production and utilization were measured before and during the insulin infusions. Blood glucose concentration decreased more rapidly with high-dose rather than low-dose insulin and declined very slowly during the saline infusion (Fig. 6A). Insulin produced an immediate reduction of 50% in hepatic glucose production (R_a); this was the only change leading to the decrease in glucose during the low-dose study. The reduction in blood glucose concentration produced by the low-dose insulin infusion was followed by a reduction in glucose uptake (R_d) in keeping with the primary importance of the glucose concentration gradient rather than the insulin concentration in the regulation of glucose disposal. Thus insulin administration lowered glucose uptake into peripheral tissues. During the high-dose infusion there was a similar immediate halving of hepatic glucose production but insulin stimulated glucose uptake as well as inhibiting glucose production, so the decrease in blood glucose was more rapid (Fig. 6B). Under physiological conditions, plasma insulin concentrations rarely increase above $100 \text{ mu litre}^{-1}$ except transiently for

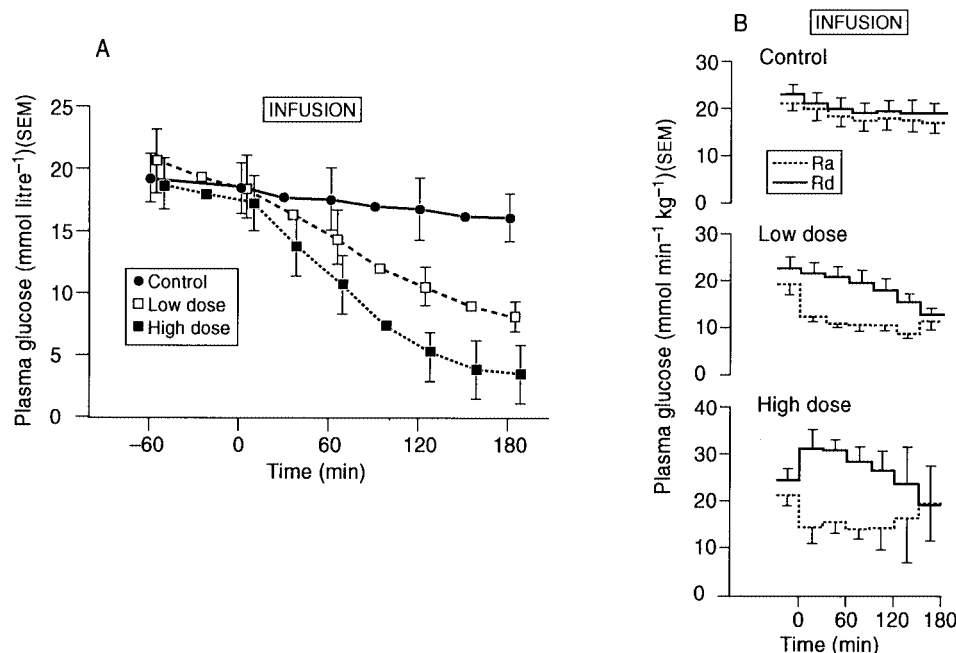


Fig 6 The data shown were obtained from a group of insulin-deprived patients with type 1 diabetes. Each volunteer was studied on three occasions 24 h after his or her last dose of insulin. (A) On each occasion the volunteers had a comparably increased fasting blood glucose concentration (as in Fig. 3). Each volunteer received an infusion of saline, low-dose insulin and high-dose insulin in random order. (B) Glucose production and utilization were measured using a continuous infusion of tracer glucose. Corrections were made for measured urinary glucose losses. The results show that fasting hyperglycaemia was caused by an elevated hepatic glucose production rate and that insulin lowered blood glucose through inhibition of hepatic glucose production. In the low-dose study (which produced insulin concentrations in the normal physiological range) glucose utilization was not increased at all; indeed, it fell in parallel with blood glucose concentration. In the high-dose study (which produced insulin concentrations of a magnitude seen only after a high-carbohydrate meal), there was a transient increase in glucose utilization but this was later followed by a progressive decrease in glucose.

a few minutes after a very high carbohydrate load. The low-dose infusion at 2.6 u h^{-1} produce insulin concentrations around $30\text{--}50 \text{ mU litre}^{-1}$ whereas with 10.6 u h^{-1} concentrations may exceed $200 \text{ mU litre}^{-1}$, which is rarely seen in normal life. The results are very much in keeping with Schafer's concept of a primary role of insulin in regulating blood glucose concentration through control of hepatic glucose storage, hence the evolution of the endocrine pancreas within the portal circulation.

Why is it important to understand the mechanisms regulating energy metabolism?

Anaesthetists are involved in the management of many types of disease, both in the perioperative setting and in intensive care. Diabetes is a common condition and people with diabetes consume proportionally more health care resources than expected purely on the basis of prevalence. Modern management of diabetes in everyday life has been shown to affect outcome profoundly; the same is true in the perioperative and intensive care setting. The excellent review in this issue on the 'Anaesthetic management of patients with diabetes mellitus', by Professor Hall,⁴ describes these advances in detail. The purpose of this article is to provide the modern science behind the practice

of medicine. A clear understanding of the physiology and pathophysiology is essential if a doctor is going to react appropriately when they are faced with a clinical scenario in which they have to make decisions based on their understanding of disease rather than their personal experience of handling similar cases. Since this happens frequently in medicine, it is one of our obligations to keep up to date in our education. It is clear from the author's experience that there is a gap between modern science and teaching in this area.

We conclude this article by examining how this basic science approach might influence the management of a common clinical scenario; the well controlled, uncomplicated diabetic scheduled for elective surgery.

The well controlled diabetic undergoing elective surgery

The importance of good metabolic control

Tight metabolic control is important at all times for both type 1 and type 2 patients; it is especially so at the time of surgery. If control has been tight for the weeks preceding surgery, then at admission these patients are essentially normal with respect to fluid balance and electrolyte status.

Therefore, in respect to these issues, they can be treated in a similar way to other patients. This should greatly simplify the immediate preoperative preparation. Probably the most useful marker of the quality of recent control is the percentage of glycosylated haemoglobin, expressed as HbA1c (normal < 6.2%).

HbA1c is only a measure of recent *glycaemic* control; fortunately, however, in the vast majority of people with diabetes, if glycaemic control is good then so are other metabolic processes. Indeed, in patients with diabetes treated with insulin who have 'tight glycaemic control', there can be 'over-control' of other metabolic processes. This is explained as follows. Exogenous insulin is given peripherally and the dose adjusted to control hepatic glucose production, not proteolysis. With subcutaneous or intravenous administration of insulin, the concentration in the portal circulation will always be less than in the systemic circulation. This is a reversal of the normal situation. Good glycaemic control thus invariably results in peripheral hyperinsulinaemia. Hence lipolysis and proteolysis are slightly 'over-controlled' (protein synthesis is normal—insulin-independent and regulated by growth hormone—but protein breakdown is reduced as a result of the 'chalone' action of peripheral hyperinsulinaemia).

A patient with an HbA1c of 7% is effectively normal. It is only when HbA1c is >9% that things begin to go wrong, especially with respect to the development of osmotic diuresis and consequent water and electrolyte loss. With an HbA1c of 12–15%, the patient is on the verge of diabetic ketoacidosis. In these patients, electrolyte imbalance, fluid balance and improvement in metabolic control must be corrected preoperatively.

Immediate preoperative metabolic control

Type 1 patients. Within a few years of the onset of diabetes, and even in the presence of excellent control, all type 1 diabetics stop producing endogenous insulin. Therefore, no type 1 diabetic can tolerate prolonged periods without exogenous insulin. Once the effect of their last dose of insulin has worn off, the brake on hepatic glucose production and catabolism is removed, resulting in the development over many hours of hyperglycaemia and ketosis. Traditionally, many of these patients have been managed with a mixture of glucose, insulin and potassium combined in one bag and administered intravenously. The use of separate infusions of glucose and insulin is also popular and probably allows tighter control, but has generated concern regarding the sudden discontinuation of the glucose if the bag finishes, a disconnection occurs or the cannula becomes obstructed (and is later flushed). Some of these concerns are equally valid if a saline infusion is used. However, as long as a well controlled type 1 patient continues their normal regimen (diet and insulin) up to the night before surgery (possibly modifying (reducing) their evening dose of insulin), then on the morning of operation their insulin

concentrations will be low and decreasing and their glucose normal or slightly increased. In all other respects they remain well controlled. From then on the consequence of continued starvation combined with avoidance of insulin has to be the gradual development of hyperglycaemia as the liver starts releasing increasing amounts of glucose. It thus seems logical that the only intervention these patients require to maintain good control in the preoperative period is insulin. They should not yet have developed any fluid or electrolyte imbalance greater than that in a non-diabetic. Therefore, they could be managed more simply, more safely and with tight control by the sole use of an insulin infusion in combination with a traditional sliding scale. This greatly simplifies the physical administration of treatment (only one cannula, no three-way taps or 'Y' connectors with anti-reflux valves, no fluid pumps) and also avoids the complicated picture of matching insulin requirements to two variables: endogenous production of glucose and exogenous administration of glucose. The requirement for calories or fluid, or both, intra-operatively or later may require the commencement of a glucose, or other intravenous fluid, infusion at that time.

Type 2 patients. All type 2 diabetics, even those who require insulin for good control, produce endogenous insulin, and their metabolic state improves with fasting. Indeed studies have shown that prolonged starvation for up to 3 weeks normalizes glucose metabolism in type 2 patients. The immediate preoperative management of well controlled type 2 patients can and should be relatively simple. Normal treatment and diet should be continued up to the commencement of their preoperative fast (except, possibly, for reducing or omitting their evening dose of insulin or oral hypoglycaemic, respectively). Following their overnight fast they can safely omit the morning dose of insulin or oral hypoglycaemic and, apart from regular monitoring, no further intervention is likely to be necessary.

Intra- and post-operative management

The further management of all diabetic patients should be dictated by frequent assessment of glycaemic control and the patient's ability to resume a normal diet (see reference 4). Hyperglycaemia in any diabetic patient unable to take a normal diet should probably be treated with an intravenous infusion of insulin from an infusion pump, the rate being guided by a sliding scale. The need for intravenous calories and fluids should be assessed, as it would be for non-diabetics in a similar setting.

References

- 1 Brown BR Jr. Anaesthesia and endocrine disease. In: Prys-Roberts C, Brown BR Jr, eds. *International Practice of Anaesthesia*. Oxford: Butterworth-Heinemann, 1996; 1/82/1–11
- 2 Brown P, Tompkins C, Juul S, Sonksen PH. Mechanism of action

- of insulin in diabetic patients: a dose-related effect on glucose production and utilisation. *BMJ* 1978; **i**: 1239–42
- 3** Hall SE, Saunders J, Sonksen, PH. Glucose and free fatty acid turnover in normal subjects and in diabetic patients before and after insulin treatment. *Diabetologia* 1979; **16**: 297–306
 - 4** McAnulty GR, Robertshaw HJ, Hall GM. Anaesthetic management of patients with diabetes mellitus. *Br J Anaesth* 2000; **85**: 80–90
 - 5** Schafer E. *The Endocrine Organs*. London: Longman, Green & Co., 1916
 - 6** Thomas SH, Wisher M, Brandenburg D, Sonksen PH. Insulin action on adipocytes: evidence that the anti-lipolytic and lipogenic effects of insulin are mediated by the same receptor. *Biochem J* 1979; **184**: 355–60