Summary  Background The gut and immune system form a complex integrated structure that has evolved to provide effective digestion and defence against ingested toxins and pathogenic bacteria. However, great variation exists in what is considered normal healthy gut and immune function. Thus, whilst it is possible to measure many aspects of digestion and immunity, it is more difficult to interpret the benefits to individuals of variation within what is considered to be a normal range. Nevertheless, it is important to set standards for optimal function for use both by the consumer, industry and those concerned with the public health.

The digestive tract is most frequently the object of functional and health claims and a large market already exists for gut-functional foods worldwide. Aim To define normal function of the gut and immune system and describe available methods of measuring it.

Results We have defined normal bowel habit and transit time, identified their role as risk factors for disease and how they may be measured. Similarly, we have tried to define what is a healthy gut flora in terms of the dominant genera and their metabolism and listed the many, varied and novel methods for determining these parameters.

It has proved less easy to provide boundaries for what constitutes optimal or improved gastric emptying, gut motility, nutrient and water absorption and the function of organs such as the liver, gallbladder and pancreas. The
many tests of these functions are described. We have discussed gastrointestinal well being. Sensations arising from the gut can be both pleasant and unpleasant. However, the characteristics of well being are ill defined and merge imperceptibly from acceptable to unacceptable, a state that is subjective. Nevertheless, we feel this is an important area for future work and method development.

The immune system is even more difficult to make quantitative judgements about. When it is defective, then clinical problems ensure, but this is an uncommon state. The innate and adaptive immune systems work synergistically together and comprise many cellular and humoral factors. The adaptive system is extremely sophisticated and between the two arms of immunity there is great redundancy, which provides robust defences. New aspects of immune function are discovered regularly. It is not clear whether immune function can be “improved”.

Measuring aspects of immune function is possible but there is no one test that will define either the status or functional capacity of the immune system. Human studies are often limited by the ability to sample only blood or secretions such as saliva but it should be remembered that only 2% of lymphocytes circulate at any given time, which limits interpretation of data.

We recommend assessing the functional capacity of the immune system by:
- measuring specific cell functions ex vivo
- measuring in vivo responses to challenge, e.g. change in antibody in blood or response to antigens
- determining the incidence and severity of infection in target populations during naturally occurring episodes or in response to attenuated pathogens.

**Key words** functional foods – biomarkers – gastrointestinal tract – microflora – digestion – immune system

**Introduction**

Around 60% of functional foods, principally pro- and prebiotics, are targeted at the gut and the immune system [1]. However, digestion and immunity are complex integrated physiological processes. They have evolved over millennia to provide effective absorption of nutrients, excretion of toxic and unwanted compounds and defence against environmental agents, especially pathogenic bacteria. Digestion involves physical processes that comminute food and transport it along the gut, chemical digestion by acid, enzymes, water/lipid/membrane interactions, neuronal and hormonal controlling systems that are modified by cortical pathways and transport and secretory processes in the gut epithelium. Similarly, gut defences include the physical barrier of the mucosa, motor activity, digestive secretions and the immune system. This latter comprises both innate and adaptive immunity, mucosal immune structures and the concept of tolerance to the commensal flora. In addition the huge numbers and diverse forms of bacteria that inhabit the gut give it a unique environment, which together with the wide variety of foods available for consumption today, make for the great variation in individual patterns of digestive and immune function.

Set against this complexity, how can we evaluate claims such as “promotes natural healthy digestion”, “maintains the balance of the intestinal flora”, or “enhances the defences of the body”? Can optimal function be defined? Can normal processes that vary considerably between healthy individuals be improved? Can we measure intestinal well being or define the limits of normal digestive and immune function? In an attempt to do this, we have described the normal functions of the digestive and immune systems and the commensal flora. Methods for measuring these are listed and common functions are defined. This is an active area of research where new products and methodologies are leading to changes in our understanding of the interdependence of these systems and providing opportunities for improved health.

**Existing claims for foods that benefit gut health and immunity**

The demand for foods that may benefit gut health is leading the functional food sector, and gut health and immune system claims on foods are frequently used health claims [2, 3]. There is broad consumer appeal for foods targeting a healthy digestive system and the body’s defences, which has accelerated market development. The appeal of products with gut health associated claims is that the product formulation can be targeted to an entire population, a sub-group or an at-risk group. Survey data show that gut health is well recognised by consumers and that they can identify the associated health claims and have some understanding of the benefits [4].

The principal products driving this market have been fermented milk or yoghurts containing “beneficial bacteria”. European interest in the gut health benefits of yoghurt containing “beneficial bacteria” began in the early 1900s and was endorsed by Metchnikoff and Tissier at the Pasteur Institute. European commercialisation of
specific yoghurts based on their gut health benefits was initiated as early as 1920 by Carasso [5]. Yoghurt intake is currently in the region of or greater than 15kg/capita/year in Finland, Greece, Germany, France, the Netherlands and Switzerland [6]. Furthermore, there is a strong tradition of regarding probiotic products as healthy in European countries such as Sweden [7].

Probiotics are defined as live microbial food ingredients that are beneficial to health. A wide range of bacteria is used in foods but lactic acid producing bacteria such as lactobacilli and bifidobacteria, commonly used as dairy cultures, tend to predominate in the probiotic food sector [8]. Probiotic products have dominated activities in the global and European gut health market and generally segment into two main lines; yoghurt (spoon-able and drinking) and single serve delivery drinks. The latter, which grew from zero to a European market value of $300 million in five years, accounts for the main current growth in the sector [9]. Non-dairy-based probiotic products are also available such as fruit juice drinks.

Dietary fibre (non-starch polysaccharide, NSP) also has a long-standing association with gut health benefits among consumers, particularly in relation to bowel habit and to satiety. Dietary fibres or plant cell-wall NSP are widely used as food ingredients [10]. More recently the gastrointestinal effects of prebiotics, which complement and extend those of probiotics and NSP, are also receiving increased attention in Europe [11, 12]. Prebiotics are defined as non-digestible food ingredients that beneficially affect the host by selectively stimulating the growth and or activity of one or a limited number of bacteria in the colon that have the potential to improve health [13]. Examples of prebiotic ingredients include some oligosaccharides, such as, galactooligosaccharides (GOS), fructooligosaccharides (FOS), inulin and isomaltooligosaccharide (IMO) [14, 15].

Synbiotic products, which are increasing in market presence, contain both a probiotic and a prebiotic. European introductions of prebiotic-based foods have been fragmented but have included yoghurts, synbiotic yoghurts, and bakery and cereal products.

Current estimates of the value of foods marketed on a gut benefit platform vary widely due to differences in the nature of products that are included in this category. One estimate suggests that the value in key markets (i.e. Europe, USA, Japan and Australia) is in the region of $3.5 billion and is forecast to reach $5 billion by 2005 [16]. Dairy-based foods are currently leading the sector in Europe, which is not unexpected, given that many European consumers have high traditional consumption patterns of such foods.

### Existing claims on foods

An ever-increasing variety of claims for gut health and immune function are found on foods on the European market. While generic claims such as ‘help keep your body in balance with probiotics’ and ‘diets rich in fibre keep your digestive system regular’ are common, product-specific claims such as ‘product X, the bifidofibres – a promoter of magnesium absorption’ or ‘help refresh and re-energise mind and body with X’s high fibre cereals’ are increasing. Furthermore, the benefits associated with intake of fibre, probiotics and prebiotics are widely reported in editorials, magazine articles, interview reports, on websites and in scientific publications (e.g. [17, 18]). Generic benefits attributed to fibre, probiotics and prebiotics in these sources are often stronger than on-pack claims and may relate to disease prevention or treatment although medical claims are not allowed on a product. Claims currently made can broadly be categorised into five main groups:

- **Content**
- **Functional**
- **Enhanced function**
- **Reduction of disease risk or disease risk factor**
- **Medical claims (not permitted on product)**

### Content

Much attention has focused on claims that give the content of the ingredient in the case of fibre and prebiotic-containing foods and the nature and numbers of bacteria in the case of probiotic-containing foods. With regard to probiotic products, specific mention of the genus, species and strain of the bacteria is not unusual and may appear in the ingredient list or as a flash on the pack. Details of actual bacterial numbers are often mentioned. This is not legally required though it is advocated that the minimum viable number of each probiotic strain should be described on labels [19]. In addition to bacteria type and count, attributes of the bacteria, such as their origin and ability to survive transit through the gut, are often mentioned. Furthermore, the concept that the bacteria used in these products are good, beneficial, and friendly or probiotic is claimed widely and, whilst not part of the traditional content claim, are generally supported by a broad basis of science. Examples of these content type claims are shown in Table 1. In the future, probiotic bacteria will be classified as “other substances” for the purpose of health claims, meaning “a substance other than a nutrient that has a nutritional or physiological effect” (Commission of the European Communities 2003/01655(COD) [20]).
**Table 1** Examples of existing content claims on foods

- Fibre: typical value 2.5 g per 100 g*
- 50 % of your daily fibre* needs in one bowl
- Ingredient: dietary fibre* (inulin)
- Ingredient: soluble fibre* (inulin)
- Bifidobacteria longum
- Lactobacillus casei Shirotai
- Product X contains special active cultures
- Product X contains beneficial active bacteria
- Product X supplies 50 million colony forming units of Lactobacillus plantarum 299v
- Yoghurt with bifidus essensis cultures
- 10 billion good bacteria
- 100 million Lactobacillus reuteri bacteria (the recommended daily intake)
- Product X provides all the good bacteria your body needs

* There is no currently agreed method for measuring dietary fibre in the EU

**Functional**

A functional claim is any claim that states, suggests or implies the role of a food category, a food or one of its constituents in growth, development or normal physiological functions of the body.

The most common functional claims in the gut and immunity area relate to “maintenance of a healthy flora”, to long-standing concepts such as a healthy or balanced digestive system, and the body’s natural defences. In addition, maintenance of natural defences and natural resistance are also popular claims though it is not usually explicitly detailed whether the basis for these claims is immunological, non-immunological or both (Table 2).

**Enhanced function**

Enhanced function implies any health claim that states, suggests or implies the consumption of a food category, a food or one of its constituents has a specific beneficial effect, beyond that normally obtained from the diet, on physiological functions of the body.

The most frequently made enhanced function claims again relate to improvement of digestion, the intestinal flora, natural defences and natural resistance. Specific claims relating to the bifidogenic or effect bifidus of prebiotics are evident on food products and have been viewed positively by some European regulators. For example, the French Food Safety Agency allows such claims but states that the bifidogenic claim cannot be accompanied by claims as to the beneficial or curative effects of bifidobacteria [21]. Claims relating to mineral absorption are also made for prebiotics. A recent innovation has been the introduction of claims relating to improvement in aspects of quality of life, such as well being, morale and stress levels, which are regarded as pertinent to health. Typical product-associated statements in this category are shown in Table 3.

Functional and enhanced functional claims will be subsumed in forthcoming EU legislation into a single category of health claims (Commission of European Communities 2003/01655(COD) [20]).

**Reduction of disease risk and risk factors**

General provisions relating to the labelling, presentation and advertising of foods prohibit the use of information that would attribute medicinal properties to food [22]. It specifically prohibits attributing to foods any properties of prevention, treatment or cure of a human disease or any reference to such properties.

Claims to reduce disease risk or risk factors for disease tend, therefore, to be quite general and implicit rather than explicit in most cases. For example, much emphasis is placed on the fortification of the intestinal flora as a barrier against opportunistic pathogens with the implication that eradication of harmful bacteria is beneficial. Explicit claims on probiotic products have also been made in relation to relief of irritable bowel symptoms and cholesterol reduction. Examples of claims in this area are outlined in Table 4.

While information relating to disease risk reduction

**Table 2** Examples of existing functional claims on foods

- Helps you stay regular
- Aids regular transit
- Maintaining the balance of the digestive system
- Helps maintain a healthy digestive system
- Helps maintain the balance of the intestinal flora
- To help balance the probiotics in your digestive system
- Supports the body’s natural defences

**Table 3** Examples of existing enhanced function claims on foods

- Proven benefits to your digestion
- Improves digestion
- Promotes natural healthy digestion
- Modulates bowel activity
- Promotes natural rhythm of the bowel
- Improves intestinal transit
- Active on intestinal comfort
- Boosts the body’s immune system
- Stimulates the immune system
- Boosts natural resistance
- Actively strengthens the body’s natural resistance
- Help strengthen your natural defences
- Enhances defences of the body
- Helps your body to protect itself
- Helps the body to defend itself against external aggressions
- Feel fabulous with fibre
- Refresh and re-energise your mind and body
- Product X stimulates mineral absorption
and the therapeutic properties of fibre, probiotics and to a lesser extent prebiotics abound in the scientific and lay literature, few food products currently on sale in the European market actually bear such claims. An example of one probiotic product that does make an explicit reference to its prophylactic and therapeutic benefits states ‘product x prevents and shortens many types of diarrhoea’. It is worth noting that specific probiotic products have been registered, based on traditional use, as natural remedies by the Swedish Medical Product Agency [7]. Generally examples of claims in this area are limited and those that exist tend to be associated with areas where innovative research is on-going such as, the dietary management of allergy [23].

**Digestive functions of the gut**

The digestion of food and absorption of nutrients is a complex but well integrated and highly efficient process involving nerves, muscle, secretory organs and gut-derived hormones. Normal function is described by a wide range of values that vary between people and with age, sex, ethnic origin and many lifestyle factors.

The oral cavity and tongue surfaces are richly supplied with sensory nerve endings for touch, temperature, taste and pain. Impulses from these endings tell the brain if food is acceptable and then mastication and mixing with saliva begins. Saliva acts primarily as a solvent for solids in food. It contains bicarbonate for buffering, and a number of proteins including amylase, which start the hydrolysis of starch and glycogen, immunoglobulins and lysozymes with antibacterial activity and epidermal growth factor-alpha to enhance mucosal resistance to injury, etc. Thereafter, coordinated pharyngeal and oesophageal motility propels food down to the stomach (deglutition).

The oesophagus at rest is closed by muscle sphincters at its upper and lower ends, so that the passage of a bolus of food requires a series of contractions and reflex sphincteric relaxations that are coordinated by the swallowing centre in the medulla. The lower oesophageal sphincter also plays an important role in the control of reflux of gastric acid into the oesophagus. Gastro-oesophageal reflux is a normal physiological phenomenon that occurs especially during the postprandial period. However, individuals with frequent and prolonged episodes of reflux develop gastro-oesophageal reflux disease, a condition of increasing prevalence in the developed world that seems to be related to food habits, smoking and stress. Up to 44% of the adult population experience symptoms of gastro-oesophageal reflux disease and 18% take digestion aids [24]. Dietary intervention may reduce this burden. Ambulatory monitoring of intraoesophageal pH is a useful technique to assess gastro-oesophageal reflux of acid (Table 5).

**Gastric function**

The stomach is a mixing chamber. It stores food and converts it to a semi-liquid slurry called chyme by mixing food with gastric juice. The stomach accommodates variable volumes of food by expanding the fundus, which is the upper part of the stomach, and acts as a reservoir. Mixing is mainly due to antral contractions, and gastric emptying results from propulsive antral waves associated with pyloric relaxation. Gastric juice contains acid and pepsins, which are endopeptidases active at low pH, and is secreted in response to the sight of food and also when a meal enters the stomach. After thorough mixing, chyme is delivered into the small bowel at a controlled rate. Chyme with high caloric density, viscosity or osmolarity is emptied more slowly than non-nutrient liquids. The inhibitory effect on gastric emptying is due to the release of cholecystokinin and gastric inhibitory polypeptide in response to fatty meals or to stimulation of osmoreceptors. Delayed gastric emptying is a common dysfunction in otherwise healthy individuals. Subjects with functional dyspepsia may exhibit delayed emptying of solid meals and an altered sensitivity of the stomach, so that they perceive and tol-

---

**Table 4** Examples of existing reduction of disease risk and risk factor claims on foods

<table>
<thead>
<tr>
<th>Claim</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keep harmful bacteria at bay</td>
</tr>
<tr>
<td>Helps reduce the number of harmful micro-organisms</td>
</tr>
<tr>
<td>Assists in elimination of harmful organisms</td>
</tr>
<tr>
<td>Counteracts potentially harmful bacteria</td>
</tr>
<tr>
<td>Produces an antimicrobial substance which fights against pathogens</td>
</tr>
<tr>
<td>Product X helps against harmful bacteria, stimulating the immune system, helping reduce the risk of infections and digestive problems</td>
</tr>
<tr>
<td>Product X helps calcium absorption and is ideal for osteoporosis or persons in growing age</td>
</tr>
<tr>
<td>Product X promotes recovery from milk allergy, alleviates allergic inflammation and reduces atopic skin symptoms</td>
</tr>
</tbody>
</table>

---

**Table 5** Normal values for gastro-oesophageal pH and reflux

<table>
<thead>
<tr>
<th>Variable</th>
<th>Normal Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gastric pH, fasting</td>
<td>Between 1 and 3</td>
</tr>
<tr>
<td>Gastric pH, after a meal</td>
<td>Up to 4 to 5</td>
</tr>
<tr>
<td>Oesophageal pH at rest</td>
<td>6 to 7</td>
</tr>
<tr>
<td>Oesophageal pH during a reflux episode</td>
<td>&lt; 4</td>
</tr>
<tr>
<td>Number of reflux episodes per day</td>
<td>&lt; 50</td>
</tr>
<tr>
<td>(most common after a meal)</td>
<td></td>
</tr>
<tr>
<td>Maximal duration of a physiologic reflux episode</td>
<td>&lt; 9 min</td>
</tr>
<tr>
<td>Total time per day with an oesophageal pH below 4 (reflux time)</td>
<td>&lt; 60 min</td>
</tr>
</tbody>
</table>

See [26]
erate significantly lower levels of intraluminal pressure and gastric wall tension than healthy subjects [25].

The gastric barostat is used to measure compliance and sensitivity [26]. A number of methods for the measurement of gastric emptying have been described (Table 6). These include direct aspiration of gastric contents after a test meal through a nasogastric tube, bio-electrical impedance, scintigraphy and stable isotope breath test. Ultrasound-based methods have also been described. Of these methods, scintigraphy is the reference method. When compared with other methods, scintigraphy is non-invasive, does not disturb normal physiology, uses commonly ingested foods rather than non-nutrient substrates or markers, permits quantification of transit of solids and liquids simultaneously and results in low radiation exposure [26] (see also Table 12 in [27]).

### Digestion, absorption and motility

During fasting, a pattern of periodic motor activity is always present in the small bowel. The normal migrating motor complex consists of a series of rhythmic and regular contractions that commence in the gastric antrum, progress into the duodenum and migrate distally towards the ileum. This is followed by a phase of motor quiescence, and thereafter a new phase of propulsive motor activity starts again. Following ingestion of a meal, small bowel motor activity becomes much more intense. Contractions are variable in frequency, amplitude and speed of propagation. Most waves travel only a short distance down the gut and disappear. Pendular activity, that moves contents backwards and forwards short distances, have also been demonstrated. This activity pattern enables the mixing of chyme with pancreatic juice and bile, and also increases the contact time of nutrients with the absorptive mucosal surface. The net result is luminal flow and transit through the small bowel up to the ileo-caecal valve. Solids and liquids travel through the small intestine at similar rates.

Pancreatic secretions include bicarbonate for buffering gastric acid and digestive enzymes for hydrolysis of polysaccharides (amylase), proteins and peptides (trypsinogen, chymotrypsinogen, elastase and pro-carboxypeptidase), lipids (lipase and co-lipase), and nucleotides (ribonuclease, deoxyribonuclease). Bile is formed by the liver and consists of an aqueous solution of organic compounds i.e. bile acids, bilirubin, cholesterol, fatty acids, phospholipids and electrolytes (i.e. Na, K, Cl, bicarbonate), that is secreted into the duodenal lumen when food enters the small bowel. Bile acids combine with cholesterol and phospholipids to form micelles, which are aggregates able to carry lipids in aqueous suspension. Most bile acids are eventually re-absorbed in the ileum returning via portal blood to the liver in a process known as enterohepatic circulation.

![Table 6 Methods used to assess gastric emptying](image-url)

<table>
<thead>
<tr>
<th>Method</th>
<th>Measurement</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epigastric impedance¹</td>
<td>The impedance of the epigastrium to a low intensity electrical current increases while liquids of low electrical conductivity are being drunk, and declines as the liquid leaves the stomach.</td>
<td>Non-invasive.</td>
<td>Only used to evaluate gastric emptying of large-volume liquid meals. The method is susceptible to movement artefacts (non-valid measurements in around one in four infants).</td>
</tr>
<tr>
<td>Stable isotope breath test²</td>
<td>Changes in the $^{12}$C to $^{13}$C ratio in expired air after oral administration of a labelled test substance, absorbed rapidly in the duodenum, metabolised in the liver and excreted as $^{13}$CO₂.</td>
<td>Non-invasive. Overcomes the risk of radiation (repeated studies are acceptable). Measures gastric emptying of solid or liquid meals. Reference method.</td>
<td>Does not permit imaging or regional assessment of gastric function. Impaired absorption or liver disease can influence the outcome.</td>
</tr>
</tbody>
</table>

¹ See [342]; ² See [26]
Table 7  Hydrogen breath tests

<table>
<thead>
<tr>
<th>Sugar</th>
<th>Application</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glucose</td>
<td>Detection of small bowel bacterial overgrowth.</td>
<td>False negative results in subjects who do not produce hydrogen.</td>
</tr>
<tr>
<td>Lactulose</td>
<td>Detection of non-producers of hydrogen. Oro-caecal transit time.</td>
<td>In subjects with small bowel bacterial overgrowth oro-caecal transit time cannot be recorded (lactulose is metabolised in the small bowel)</td>
</tr>
<tr>
<td>Lactose</td>
<td>Detection of lactase deficiency.</td>
<td>False negative results in subjects that do not produce hydrogen. False positive results are found in small bowel bacterial overgrowth.</td>
</tr>
<tr>
<td>Xylose</td>
<td>Detection of malabsorption. Screening of coeliac disease.</td>
<td>False negative results in subjects that do not produce hydrogen. False positive results are found in small bowel bacterial overgrowth.</td>
</tr>
</tbody>
</table>

The bile acids that escape are metabolised (deconjugated and dehydroxylated) by colonic bacteria to secondary bile acids and are either partly reabsorbed from the colon or excreted in the faeces. Colonic bacteria also degrade bilirubin and its end product, stercobilin, is responsible for the brown colour of faeces.

The small bowel mucosa exhibits a vast surface adapted to the absorption of nutrients. The luminal surface of the enterocyte, known as the brush border, contains many enzymes responsible for the terminal hydrolysis of oligosaccharides and peptides to monosaccharides and amino acids respectively. The end products of enzymatic digestion by pancreatic and brush border enzymes, as well as vitamins, minerals and electrolytes are absorbed into the enterocyte by active mechanisms involving specific carrier proteins. Water-soluble substances also cross the cell membrane by passive diffusion through aqueous pores (transcellular pathway) or enter through the lateral spaces and tight junctions between enterocytes (paracellular pathway). Triglycerides and phospholipids are hydrolysed by pancreatic enzymes. Free fatty acids and mono-glycerides need to form micelles to diffuse through the layer of unstirred water and reach the absorptive surface. Then, the micelle is disrupted and the lipids diffuse through the membrane into the enterocyte.

Malabsorption may be due to villus atrophy as occurs in coeliac disease or as a consequence of long-term starvation or enzymatic deficiencies, for instance in severe pancreatic insufficiency or loss of β-galactosidase activity. Pancreatic insufficiency is detected by measurement of the enzyme elastase in faecal samples, which is a non-invasive method to evaluate exocrine pancreatic function. Whilst severe pancreatic insufficiency is uncommon, the prevalence of lactose malabsorption in adult populations varies between 5 and 15% in Northern European and American countries and 50 to 100% in African, Asian and South American countries [28]. Malabsorption of lactose or monosaccharides such as sorbitol or fructose has been suggested to be responsible for symptoms in some subjects with irritable bowel syndrome [29].

**Breath tests**

Malabsorption of a specific sugar can be investigated by non-invasive and safe techniques using a breath test [30]. Table 7 summarises the most commonly used tests based on the detection of an increased excretion of hydrogen in breath after oral administration of a sugar. The source of hydrogen is the fermentation of the carbohydrate by bacteria in the gut lumen (germ-free animals or newborn infants have no significant breath hydrogen). A limitation of this technique is that a considerable fraction of the normal population (around 15%) are methane producers rather than hydrogen producers, and do not exhibit a peak excretion of hydrogen in breath when a carbohydrate enters the caecum. Tests are negative in those subjects and in patients on antibiotic therapy. By contrast, some individuals exhibit a high baseline excretion of hydrogen in breath that makes the test unsuccessful. Rigid dietary restrictions, such as avoiding food with nonabsorbable carbohydrates for several days prior to the test, can overcome this problem. Some laboratories simultaneously measure hydrogen and methane with every test, whereas others measure methane selectively after negative lactulose tests. However, methane breath tests have not been standardised, and recently developed 13C breath tests offer a better alternative [31]. These tests measure the pulmonary excretion of labelled carbon dioxide produced from the fermentation of labelled substrates using the stable 13C isotope. These are useful in subjects who are not hydrogen producers, and also offer the possibility of developing a wider variety of breath tests. In addition, stable isotope-labelling is a very useful tool to investigate absorption of specific substances (calcium, lactose, amino acids, etc) which then can be detected in blood and/or urine samples.

**The colon**

The colon completes the digestive process through fermentation of unabsorbed residues from the small bowel. The slow motility patterns of the colon allow bacterial proliferation at high densities, so that bacteria play a major role in colonic physiology (see later). Fermentation provides energy for bacterial growth and produces hydrogen, methane and carbon dioxide and short chain fatty acids, than can be absorbed by the host. Short chain fatty acids constitute the principal source of metabolic energy for the colonic epithelium. As a result of fermenta-
The colonic epithelial cells recover water and electrolytes from the lumen. The colon absorbs sodium by an active carrier and chloride in exchange for bicarbonate. Potassium is the dominant cation excreted in faecal fluid. Transit time and stool weight are biomarkers of overall colonic function (see later).

**Gastrointestinal well being**

A sense of intestinal well being is something to which we all aspire, yet it is an ill-defined state that is often equated simply with an absence of symptoms. “Being well” is frequently talked and written about, but rarely if ever measured, because the boundaries and characteristics of “wellness” are ill defined. However, we all know when we are well, and may be aware of those lifestyle factors, such as diet, that make or keep us well. Achieving and maintaining gastrointestinal well being is important, but neglected by medical and nutritional science, yet is often a target for functional foods and health claims. The sense of well-being is probably the outcome of several neuronal, hormonal and cortical events in response to digestion.

**Mechanism of conscious sensation of the gut**

To accomplish its many functions, the gut has a complex network of regulatory neuro–humoral reflexes that control motility, secretion, absorption, blood flow, and, possibly, the immune system. This integrated network modulates the responses to a meal, the preparatory processes of the gut during the interdigestive period, the progression of residues through the different colonic compartments and faecal evacuation. The digestive system also has an afferent network that gives rise to conscious sensations in response to gut stimuli [32]. A series of modulatory mechanisms tune the afferent input at different points between peripheral receptors and the brain cortex, and thereby, determine final perception.

The way abnormal sensations, that is gastrointestinal symptoms, are generated has been studied in recent years [32]. Although the mucosal surface is not sensitive to touch, distension by gas, for example, and thermal changes induce perception. The type and referral area of the sensations, depend on the type of the stimulus and its location in the digestive tract. The intensity of perception depends on the magnitude of the stimulus and the size of the area stimulated, that is the number of receptors activated. Perception may also be influenced by the response of the gut to mechanical stimuli. For example, the same distending volume will induce less symptoms if the gut is relaxed, and conversely, abnormal relaxation may result in symptoms in response to normal filling of the gut. Digestive symptoms, which occur in the absence of a detectable cause, that is in subjects with functional gastrointestinal disorders, may be related either to abnormal control of gut function or to visceral hypersensitivity and exaggerated gut perception [32]. Hence, symptoms in these subjects are induced by normal stimuli that are unperceived by healthy subjects.

**Sensations arising from the gastrointestinal tract**

Gut derived symptoms are commonly experienced by the general population and form part of the normal physiological digestive process. The boundary between normal, pleasant sensations and what individuals consider abnormal or troublesome is unclear. More than half the population frequently experience one or more gut sensations such as pain, change in stool consistency, bloating, incomplete evacuation or urgency [33, 34]. Such sensations when codified [35] merge into the irritable bowel syndrome (IBS), which is one of the most frequent problems in hospital gastroenterological practice. Prevalence of IBS ranges from 3–22 % and clearly indicates that some people find sensations such as pain and disordered bowel habit clearly objectionable [36–38] and a restraint on GI well being. Perception of sensation has been studied both in response to probe stimuli in experimental conditions as well as in free-living in subjects with clinical symptoms. However, the concept of gastrointestinal well being which has been proposed [39–41] has been much less studied. The existence of specific pleasant sensations derived from the gastrointestinal tract is supported by uncontrolled observations. Such sensations are primarily related to the intake of meals and the evacuation of faeces, i.e. gratifying sensations, such as satiation and complete rectal evacuation, and conceivably also preparatory sensations, such as appetite and the call to stool. Other physiological events, such as eructation and the emission of wind from the anus, may also contribute to gastrointestinal well being. Sensations related to thirst, taste, smell, the desire of specific types of foods (salt, sweet) and nonspecific sensations, such as “easy digestion”, for instance related to some foods, digestive liquors and infusions could also be considered part of this concept. The perception of all of these will be modulated by age, sex and cultural background.

**Subjects and methods**

It has been shown that subjects with symptoms primarily originating from the lower gastrointestinal tract, i.e. irritable bowel syndrome, have intestinal hypersensitivity and exhibit exaggerated perception of gut stimuli [42]. It has further been shown that these patients may have abnormal control of gut motility [43]. Similarly, pa-
tients with symptoms probably originating from the upper gastrointestinal tract, such as functional dyspepsia, have gastric hypersensitivity and abnormal accommodation reflexes, that can be measured by specific laboratory tests [44, 45]. Conceivably, in both irritable bowel syndrome and functional dyspepsia, gut sensory and reflex abnormalities interact to produce the clinical symptoms. The same methodologies can be applied to study the potentially beneficial effects of specific food components and extended to investigate the potential induction of pleasant sensations or gastrointestinal well being.

Studies may target the general population, in an effort to show the specific effects of foods on gastrointestinal well being (functional effects), or specific sub-populations that lack gastrointestinal well being (enhanced factor) or even those with gastrointestinal symptoms such as irritable bowel syndrome to try to improve unpleasant sensations (reduction of disease risk).

For each type of sensation, two types of scales should be used, one to measure the intensity of perception and the other to measure its affective dimension (pleasantness/unpleasantness) [46]. The sensations potentially related to well being outlined above can be targeted as outcomes. Sensations opposing well being i.e. gastrointestinal symptoms, would include nausea, abdominal pressure/fullness, bloating, colic, soreness, and incomplete rectal emptying [32].

The effect of the consumption of specific foods should be compared with consumption of control food, ideally using a double blind randomised design. Studies can be done under free-living conditions, or with the use of experimental probe stimuli in the laboratory.

Mechanisms of action

Gastrointestinal symptoms can be related to gastrointestinal dysfunction, particularly altered sensitivity and altered reflexes, and these dysfunctions can be demonstrated experimentally in the laboratory. Subjects with gastrointestinal complaints frequently attribute their symptoms to gas. The methodology to study gastrointestinal gas handling and tolerance has been well developed and could provide a marker for some specific claims in this respect [47, 48].

In contrast to the study of abnormal sensation, the area of gastrointestinal well being is less developed, and it is not known which type of gut stimuli and gut responses may induce pleasant sensations. It seems very unlikely that gastric emptying, small intestinal transit or colonic transit can produce pleasant sensations. Other physiological events, such as gut secretion, absorption, intestinal blood flow, and microflora metabolism may be related to pleasant sensations and be used as markers that could be explored. Potential markers would depend on the specific pleasant sensations considered; for instance, the characteristics of the faeces (volume, consistency, dryness, compactness) could be related to satisfactory evacuation. Since gut stimuli induce reflex responses and hence perception [32], the reflex responses may correlate with pleasant sensations and could serve as markers of these. In this respect, measurable physiological responses to gastric or rectal filling could be considered. The identification of objective markers of physiological gastrointestinal well being is complex and will require appropriate validation studies, but is important for the concept of functional foods and for making claims.

Bowel habit and gut transit time

Bowel habit

Bowel habit is a useful overall biomarker of gut, especially colonic, function and is usually defined in terms of frequency of defaecation, stool consistency and form and stool weight. In pathological conditions, other descriptive parameters may be recorded including incomplete rectal emptying, painful defaecation, the presence of blood or mucus, urgency and incontinence.

Frequency

The modal frequency of defaecation for most adults living in western countries is one stool a day, with 95% of the population passing between three a day and three a week [49–53]. The frequency of defaecation is greater in early life at four times a day in the neonatal period (range 1–9), with breast-fed babies having a higher rate than bottle-fed babies. The adult pattern of once a day is reached by the age of four years [54]. As old age (over 70 years) approaches, there is a tendency for bowel habit to become less frequent and laxative use more common [55]. Bowel frequency is very variable and in developing countries the modal frequency is nearer twice a day [56, 57]. Bowel frequency, and consistency, can conveniently be recorded using a bowel habit diary [39, 50].

Consistency and form

In the normal population, faecal form and consistency varies from hard, small, fragmented droppings through cylindrically shaped stools to soft porridge-like amorphous material. It has been measured by a variety of means including visual assessment [50, 58], subjective evaluation [59, 60] and by using a penetrometer [51] or analysis of water content [61]. Stool consistency is not uniform and may vary within a single motion. In general, as daily stool weight increases so stools become softer and less well formed. Consistency is related to wa-
ter content, which is normally 70–80%. Men pass softer stools than women. Constipated subjects generally pass harder stools with a water content of less than 70% [59, 62]. It is the hardness and dryness of stools that is thought to cause discomfort on defaecation. Physical methods of stool consistency are impractical for most purposes and consistency is usually recorded in a diary after visual inspection.

**Stool weight**

Stool weight varies substantially in different populations (Table 8). In the United Kingdom, daily stool weight is about 110 g/d with a wide range. Men pass significantly more than women and the young more than the elderly [52]. 47% of men and 51% of women have stool weights less than 100 g/d, and 17% of women and 1% of men pass less than 50 g/d, i.e. an amount of that overlaps with that seen in constipation (see later). Stool output is very variable from day to day in individuals [51, 56, 58, 63]. The variation is the result of hormonal cycles, such as the menstrual cycle, dietary patterns, mood changes, social pressures, e.g., whether defaecation has to be suppressed or not, travel and, possibly in hot climates, fluid intake. Changes in bowel habit are reported in long-distance runners. However, moderate exercise sufficient to improve physical fitness, but with diet held constant does not affect stool weight or transit time [64]. To measure average daily stool weight in individuals requires complete faecal collections for at least five consecutive days preferably using markers [65]. For population studies, timed collections over two or more days are preferable, rather than single stool studies, which tend to overestimate population averages. An assessment of markers used for bowel transit studies, and as balance markers in faecal collections is given in Table 9.

### Determinants of bowel habit

Bowel habit is determined primarily by diet and transit time. Any carbohydrate that reaches the large bowel will affect bowel habit. Of the carbohydrates that have been studied, dietary fibre, specifically non-starch polysaccharide (NSP), is the only dietary component to have been shown consistently over many years to control bowel habit. Feeding major components of the diet such

<table>
<thead>
<tr>
<th>Population</th>
<th>Mean Daily Stool Weight g</th>
<th>Range or SD</th>
<th>Sex</th>
<th>N</th>
<th>Source of data</th>
</tr>
</thead>
<tbody>
<tr>
<td>UK</td>
<td>122</td>
<td>46–415</td>
<td>M</td>
<td>106</td>
<td>Cummings et al. 1992 [52]</td>
</tr>
<tr>
<td></td>
<td>102</td>
<td>19–259</td>
<td>F</td>
<td>84</td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban*</td>
<td>136</td>
<td>71</td>
<td>M</td>
<td>30</td>
<td>Cummings et al. 1982 [343]</td>
</tr>
<tr>
<td>Rural*</td>
<td>169</td>
<td>93</td>
<td>M</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban*</td>
<td>176</td>
<td>88</td>
<td>M</td>
<td>30</td>
<td>Reddy et al. 1978 [344]</td>
</tr>
<tr>
<td>Rural*</td>
<td>196</td>
<td>82</td>
<td>M</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Rural*</td>
<td>274</td>
<td>–</td>
<td>M</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Tonga (NZ)</td>
<td>162</td>
<td>68</td>
<td>F</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Maori</td>
<td>119</td>
<td>48</td>
<td>F</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>New Zealand (E)</td>
<td>113</td>
<td>46</td>
<td>F</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>311</td>
<td>19–1505</td>
<td>M, F.</td>
<td>514</td>
<td>Tandon and Tandon 1975 [346]</td>
</tr>
<tr>
<td>Japan</td>
<td>165</td>
<td>69</td>
<td>M, F.</td>
<td>44</td>
<td>Tarida 1984 [347]</td>
</tr>
<tr>
<td></td>
<td>195</td>
<td></td>
<td>M</td>
<td>17</td>
<td>Alder et al. 1977</td>
</tr>
<tr>
<td>South Africa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White students</td>
<td>173</td>
<td>120–195</td>
<td>M, F.</td>
<td>100</td>
<td>Burkitt et al. 1972 [57]</td>
</tr>
<tr>
<td>Black urban school children</td>
<td>165</td>
<td>120–160</td>
<td>M, F.</td>
<td>500</td>
<td></td>
</tr>
<tr>
<td>Black rural school children</td>
<td>275</td>
<td>150–350</td>
<td>M, F.</td>
<td>500</td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>120</td>
<td>–</td>
<td>M</td>
<td>15</td>
<td>Glober et al. 1977 [348]</td>
</tr>
<tr>
<td>Japanese</td>
<td>121</td>
<td>–</td>
<td>M</td>
<td>47</td>
<td>Reddy et al. 1978 [344]</td>
</tr>
<tr>
<td>New York*</td>
<td>78</td>
<td>–</td>
<td>M</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

* Based on collection of single stool
NZ Living in New Zealand; E Of European origin
as fat and meat have no effect on bowel habit [66, 67]. Resistant starches have a small effect [68]. Probiotics have not been convincingly shown to affect bowel habit, whilst prebiotics have a small effect [39]. Certain foods such as prunes, bananas, beer or spicy dishes are reported to affect the bowel habit of individuals.

## Transit time

Transit time is the time it takes for a substance to pass through the gut. Of the 24–72 hours which on average most substances spend in the human gut the greater part of this, about 18–64 hours, is spent in the large bowel with gastric and small-gut transit accounting for 4–8 hours only [69–73]. Transit time is, therefore, mainly a colonic event.

Transit has interested scientists and clinicians for many years if only to explain the occasional appearance of identifiable foodstuffs in the faeces. It is also recognised that patients with diarrhoea have rapid transit and those with constipation the reverse [74]. It has been observed that slow transit time is associated epidemiologically with a high prevalence of large-bowel disorders, particularly diverticular disease and colon cancer [57], and it is implied that slow transit itself is important in determining metabolic events in the colon which are important in the aetiology of these disorders. Populations with fast transit often have a low prevalence of colonic disorders and are thought to have much higher dietary fibre intakes than those with slow transit. Adding fibre to the diet speeds up transit [72, 75–77] and so may be one factor in the prevention of these disorders.

### Physiological significance of transit

Many variables interact in the colon, the result of which is a measurable physiological event, which is transit. These variables include diet, bowel habit, stress, exercise, hormones, the overall effect of small-intestinal absorption and secretion, colonic anatomy, the way in which water and electrolyte secretion is controlled, and other transluminal events. Transit is an overall measure of colonic function. It gives an indication of what is going on in the ‘black box’, which is the colon, and as such it is a useful measure.

Transit, along with diet, is a major determinant of faecal weight and bowel habit [57, 78]. Experiments in humans in which transit has been manipulated by drugs have shown clear effects on bowel habit with slowing of transit producing reduced faecal output and decreased faecal frequency [79].

Certain metabolic changes in the colon have also been associated with changes in transit. Evidence for a relationship between transit and colonic steroid metabolism can be found in Davignon et al. [80], in which on equivalent diets, subjects showed an inverse relationship between transit time and the formation of neutral-steroid conversion products (coprostanol and coprostanone). Other bacterial metabolites in the colon can also be related to transit. A study of the effect of fibre from carrot, cabbage, bran, apple and guar [81], showed faecal-ammonia increased with lengthening transit time and that there is a significant relationship between the two. Both ammonia [82, 83] and degraded bile acids [84, 85] have been implicated in colonic tumour formation and if their production is determined by transit, the association of slow transit with colon cancer [86] could be important.

In other studies it has been shown that high urinary phenol excretion is associated with slow transit [87]. Many years ago Macy [88] reported a similar relationship between transit and urinary sulphate excretion in children. Both the ethereal sulphate and urinary phenol are products formed as a result of bacterial metabolism of dietary protein. Macy also noted a relationship between transit and the digestion of fibre in the gut, an observation also made by Southgate and Durnin [89].

All these studies point to a relationship between transit and metabolism in the colon. Whilst none furnish

---

**Table 9** Markers used for bowel transit time and to determine accuracy of faecal collections (Cummings 1978 [65]; Warner 1981)

<table>
<thead>
<tr>
<th>Marker</th>
<th>Use</th>
<th>Advantage</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carmin and chromic oxide (Whitby and Lang 1960 [349]; Rose 1964 [350]; Branch and Cummings 1978 [65])</td>
<td>Single dose with a meal</td>
<td>Simple and cheap</td>
<td>Measures only head of meal transit, and so underestimates true mean transit time. Inaccurate for faecal collections</td>
</tr>
<tr>
<td>Radio-opaque pellets (Hinton et al. 1969 [91]; Cummings et al. 1976a [76]; Cummings et al. 1976b [351]; Casanovas et al. 1991 [352]; Fotherby and Hunter 1987 [99]; Bouchouca et al. 1995 [353])</td>
<td>Single or multiple doses</td>
<td>Continuous dosing is gold standard method. Cheap and accurate for both transit and faecal collections Can measure regional transit in large bowel</td>
<td>Requires several faecal collections or x-ray</td>
</tr>
<tr>
<td>Isotope labelled test meal (Krevesky et al. 1986 [354]; Read et al. 1986; Metcalf et al. 1987 [92]; Bennink et al. 1999 [355]; Gryback et al. 2002 [356])</td>
<td>Single dose</td>
<td>Useful for tracking meals through both small and large bowel. No faecal collections</td>
<td>Radiation exposure. Cost. Requires scanning equipment</td>
</tr>
</tbody>
</table>

---

Markers used for bowel transit time and to determine accuracy of faecal collections (Cummings 1978 [65]; Warner 1981)
proof that transit determines metabolism, the urinary excretion data suggest strongly that transit exerts a controlling influence.

Finally, it is worth remembering that transit reflects total transit through the gut. The colon, like the rest of the gut, may functionally divide itself into a metabolic area (caecum and ascending colon) and transport, absorptive and storage area (transverse, descending and sigmoid colon). Transit through the metabolic region may be the important factor controlling colonic metabolism rather than transit through the rest. Transit through the right colon is only about one third of total colonic transit [90]. Measurement of total transit will tend to obscure any relationship between colonic events and transit; so methods of studying regional colonic function are also needed.

Measurement of transit

A variety of methods for measuring transit time have been described, but all are essentially variations on two themes (see Table 10). Transit time may be measured by giving a single dose of a marker substance, usually radio-opaque pellets, collecting faeces until all are recovered, and calculating the mean time the markers have taken to traverse the gut. Alternatively, markers may be given with every meal, and faeces collected until a steady state is reached (where input is approximately equal to output of markers), when the number of markers retained in the gut is a measure of transit time (number of markers retained divided by marker input rate equals transit time) [76, 91].

To avoid collecting faeces, abdominal radiographs can be taken to assess marker retention and distribution. This has clear advantages for the subject and allows transit time in the different segments of the colon to be determined but radiation exposure limits the wide applicability of this technique [71, 92–96]. Another transit method that involves some exposure to radiation is the use of a radio-labelled test meal. This can be followed through the gut using a gamma-camera and gives useful information about transit through different regions of the gut, including the small intestine. Furthermore, it does not require faecal collection [97] (Table 9).

One of the earliest methods described is that of Hinton and co-workers [94], in which a single dose of 20 radio-opaque pellets is given to a subject and the appearance of 80% of these in the stool taken as the transit time (80% TT). The 80% TT is an arbitrary and also poorly reproducible measure of transit which is about 1.16 times a true mean transit time [76]. It is imprecise because it relies on the excretion of less than the total markers and is particularly vulnerable to any change in defaecation, which affects the passing of the 16th pellet. It also suffers, along with other methods, by relying on a single dose of marker to estimate transit, in that it gives transit time at a single point in time when, in fact, in a given individual transit varies from week to week. Dissatisfaction with these methods led to the development of a new way of measuring transit in which a small dose of marker is given to subjects with each meal continuously over a period of weeks and its excretion estimated in the faeces. The amount of marker retained in the gut at any time may be calculated and, hence, turnover or mean transit time [76]. A simple method suitable for clinical studies and which obviates stool collection is to give the patient ten radio-opaque pellets daily with breakfast for 14 days, and then take one plain abdominal radiograph and count the number of pellets retained. Transit (days) is calculated from the number of pellets retained divided by ten. The normal range of transit is 1–4 days [98, 99]. A 14-day period is chosen because it takes at least twice the transit for a patient to reach a steady state; otherwise transit will be underestimated. The method gives an absolute value for transit. Its accuracy can be improved by giving more frequent doses of marker, and by taking more radiographs and averaging the results. Alternatively, a single dose of markers is given and faeces collected daily until all the markers are recovered [76], or daily abdominal radiographs are taken [93, 100, 101]. Various mathematical models have been suggested for calculation of the results [101–103].

X-raying the patient also allows segmental transit time in the colon to be measured. Using bony landmarks in the spine and pelvis, the number of markers present in the right, left and pelvic colon can be counted and a calculation of segmental transit time made. This is particularly useful in identifying hold-up of markers in the pelvic region in disordered defaecation. If possible, subjects should be asked to eat their usual diet and not take any laxatives during these studies [104].

For population studies transit can be measured by collection of a single stool on day 4 after giving a dose of radio-opaque markers of different type on days 1–3 [76]. The accuracy of this method can be improved by giving either more doses of marker, or collecting two stools [105]. An abdominal x-ray is also valid, but not suitable for population studies.

Normal values

In the United Kingdom, mean transit time (MTT) is 70 h but the distribution is very skewed and median MTT is 60 h, with men 55 h and women 72 h. The range (95%) is 30–168 h. Transit time in other countries is faster, particularly in Africa, where 24–48 h is the norm [57]. However, differences in methods used make comparisons between studies difficult.

Transit time varies greatly from day to day and from week to week in individuals. In a study of daily transits measured over 12 weeks in two healthy subjects, the coefficient of variation was 20–25% for the two, with a
range of 64–157 h in subject 1 and 25–70 h in subject 2. Such variation is presumably due to the same factors that influence bowel habit, though much is still unknown about this important parameter of colonic function [98].

MTT and stool weight are closely related [57,103,106,107]. At transit times of 80 h or more, stool weight is low but increases as transit time decreases to 40–50 h, after which changes in transit time are not associated with particularly great changes in stool weight. Transit determines stool weight, not vice versa.

Transit times through different regions of the colon can be measured using oral doses of radio-opaque markers or by isotope scanning and taking radiographs of subjects at intervals. Reported times are 7–24 h for the right colon, 9–30 h for the left colon, and 12–44 h for the sigmoid colon and rectum [92, 95, 96, 101, 102]. In a healthy individual, transit time is 62 h and the markers are distributed fairly evenly around the colon, giving segmental transit times for the right colon of 19 h, left colon 38 h, and rectosigmoid colon 5 h.

Constipation

Constipation is a disorder of motor activity of the large bowel. For more detailed information, any of a number of reviews should be consulted [104, 108–110]. Although traditionally defined in terms of bowel regularity, stool consistency and weight, in practice the main symptom in constipation is straining on defaecation. In addition, abdominal discomfort, distension and incomplete rectal emptying are all considered part of the condition [53, 60, 111]. Constipation may present as a very infrequent bowel habit, some patients passing a stool only weekly or even monthly. When people are questioned about their bowel habit, however, a number will complain of constipation but nevertheless pass a stool daily [53, 60, 112]. Stool frequency is thus not an infallible guide to the diagnosis of constipation. Clinically, a change in bowel habit is equally as important as defining defaecation pattern frequency.

Despite the belief that constipation is characterised by low stool weight (among other criteria), stool collections are seldom made in constipated subjects. In a meta analysis of data from eight groups of subjects who complained of constipation, mean daily stool weight for the whole series (n = 209) was 48.6 g/d [113]. As some of these subjects, chosen from various clinical trials, probably represent the more extreme end of the spectrum of constipation, the true mean may be higher. Thus, stool weight in constipation is low, but the range overlaps substantially with what is generally regarded as normal. It is not surprising, therefore, that constipation is a common complaint in many countries.

Total gut transit time is prolonged in many patients with constipation [99, 114]. Absolute values are not readily available, however, because the many techniques using radio-opaque pellets adopt widely differing criteria for arriving at transit time value. Thus, the dividing point between normal and constipated whole gut transit times, reported as between 67 and 120 h, is dependent on the method of measurement and the population studied [95, 103, 115, 116]. Another difficulty in setting a dividing line between normality and constipation is that many constipated patients have transit times well within the normal range whatever method is chosen [71, 116, 117]; a reasonable cut-off point is probably 120 h.

Studies of transit through the major regions of the large intestine have proved useful in distinguishing different types of constipation. Several patterns have been described, including generalised slowing throughout the whole colon (colonic inertia, slow transit constipation) seen in young women and in response to drugs or systemic disease [118]. Right-sided slowing is most often seen in ulcerative colitis [119] and in some lesions of the central nervous system [120]. Delay on the left side or in the rectosigmoid region is associated with disorders of defaecation and has been called ‘outlet obstruction’ [101, 116, 121].

Measurement of colonic transit is probably the simplest and most reliable way of assessing an individual’s constipation. Measurements of transit also allow identification of patients who do not have a slow rate of passage, and who may even have rapid transit, if taking laxatives covertly.

Risks of low stool weight and constipation

Burkitt and colleagues popularised the view that low stool weight is associated with certain western diseases [57] especially bowel cancer [86], diverticular disease [122], appendicitis [123] and various anal conditions [124]. He reported stool output, and transit time, in several population groups and related these data to national mortality and morbidity statistics where available, and to his own records. In general stool weights below about 150 g/d and slow transit of more than 4–5 days were associated with greater risk of bowel disease. Is low stool weight, therefore, a risk factor for bowel or other diseases?

Experimentally induced constipation leads to irritable bowel-like symptoms [125]. When healthy subjects are put on NSP-free diets constipation is a common complaint where stool weight falls below about 40–60 g/d [126]. If constipated subjects are given laxatives to speed up transit the amount of deoxycholic acid in bile falls (25.9 ± 8.6% constipated to 17.2 ± 8.3% + laxative) and when healthy subjects have symptomatic constipation induced and transit time slowed from 48 to 103 hours the deoxycholic acid pool
size increases significantly [127]. High deoxycholic acid levels are a risk factor for gallstones [128].

A number of anal problems such as haemorrhoids, and fissure are generally accepted as being related to constipation although few data are reported. Hard stools cause straining which may also lead to stretching of pelvic floor muscles and neuropathy with disordered ano-rectal function [129]. There is some evidence linking abnormal cells in breast ducts with constipation [130].

Diverticular disease of the colon, which affects around one-third of the elderly population is characterised by low stool weights and slow transit time in most cases [105, 131–141]. In ten publications of UK patients, mostly with symptomatic diverticular disease where stool weight has been measured, the average is 95 g/d (n = 285) [131–140]. This may, of course, be a result of the disease rather than its cause.

There is also a significant relation between stool weight and colon cancer. Stool weights of around 100 g/d are associated with a high risk of colon cancer (25/100,000 of the population) and, other factors being equal, need to be around 150 g/d to cut risk by about 50% (around 10/100,000) [52] (Fig. 1).

Defining a healthy bowel habit

What, therefore, would constitute a significant beneficial change in bowel habit? Since the response of bowel habit to carbohydrates is always, in individuals, proportional to their initial stool weight [81] then any effect should be expressed as a percentage or proportion of the group mean. Small changes of around 10–15%, are very difficult to detect and will be unnoticed by the subjects. Increases of 25% are measurable, but can they be said to be beneficial? Bowel habit is undoubtedly related to disease risks, especially that of bowel cancer (Fig. 1). To obtain any protection from colorectal cancer, stool weight needs to be above 150 g/day. For most EU populations, this would mean an increase in mean daily stool weight of about 50% from 100 g/day. For this population, therefore, a 25% increase might, therefore, be described as significant, 50% as beneficial. For transit time the data are more difficult to interpret because of the high dependence on methodology. Any dietary change that brings transit time into the normal range of 1–4 days should be seen as beneficial. Whether changes within this range are also beneficial is not easy to ascertain. However, they will be associated with changes in stool weight, which are more reliably related to health.

Composition and activities of the gut microbiota

The gastrointestinal tract exists in symbiosis with very great numbers of bacteria that contribute substantially to normal digestive function. The human gut is sterile at birth, when microbial colonisation begins during delivery. The inoculum is derived largely from the mother's vaginal and faecal flora in a conventional birth or from the environment in a caesarean delivery. Initially, facultative bacterial strains, such as *Escherichia coli* or streptococci, are transferred. These nutritionally undemanding bacteria create a reduced environment which allows the development of strictly anaerobic species that later dominate the gut [142]. Microflora development is then dependent on the type of feeding regime given in early life [143]. The breast-fed infant has a predominance of bifidobacteria, which easily outcompete other genera and are thought to depend on the occurrence of certain glycoproteins in human breast milk. In contrast, the formula-fed infant has a more complex flora which resembles the adult gut in that bacteroides, clostridia, bifidobacteria, lactobacilli, Gram positive cocci, coliforms and other groups are all represented in fairly equal proportions. During weaning, the microbiota becomes more developed and the ecosystem is thought to be fairly stable at around 2 years of age.

The entire gastrointestinal tract is populated by micro-organisms. However, the numbers and species composition varies greatly according to the region. In the oral cavity a particularly complex microbiota exists [144]. Bacteria can be found on the posterior and anterior tongue, sub- and supragingival plaque, buccal mucosa and vestibular mucosa [145]. These include members of the *Prevotella*, *Porphyromonas*, *Peptostreptococcus*, *Bacteroides*, *Fusobacterium*, *Eubacterium* and *Desulfovibrio* genera. The stomach is not heavily colonised due to its low pH, and typically harbours less than 10^3 CFU/g, mainly lactobacilli, streptococci and yeasts [146, 147]. The lumen of the human stomach is essentially sterile, therefore, except after meals. However, microorganisms are known to reside in the mucosal
layer that overlies the gastric epithelium. These include *Helicobacter pylori*, which has attracted a great deal of research interest. The organism uses its flagellae to invade gastric mucus and thereafter adhere to epithelial cells. This, in conjunction with production of ammonia, allows its effective long-term survival in the stomach [148]. An epidemiological link between *H. pylori* carriage and peptic ulcers, type B gastritis and, to a less proven extent, stomach cancer has been hypothesised.

The duodenum also has low microbial populations due to its short transit time and the secretion of biliary and pancreatic fluids, which create a hostile environment. However, there is a progressive increase in both numbers and species along the jejunum and ileum, from approximately $10^4$ in the jejunum to $10^{6-7}$/g of contents at the ileo-caecal junction [143] with the appearance of Gram-negative facultative organisms and obligate anaerobes [147].

The colon is the most heavily populated area of the gastrointestinal tract, with numbers typically in the region of $10^{12}$/g of contents [13]. The environment is favourable for bacterial growth with a slow transit time, ready availability of nutrients and favourable pH. There are several hundred identifiable species present, and a significant number that cannot be cultivated by conventional methods [149]. The majority of bacteria are non-spore-forming anaerobes, of which the numerically dominant are *Bacteroides* spp. and *Bifidobacterium* spp., *Eubacterium* spp., *Clostridium* spp., *Lactobacillus* spp., *Fusobacterium* spp. and various Gram-positive cocci [143]. Bacteria present in lower numbers include *Enterococcus* spp., *Enterobacteriaceae*, methanogens and dissimilatory sulphate-reducing bacteria [13, 143]. Yeasts, including the opportunistic pathogen *Candida albicans*, are also present in the gut, although in healthy individuals counts do not exceed $10^4$/g faeces [146, 150]. The vast majority (> 90 %) of the total cells in the body are bacteria in the colon. It is thought that over 60 % of the faecal mass exists as prokaryotic cells [151].

Whilst bacteria in the upper gut may impact on immune function, the transit of food through the stomach and small intestine is probably too rapid for the microflora to exert a significant impact on digestion. However, transit slows markedly in the colon. With typical transit times around 70 h, and longer, colonic microorganisms have ample opportunity to degrade available substrates. These may be derived from either the diet or endogenous secretions [152]. Major substrates available for colonic fermentation are starches that for various reasons are resistant to the action of pancreatic amylases, dietary fibres like cellulose, pectins, xylans and, in lower concentrations, oligosaccharides, a variety of sugars and sugar alcohols. In addition, proteins and amino acids can be effective growth substrates for colonic bacteria, whilst bacterial secretions, lysis products, sloughed epithelial cells and mucins also make a contribution. However, diet provides, by far, the predominant source of nutrients with around 70–100 g/d of dietary residues being available to the colonic microbiota. These materials are degraded by a wide range of bacterial polysaccharides, glycosidases, proteases and amino-peptidases to smaller oligomers and their component sugars and amino acids. Intestinal bacteria are then able to ferment these intermediates to hydroxy and dicarboxylic organic acids, $\text{H}_2$, $\text{CO}_2$ and other neutral, acidic and basic end products [153, 154].

In its entirety the gut microflora act as an effective barrier against opportunistic and pathogenic microorganisms, and this 'colonisation resistance' is one of their most important functions. However, the gut flora itself can be sub-divided into benign, beneficial and potentially harmful species, although certain genera contain species belonging to both groups, e.g. bacteroides may be saccharolytic (beneficial) or proteolytic (potentially harmful). Bacterial metabolism can result in a number of advantageous effects, including the production of vitamins, modulation of the immune system, enhanced digestion and absorption, inhibition of harmful species and removal of carcinogens and toxins. Furthermore, bacterial metabolism results in the production of short chain fatty acids (SCFA) upon which the colonic mucosa is dependent. Negative effects include the production of toxins and carcinogens, constipation or diarrhoea, liver damage, predisposition towards gut disorder and intestinal putrefaction [13, 147, 155].

As such, the use of dietary intervention has much promise for reducing the risk of disease onset – by fortifying components of the flora seen as beneficial (e.g. bifidobacteria, lactobacilli) and/or disturbing the more detrimental forms (e.g. clostridia, sulphate-reducing bacteria). A healthy, or balanced, flora is, therefore, one that is predominantly saccharolytic and comprises significant numbers of bifidobacteria and lactobacilli (see below). The exact numbers are difficult to give at present because a proportion of the gut flora have yet to be identified.

### Methodologies

The easiest way to determine how bacterial substrates such as prebiotics are metabolised, or how probiotics grow, is to use pure cultures of selected microorganisms [156]. For gut microorganisms, the approach involves something of a challenge in that anaerobic growth conditions must be induced along with standard microbiological aseptic techniques. The usual way to monitor the growth response is by measuring the optical density of the culture or by viable counting. This gives a reasonable comparative assessment of metabolism in mono-culture, but does not include any element of competition, which is prevalent in the gut ecosystem.
To mimic more closely the gut environment, a common approach towards the determination of gut microbial activities, is to use anaerobic batch culture fermenters inoculated with faecal bacteria [157]. However, these are closed systems where the substrate is limited, so are only appropriate for short time course experiments.

A more physiologically relevant approach is continuous culture, whereby a constant input of nutrients is supplied and other physiological parameters controlled [158]. Semi-continuous culture is a variation of this technique in which medium is added and spent culture removed at specific intervals. The most physiological approach is the chemostat where a continuous supply of growth medium is fed to the culture. The drawback is that the one stage continuous culture chemostat is a homogeneous system and varying physicochemical determinants cannot be imposed.

The human gastrointestinal tract is a heterogeneous microbial ecosystem. As such, an efficient gut model system would mimic these different physicochemical parameters. One model is the three-phase chemostat based on gut model simulation in different anatomical areas of the large gut, such as the right, transverse and left sides. This system has been validated against samples taken at autopsy and gives a very close approximation to fermentative bacterial events that occur in situ [159]. Five-stage continuous fermenters have also been used to simulate the intestinal tract from the jejunum to the descending colon [160]. Another popular model [159] is the “gastrointestinal simulator” which attempts to mimic the absorptive, degradative and microbial interactions of the alimentary tract. Such gut models have been applied to probiotic and prebiotic research and give useful mechanistic data.

Animals, often rats or mice, have been used to determine probiotic and prebiotic effects [161]. Conventional, gnotobiotic (germ-free) rats or those inoculated with one or a limited number of species of microorganisms may be used to investigate probiotic and prebiotic interactions, although this does not resemble the usual situation in the human gut. Rats, known as human flora associated (HFA) rats, may also be inoculated with a human faecal flora, and give a further representation of the situation in the human intestine, although the gut physiology is not the same. One drawback with laboratory animal experiments is the differing (gut) anatomy as well as coprophagy and the much greater food intake relative to body weight.

The definitive assessment of an effect on the gut flora is to feed candidate substrates or food to human volunteers and assess resultant changes in well being, stools, urine or blood markers [162].

For a microbial assessment, faeces is the only readily accessible area of gut contents, but study of faeces does not predict fermentation reactions in more proximal gut contents or at the host-microbe interface. Another major drawback of microbial assessments of faeces is that conventional cultural bacteriology dictates that freshly voided samples are processed. This is because gut bacteria react differently to storage, thus grossly affecting characterisation traits. It is not possible to derive a true quantitative assessment from culturing of stored stools. In this case, discriminatory techniques that involve a molecular biological approach towards bacterial characterisation are of much value. Such technology is highly reliable and allows the processing of frozen specimens and therefore multiple centre trials. These molecular techniques are also applicable to studies on gut mucosal biopsy material, which gives a useful indicator of interactions at the host-mucosal layer.

Gut microbiology is usually carried out by plating faecal microorganisms onto selective culture media designed to recover numerically predominant groups. However, the agars used are only semi-selective, do not recover non-culturable bacteria (which may represent over 50% of the overall diversity) and allow operator subjectivity in terms of microbial characterisation – which is usually based on limited phenotypic procedures. As such, alternative mechanisms, based around molecular principles, to characterise more effectively the microflora involved in fermentation studies, have been developed [163–168]. These involve gene sequencing procedures for colonies that grow on agars as well as molecular techniques that are culture independent (Table 10).

In conclusion, human trials are preferred for monitoring gut benefit related diet claims. However, these may be supported by mechanistic data, perhaps derived from model fermenter systems. Microflora analysis should be carried out using the high throughput and reliable molecular approaches that are now available. Moreover, the entirety of the gut flora diversity should be analysed, as far as is possible. It is not useful to claim effects on the basis of the recovery of a restricted number of microbial species from faeces.

Types of probiotics and prebiotics used to modulate the gut flora

The most frequently used dietary method of influencing the composition of gut flora is that of probiotics. Over the years many species of microorganisms have been used. They consist mainly of lactic acid producing bacteria (lactobacilli, streptococci, enterococci, lactococci, bifidobacteria) but also Bacillus spp. and fungi such as Saccharomyces spp. and Aspergillus spp. The most frequently used probiotics belong to the genera Lactobacillus (e.g. L. casei, L. acidophilus, L. rhamnosus, L. plantarum, L. johnsonii, L. reuteri) and Bifidobacterium (e.g. B. bifidum, B. lactis, B. longum, B. breve). New probiotic strains are being isolated regularly [169–171]. Probiotics
are commercially available for consumption, either as freeze dried preparations such as tablets, sprays, capsules or powders, or as foods, e.g. in dairy products, fruit juices.

Any dietary component that reaches the colon intact is a potential prebiotic. However much of the interest in the development of prebiotics is aimed at non-digestible oligosaccharides such as FOS, TOS, IMO, xylooligosaccharides (XOS), soyoligosaccharides (SOS), GOS and lactosucrose. Oligosaccharides are sugars consisting of between approximately 2 and 20 saccharide units, i.e. they are short-chain polysaccharides [172]. Some occur naturally in several foods such as leek, asparagus, chicory, Jerusalem artichoke, garlic, globe artichoke, onion, wheat, banana and oats [173], as well as soybean. However, these foods contain only small amounts of prebiotics so developments in functional foods have taken the approach of removing the active ingredients from such sources and adding them to more frequently consumed products in order to attain levels whereby a prebiotic effect may occur, e.g. cereals, confectionery, biscuits, infant feeds, yoghurts, table spreads, bread, sauces, drinks, etc. FOS, GOS and lactulose have been shown to prebiotics, through numerous volunteer trials, as evidence by their ability to change selectively gut flora composition after a short feeding period [173].

A new development is that of synbiotics. Here, a useful probiotic is incorporated into an appropriate dietary vehicle with a suitable prebiotic. The premise is that the selected substrate is metabolised by the added probiotic in the gut. This should enhance probiotic survival, as well as offer the advantages of both gut microbiota management techniques. A synbiotic has been defined as ‘a mixture of probiotics and prebiotics that beneficially affects the host by improving the survival and implantation of live microbial dietary supplements in the gastrointestinal tract.’ One example is a mixture of probiotic bifidobacteria with prebiotic FOS.

### Assessment of effect

The expansion on the market of probiotics and prebiotics is easily explained by the various health benefits ascribed to their intake. It is important that these claimed health benefits are proven and are preferably underpinned by identification of the mechanisms of the effect [174]. This requires well controlled clinical trials that use up to date methodologies such as molecular typing of gut flora changes in response to diet. For gut flora effects various claims have been made. These normally refer to a beneficial influence upon the microbiota composition as measured in faecal specimens. Several possible bacterial defence mechanisms exist such as improved immune function, excretion of acids, production of specific antimicrobial agents, competition for nutrients and colonisation sites, anti-adhesive strategies and attenuation functions.

In human studies, changes in the gut flora composition can be assayed in faeces but should include technologies that encompass the non-culturable flora (Table 11). These could be linked to an assessment of functional biomarkers relevant to the claim being made. Examples include the detection of organic acids in blood or faeces (although most SCFA formed by the gut flora are metabolised in situ), enzymes, genotoxic markers, toxins, immune relevant molecules and blood lipids. For the gut flora itself, it is likely that discernible changes are required for health benefits to accrue. For example, with a typical faecal lactobacilli count of around 10^5/g faeces or bifidobacterial enumeration of about 10^6/g, a one log increase in response to a test material may be expected. In some cases where the indigenous counts are higher, then lower numerical increases are likely to occur. It should be borne in mind that ostensibly lower increases from a higher starting level may actually represent a large stimulation in absolute numbers (Table 11).

Probiotic recovery from faeces is often used as an in-VI/134 European Journal of Nutrition (2004) Vol. 43, Supplement 2

---

**Table 10** Summary of molecular techniques that may be applied for discrimination of gut bacteria (from McCartney and Gibson 1998 [357])

<table>
<thead>
<tr>
<th>Molecular Technique</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restriction fragment length polymorphism (RFLP)</td>
<td>Compares banding patterns after digestion of chromosomal DNA with endonucleases</td>
</tr>
<tr>
<td>Pulsed field gel electrophoresis (PFGE)</td>
<td>Cutting restriction enzymes are used to reduce band numbers from RFLP</td>
</tr>
<tr>
<td>Ribotyping</td>
<td>A ribosomal DNA probe highlights bands within RFLP</td>
</tr>
<tr>
<td>Ribosomal DNA sequencing</td>
<td>Direct comparison of 16S rDNA gene sequences from sequencing of PCR amplified rDNA</td>
</tr>
<tr>
<td>Amplified fragment length polymorphism</td>
<td>RFLP amplified by polymerase chain reaction (PCR)</td>
</tr>
<tr>
<td>Direct amplification</td>
<td>Culture independent PCR amplification of bacterial DNA in environmental samples</td>
</tr>
<tr>
<td>Genetic probes</td>
<td>Detection and/or identification of specific microbial groups within environmental samples by labelled hybridisation probes</td>
</tr>
<tr>
<td>Molecular marking</td>
<td>Use of a genetic tag that enables discrimination of target micro-organisms within complex ecosystems</td>
</tr>
</tbody>
</table>
dicator of ‘colonisation.’ A preferable approach is to assay ‘persistence’ by detecting the number of organisms present in faeces following a feeding period. This brings in the issue of discriminating the strain fed from the indigenous flora. Cultural methodologies are not refined enough for such a high level of diagnostic microbiology. However, it is possible to design genotypic probes that specifically hybridise to a unique sequence(s) of the 16S rRNA of the target probiotic. An oligonucleotide sequence specific to the probiotic strain would facilitate the use of conventional polymerase chain reaction (PCR) amplification techniques and exploitation of the probing strategy. For this, a specific oligonucleotide probe would be labelled with a fluorescent protein so as to mark the probiotic strain, thus facilitating differentiation from commensal organisms. This would remove the need, firstly to isolate all colonies on agars used to enumerate these genera, (which tends not to be particularly specific for the intended organisms), and then subject each colony type to DNA characterisation. Fluorescently-labelled oligonucleotide probes targeted at the 16S rRNA sequences specific for a bacterial genus have been used by Welling et al. [175] to characterise the composition of faecal bacteria. However, species or strain specific probes are limited and the use of an introduced visible marker for the probiotic strain may be necessary.

On present evidence, a prebiotic dose of 5 g/d is sufficient to elicit a positive effect upon the gut microbiota (in some cases this may be nearer to 8 g/d) [175]. Doses below 5 g/d are not likely to be effective as the overall dietary carbohydrate load into the large bowel will mask any prebiotic effect. A possible side effect of prebiotic intake is intestinal discomfort from gas production. However, bifidobacteria and lactobacilli do not produce gas as part of their metabolic processes. Therefore, even at a dose of 20 g/d, gas symptoms should not occur. If gas is being generated then the carbohydrate is not acting as an authentic prebiotic. This is perhaps because dosage is too high and the prebiotic effect is being compromised, i.e. bacteria other than the target organisms are becoming involved in the fermentation.

### Natural gut defence

The intestinal mucosa exhibits a very large surface (considering the villus-crypt structure in an unfolded disposition, a flat extension of 400 m² is estimated), and constitutes a major interface with the external environment.
This interface is the principal site of interaction with foreign substances and microorganisms, which constantly challenge the mucosal surface. Whilst humoral and cellular immunity mechanisms play a key role in the defence of the gut, a primary line of defence is provided by non-immune structural and dynamic processes. There are three levels of non-immune defence: the gut flora, the mucosal barrier and gastrointestinal motility.

**The flora as a barrier**

Bacteria colonising the gut mucosa play an important role in host defence. Animals bred in a germ-free environment have increased susceptibility to infection. Disruption of the indigenous flora by antibiotic therapy is often associated with diarrhoea. Several factors are incriminated in the barrier effect provided by the flora. For example, commensal bacteria can secrete antimicrobial substances that inhibit the growth of pathogens, or compete with invading organisms for binding sites and nutrients [176]. The commensal flora also stimulates immune system function.

**Mucosal barrier function**

**Physical barrier**

The mucosal barrier is a complex structure that separates the internal milieu from the luminal environment. Physically, the barrier includes cellular and stromal components, from the vascular endothelium to the epithelial cell lining, and the mucus layer, which consists of a gel formed by the interaction of various mucosal secretions, namely, mucins, trefoil peptides and surfactant lipids [177, 178].

The intestinal epithelium is replaced continually by a process involving proliferation of stem cells and this rapid and constant epithelial cell turnover is essential for the barrier. The newly generated cells which arise at the junction of the villus and the crypt, migrate either to the top of the villus or to the base of the crypt. The cells differentiate into enterocytes, goblet cells or Paneth cells, with specific functions that are described below. Apoptosis and exfoliation occur at the top of the villi, so that the turnover of villus enterocytes is rapid; cells being replaced every 2–3 days. The re-population of crypts and villi is sustained by the rate of replication of stem cells. An increase in the rate of replication of stem cells may occur during pathogen infection. The luminal loss of epithelial cells together with exudation of fluid from the crypt results in a washing-away of pathogens adhering to the mucosa. This mechanism contributes to the protection of the epithelial surface from bacterial invasion.

Columnar epithelial cells are held together by tight junctions and adherens junctions. These structures comprise a complex of proteins, of which more than 40 have been described. ‘Occludin’ and ‘claudins’ are integral proteins capable of interacting adhesively with complementary molecules on adjacent cells. It has been shown in patients with inflammatory bowel disease that the expression of occludin is down-regulated, which could explain the enhancement of paracellular permeability and neutrophil transmigration [179]. Tight junction function can be affected by dietary factors [180].

Goblet cells secrete a combination of trefoil peptides and mucin glycoproteins to form mucus, a continuous gel that covers the epithelial surface [177, 181]. Mucins are glycoproteins in which O-linked glycosylated regions comprise 70–80% of the polymer. Mucin oligosaccharide chains are often terminated with sialic acid or sulphate groups. These proteins are classified into neutral or acidic mucins and the latter are further classified into sulfo- or sialomucins. Mucin subtypes vary spatially throughout the gastrointestinal tract. The secretion of mucins at the apical surface of goblet cells is constitutive by simple exocytosis but can also be induced by mucin secretagogues. The regulation of mucin genes also contributes to mucin heterogeneity and the dynamic nature of the mucus gel layer. Nine epithelial mucin (MUC) genes have been identified in humans [182]. The secretion of mucus is typically enhanced by both commensal bacteria [183] and intestinal pathogens like *Vibrio cholerae* and *Entamoeba histolytica* for example [182]. The defensive nature of mucins lies in their capacity to entrap microbes. Adhesion to specific mucins also facilitates colonisation of mucus by commensal bacteria.

Trefoil peptides are a family of peptides described in recent years. Characteristically, these peptides exhibit a trefoil domain that renders them resistant to digestion by proteases. They are among the most abundant products of the gastrointestinal mucosa and in conjunction with mucins confer the viscoelastic properties of mucus [181]. Trefoil peptides are also involved in the maintenance and repair of the intestinal mucosa. In fact, mice deficient in ITF, a member of the intestinal trefoil peptide family, have impaired mucosal healing after induction of experimental colitis [184].

The mucus layer provides a physical separation between lumen and epithelium and an important framework for host-bacteria and bacteria-bacteria interactions. The strong hydrophobic surface of the mucus layer prevents the influx of water-soluble toxins into the epithelium [185]. Surface hydrophobicity is attributable to a layer of surfactant lipids which are secreted by epithelial cells and align along the top of the mucus gel [178]. In several mammalian species including humans, surface hydrophobicity is very high on top of the gastric and colonic mucosa, whereas it is much lower throughout the small intestine, which is the absorptive surface.
Thickness of the mucus layer can reach up to 450 micrometer in the stomach. The small intestine is covered with a much thinner mucus layer, which is discontinuous above the Peyer's patches. Beyond the ileo-caecal valve, hydrophobicity and mucus layer thickness increase gradually from the ascending colon to the rectum [177, 185]. Ample experimental evidence shows that the mucus layer is an important defence mechanism protecting the mucosa from injury as well as facilitating repair after injury has occurred.

Chemical barrier

■ Digestive secretions. Gastric acid facilitates digestion of nutrients and enhances iron and calcium absorption. A low pH (<3.0) confers bactericidal properties on the gastric juice. Several studies have clearly demonstrated the antimicrobial activity of gastric juice in vivo. For example, the enteric multiplication rates of Shigella flexneri increase three fold after sodium bicarbonate neutralisation of gastric juice (as reviewed in [176]). Patients with achlorhydria or hypochlorhydria are much more susceptible to infectious agents by oral route.

Pancreatic secretions also play a role in the gut defence barrier. Patients with pancreatic insufficiency are more susceptible to acute diarrhoeal infections. It has been demonstrated that pancreatic juice has bacteriostatic properties [176]. Deconjugated bile acids also exert an inhibitory effect on microbial growth in vitro.

■ Antimicrobial peptides. An important mechanism of innate defence is linked to the production of endogenous antimicrobial peptides. These peptides are secreted into the lumen both by enterocytes and Paneth cells, the best characterised of which are the alpha-defensins of the small intestinal Paneth cells [186, 187]. Paneth cells develop early after birth, giving the newborn immediate resistance to environmental and pathogenic microorganisms. Considerable amounts of these cells are found in the small intestine where the number of bacteria is low when compared to the colon. Paneth cells are found at the base of the intestinal crypts. They originate from intestinal stem cells located at the interface of the villus and the crypt and migrate downward to the bottom of the crypt. On average, 5 to 15 Paneth cells are found in the base of each crypt. It is thought that secretion of antimicrobial peptides by Paneth cells contributes to the protection of stem cells.

Human Paneth cells secrete products that are homologous to those found in phagocytic leukocytes (see below). Human and murine Paneth cells produce lysozyme, secretory phospholipase-A2, DNAses, trypsin and alpha-defensins, the latter being the most abundant proteins produced by these cells [186, 187]. The defensins produced in the mouse intestine, so-called cryptdins, and the human small bowel defensins (HDS and HD6) belong to the alpha-defensin family. Three subfamilies of defensins are known: the alpha-, beta- and theta-defensins. The alpha- and beta-defensins are characterised by a different spatial distribution of three intramolecular disulphide bonds. The single theta-defensin identified so far is a circular peptide [188].

Defensins are stored with other antimicrobial peptides in apical cytoplasmic granules. Secretion may be stimulated by Gram-negative and Gram-positive bacteria, and by bacterial products (lipopolysaccharide, lipoteichoic acid, muramyl dipeptide) [189]. In addition, stimulation of other degranulation pathways includes cholecystokinin, cholinergic agonists and cytokines such as TNF-α and Interferon-alpha may occur. The expression of several inducible genes of epithelial antimicrobial peptides from mammals is regulated by the NF-κB pathway.

Alpha-defensins are stored as propeptides and need to be modified prior to becoming fully functional. In mice, the metalloproteinase matrylsin is required to cleave the propeptide in active form. As recently reported by [190], trypsin is the enzyme that cleaves human defensin-5. Trypsin is also stored in Paneth cell granules in an inactive form. The alpha-defensins have a wide acting spectrum and their antimicrobial activity is effective against both Gram-positive and Gram-negative bacteria.

The way by which antimicrobial peptides specifically target bacteria is based on their structural properties. Many antimicrobial peptides exhibit some domains with cationic amino acids and other domains with hydrophobic amino acids. The positively charged domains interact with negatively charged phospholipids of bacterial membranes. These electrostatic interactions lead to either a fatal depolarisation, production of physical holes or scrambling of the usual distribution of lipids that disrupt the bacteria cell membrane. In contrast to bacteria, the outer leaflet of the cell membrane in plants and animals is composed principally of lipids with no net charge [191].

■ The intestinal epithelial cell as a sensor of the environment

Intestinal epithelial cells are in close contact with luminal contents and play a crucial role in signalling and mediating host innate and adaptive mucosal immune responses. As described later, activation of innate host defence mechanisms is based on the rapid recognition of conserved pathogen-associated molecular patterns (PAMPs) of microbes. Epithelial cells express both Toll-like receptors and NOD-family receptors that bind to PAMPs. In response to invading bacteria, the signals converge on the transcription factor NF-κB, which activates transcription of genes responsible for the synthe-
sis of proinflammatory proteins. Hence, epithelial cells secrete mediators, including chemoattractants for neutrophils and proinflammatory cytokines, and express inducible enzymes for the production of nitric oxide, prostaglandins and leukotrienes. Intestinal epithelial cells also express MHC class II and non-classical MHC class I molecules, suggesting that they can function as antigen-presenting cells [179]. Taken together, epithelial cells produce the essential signals for the onset of mucosal innate responses and recruitment of appropriate cell populations.

### Assessment of mucosal barrier function

An efficient gut mucosal barrier should prevent the passage of viable bacteria from the gastrointestinal lumen through the epithelial mucosa (bacterial translocation). Translocation of viable or dead bacteria in minute amounts constitutes a physiologically important boost to the immune system. However, dysfunction of the gut mucosal barrier results in translocation of a conspicuous quantity of viable microorganisms, usually belonging to Gram-negative aerobic genera, that may travel via the lymph to extraintestinal sites. However, the mesenteric lymph nodes, liver, spleen and lungs all help to prevent dissemination throughout the body and sepsis. Bacterial translocation and its complications have been shown clearly to occur in humans in several pathologic conditions such as postoperative sepsis, acute severe pancreatitis, advanced liver cirrhosis, multisystem organ failure, etc. [192]. In these conditions, increased rates of bacterial translocation have been demonstrated by positive culture of mesenteric lymph nodes. This approach to investigate bacterial translocation is invasive and not suitable for physiological human studies, so that indirect methods are needed to evaluate this phenomenon. For instance, bacterial DNA detection and identification in serum samples was proven to be a sensitive non-invasive method to demonstrate translocation (Table 11) [193].

Overall competence of mucosal barrier function can be assessed by intestinal permeability studies. Non-invasive methods used in human studies are based on the measurement of the urinary excretion of orally administered test substances (Table 11). The ideal probe should be inert, absorbed by passive diffusion, not metabolised and excreted unchanged in the urine in easily measurable form [194]. Probes of different molecular size are commonly used together (e.g. lactulose/mannitol test), so that transcellular and paracellular permeability pathways can be evaluated at the same time [194]. Dysfunction of the barrier is recognised by an increased urinary recovery of the paracellular probe (lactulose), and a reduced recovery of the transcellular marker (mannitol), and it is expressed as an increased lactulose/mannitol ratio. Since these techniques lack site specificity, colonic permeability is not specifically addressed. However, mucosal barrier function is particularly relevant at the colonic level, as the slow motility patterns of the colon allow bacterial proliferation at high densities and remarkably high concentrations of toxins to accumulate. Sugar probes of similar size that are selectively destroyed at different levels of the gut have been successfully used in the rat [195].

Mucosal barrier function can also be assessed ex vivo with specimens of intestinal mucosa (Table 12). Surface hydrophobicity is assessed by determination of the contact angle that conforms a microdrop of water deposited on the mucosal surface using a special goniometer [178, 185]. The contact angle is the angle between the solid surface and the tangent to the liquid-air interface at the triple point where all three phases meet (solid, liquid, air). Large contact angles are found on hydrophobic surfaces whereas small angles are formed on hydrophilic surfaces. The most commonly used method to investigate barrier function in vitro is the Ussing chamber. Isolated mucosal segments are mounted in a device containing two fluid reservoirs separated by the test mucosal specimen. The luminal surface is bathed by one reservoir and the basolateral interface is bathed by the other. The passage of specific substances, ions or markers (mannitol, EDTA or macromolecules such as horseradish peroxidase) across the mucosa can be investigated by measurements in the two reservoirs. Electrical potential difference across the epithelial sheet can be recorded and used for real-time assessment of barrier function [196].

### Table 12  Ex vivo methods used to assess the mucosal barrier function

<table>
<thead>
<tr>
<th>Method</th>
<th>Measurement</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface hydrophobicity¹</td>
<td>Contact angle of a microdrop of water deposited on the mucosal surface.</td>
<td>Straightforward technique.</td>
<td>Invasive; samples are obtained by endoscopy or at surgery.</td>
</tr>
<tr>
<td>Ussing chamber²</td>
<td>Passage of ions or markers (mannitol, EDTA, horseradish peroxidase) and electrical potential difference across the mucosa.</td>
<td>Most commonly used method for research purposes on mucosal barrier function. Allows real-time monitoring by recording electrical potential difference.</td>
<td>Invasive; surgical mucosal explants are needed (endoscopic biopsies are not suitable).</td>
</tr>
</tbody>
</table>

¹ See [185]; ² See [196]
Motility as a defence mechanism

Gastrointestinal motility is an important factor in the control of bacterial growth and proliferation within the intestinal lumen. Thus, the phasic and propulsive motor activity towards the ileal end, which is constantly taking place in the small bowel, impedes stable colonisation of the lumen. Only in pathological conditions with impaired small bowel motility can bacterial overgrowth be demonstrated within the small bowel lumen. In these cases, abdominal distension and malabsorption may occur. On the contrary, infectious gastroenteritis may induce rapid motility patterns in the small bowel associated with enhanced fluid secretion leading to diarrhoea. This response is regarded as a mechanism of defence.

The immune system: systemic and mucosal immunity

The body is under constant threat of attack by viruses, bacteria and parasites. Evolution has therefore provided mammals with numerous complex and potent layers of immunological defence. Microorganisms have inhabited Earth for at least 2.5 billion years, and the power of the immune system is a result of coevolution in which indigenous bacteria particularly have shaped the body’s defence functions [197, 198].

In humans, the critical role of the immune system, which in principle is partly open to the external environment, becomes clinically apparent when it is defective. Thus, inherited and acquired immunodeficiency states are characterised by increased susceptibility to infections, sometimes caused by commensal organisms not normally considered to be pathogenic.

The immune system can be divided into two general arms: innate (natural or non-specific) and adaptive (acquired or specific) immunity, which work together synergistically [199, 200]. The adaptive immune system developed late in the phylogeny, and most species survive without it. However, this is not true for mammals, which have an extremely sophisticated adaptive immune system of both systemic and mucosal (local) type. There appears to be great redundancy of mechanisms in both systems providing robustness to ensure that essential defence functions are preserved.

Innate immunity

The attempt of an infectious agent to enter the body will immediately be recognised and counteracted by the innate immune system, which comprises surface barriers (see above), soluble factors, specialized phagocytes and dendritic cells (DCs). Together, these functions constitute a primary layer of natural defence against invading microorganisms, with the common goal of restricting their entry into the body by providing: (a) physical/structural hindrance and clearance mechanisms such as epithelial linings of skin and mucosae, mucus, ciliary function and peristalsis; (b) chemical factors such as pH of body fluids, numerous antimicrobial peptides and proteins; and (c) phagocytic cells, e.g. neutrophils, eosinophils, monocytes/macrophages and DCs. Challenges to the innate system often lead to activation of the adaptive immune system, which aids substantially recovery from infection, as discussed below.

Triggering of innate immunity

The recognition molecules involved in innate immunity are encoded in the germline. This system is therefore quite similar among healthy individuals and shows no apparent memory effect – that is, re-exposure to the same pathogen will normally elicit more or less the same type of response. These receptors sense conserved molecular structures that are essential for microbial survival and are present in many types of bacteria, including endotoxin or lipopolysaccharide, teichoic acids and unmethylated CpG motifs of DNA [201]. Although such structures are generally called pathogen-associated molecular patterns (PAMPs), they also occur in commensal bacteria [202]. However, the intestinal microflora may induce distinct molecular programming of the innate immune system, which may explain why the indigenous microbiota is normally tolerated by the host [203, 204].

The cellular receptors of the innate immune system that recognise PAMPs as ‘danger signals’, are called pattern recognition receptors (PRRs), many of them belonging to the so-called Toll-like receptors (TLRs). They are expressed mainly by macrophages and DCs, but also by a variety of other cell types such as B cells and epithelial cells [202]. The engagement of PRRs causes cellular activation. In the case of professional antigen-presenting cells (APCs) such as DCs, this leads to maturation accompanied by production of cytokines and upregulation or downregulation of cell-surface molecules according to strictly defined kinetics [205]. Such signalling molecules will critically influence further induction of both innate and adaptive immunity with regard to effector potency, particularly the polarisation of T-cell responses in terms of cytokines (Fig. 2). This will be further discussed below under adaptive immunity.

In summary, there are both stereotypical and selective responses of innate host cells to different types of microorganisms. In this manner, pathogens can, early in an infection, imprint their ‘signatures’ on APCs and thereby on subsequent immune responses (Fig. 2). Thus, the plasticity of the innate immune system prepares the ground for a targeted and powerful protective function of the adaptive immune system [206].
cells such as immunoglobulin (Ig) Fc receptors and cytokines such as IFN-\(\gamma\), TNF-\(\alpha\), and IL-2, while activated Th2 cells are mainly capable of IL-4, IL-5, and IL-13 secretion. Distinct Th1 and Th2 profiles are further promoted by positive and inhibitory feedback loops as indicated. In addition, under certain unclear conditions, apparently immature APCs may induce regulatory T (Treg) cells, which by their cytokines IL-10 and TGF-\(\beta\) (or by direct cellular interactions) may suppress both Th1 and Th2 responses. Ag antigen; IFN interferon; IL interleukin; MHC II major histocompatibility complex class II molecules; TCR T-cell receptor; TGF transforming growth factor; TNF tumour necrosis factor

**Fig. 2** Decision-making to elicit appropriate adaptive immunity is determined by costimulatory signals from the innate immune system. Antigen-presenting cells (APCs) take up Ag and process (degrade) it to immunogenic peptides which are displayed to TCRs in the polymorphic groove of MHC molecules after their appearance at the cell surface. An immunological synopsis is formed between the APC and the T cell as indicated, usually resulting in cellular activation. When naive CD4+ T helper (Th) cells are activated by APCs that provide appropriate costimulatory signals (soluble factors and/or adhesion molecules), they differentiate into Th1 or Th2 cells with polarised cytokine secretion. Such skewing of the adaptive immune response depends on the presence of microenvironmental factors, including cytokines (see Fig. 4) as well as danger signals from microbial products, including endotoxins (lipopolysaccharide), lipoproteins and unmethylated CpG nucleotide motifs from microbes (Table 2). Signalling via PRRs and other receptors stimulates activation and functional maturation of APCs along different pathways and will thereby be a microbial “signature” dictate the provision of various costimulatory signals. Subsequent activation of Th1 cells leads to predominant production of cytokines such as IFN-\(\gamma\), TNF-\(\alpha\) and IL-2, while activated Th2 cells are mainly capable of IL-4, IL-5, and IL-13 secretion. Distinct Th1 and Th2 profiles are further promoted by positive and inhibitory feedback loops as indicated. In addition, under certain unclear conditions, apparently immature APCs may induce regulatory T (Treg) cells, which by their cytokines IL-10 and TGF-\(\beta\) (or by direct cellular interactions) may suppress both Th1 and Th2 responses. Ag antigen; IFN interferon; IL interleukin; MHC II major histocompatibility complex class II molecules; TCR T-cell receptor; TGF transforming growth factor; TNF tumour necrosis factor

Adaptive immunity

**T and B cells**

In peripheral blood, the lymphocytes comprise 20–25% of the leukocytes. Various lymphocyte subsets can be identified by the use of monoclonal antibodies (usually of mouse origin), which recognise specific proteins - that is, cellular markers known as cluster of differentiation (CD) molecules. Thus, all T lymphocytes (or T cells) express selectively CD3, and all B lymphocytes (or B cells) express selectively CD19 and CD20. A particular subset of T lymphocytes called T-helper (Th) cells express CD4, whereas most cytotoxic T cells express CD8. Adaptive immunity depends on the functional properties of T and B cells and is directed by their antigen-specific surface receptors, which show a random and highly diverse repertoire [199, 200, 207].

Lymphocytes originate in the bone marrow from a common lymphoid stem cell. Further development and maturation of T and B cells occur in the thymus and bone marrow, respectively (so-called primary lymphoid organs). Mature but yet unprimed (naïve) T and B cells enter the bloodstream and become disseminated to secondary lymphoid organs such as the spleen, lymph nodes and mucosa-associated lymphoid tissue (MALT). Certain adhesion molecules and receptors for chemokines (chemoattractant cytokines) enable adherence of immune cells to specialised vascular endothelium and their migration into the lymphoid organs, which are anatomically and functionally organised to facilitate interactions between lymphocytes and various types of APCs [210–212]. Antigens are carried into these immune-inductive structures from peripheral tissues via draining lymph, passively as soluble molecules and dead or live particles, and actively by migrating DCs, as well as directly from mucosal surfaces by ‘membrane’ or ‘microfold’ (M) cells in MALT (Table 13). Lymphocytes that do not encounter cognate antigen re-enter the bloodstream by way of efferent lymphatics and then the thoracic duct. The functional consequence of this recirculation of T and B cells is that all parts of the body are

---

**Natural killer cells**

Approximately 10–15% of peripheral blood lymphocytes are neither T nor B cells. Despite the fact that these previously so-called ‘null cells’ employ recognition mechanisms somewhat similar to T cells, they are considered to belong to the innate immune system and are therefore currently referred to as natural killer (NK) cells. The NK-cell receptors, which are encoded in the germline, recognise structures of high molecular weight glycoproteins expressed on virus-infected cells. After activation, NK cells release their granule content, such as perforin and cytolyisin, and kill virally-infected host cells and a variety of tumour cells without prior sensitisation [208]. There are also natural killer T (NKT) cells that exhibit features of both NK cells and highly specialised T-cell subsets (see below) with cytotoxic properties and polarised cytokine profiles [209].

**Fig. 2** Decision-making to elicit appropriate adaptive immunity is determined by costimulatory signals from the innate immune system. Antigen-presenting cells (APCs) take up Ag and process (degrade) it to immunogenic peptides which are displayed to TCRs in the polymorphic groove of MHC molecules after their appearance at the cell surface. An immunological synopsis is formed between the APC and the T cell as indicated, usually resulting in cellular activation. When naïve CD4+ T helper (Th) cells are activated by APCs that provide appropriate costimulatory signals (soluble factors and/or adhesion molecules), they differentiate into Th1 or Th2 cells with polarised cytokine secretion. Such skewing of the adaptive immune response depends on the presence of microenvironmental factors, including cytokines (see Fig. 4) as well as danger signals from microbial products, including endotoxins (lipopolysaccharide), lipoproteins and unmethylated CpG nucleotide motifs from microbes (Table 2). Signalling via PRRs and other receptors stimulates activation and functional maturation of APCs along different pathways and will thereby be a microbial “signature” dictate the provision of various costimulatory signals. Subsequent activation of Th1 cells leads to predominant production of cytokines such as IFN-\(\gamma\), TNF-\(\alpha\) and IL-2, while activated Th2 cells are mainly capable of IL-4, IL-5, and IL-13 secretion. Distinct Th1 and Th2 profiles are further promoted by positive and inhibitory feedback loops as indicated. In addition, under certain unclear conditions, apparently immature APCs may induce regulatory T (Treg) cells, which by their cytokines IL-10 and TGF-\(\beta\) (or by direct cellular interactions) may suppress both Th1 and Th2 responses. Ag antigen; IFN interferon; IL interleukin; MHC II major histocompatibility complex class II molecules; TCR T-cell receptor; TGF transforming growth factor; TNF tumour necrosis factor

Engagement of other types of receptor on phagocytic cells such as immunoglobulin (Ig) Fc receptors and complement receptors, triggers phagocytosis and elimination of invading microorganisms [207]. Although pathogens have evolved mechanisms to evade innate immunity (e.g., bacterial capsules), they can usually not persist within the body when an adaptive immune response reinforces innate immunity by providing specific antibodies directed against the invading pathogen or its toxins. Thus, the innate and adaptive immune systems are not independent; innate immunity influences the character of the adaptive response, and the effector arm of the adaptive response support several innate defence mechanisms.

Natural killer cells

Approximately 10–15% of peripheral blood lymphocytes are neither T nor B cells. Despite the fact that these previously so-called ‘null cells’ employ recognition mechanisms somewhat similar to T cells, they are con-
under continuous antigen-specific immunological surveillance.

Each T and B cell bears antigen receptors with a certain specificity, which differ between individual clones of lymphocytes [207]. A clone consists of daughter cells derived by proliferation from a single ancestor cell, so-called clonal expansion. The total population of T and B cells in a human may be able to recognise some $10^{11}$ different antigens. This remarkably diverse antigen receptor repertoire is generated during lymphocyte development by random rearrangement of a limited number of receptor genes. Even without priming, the adaptive immune system would thus be able to respond to an enormous number of antigens, but the detection of any single antigen could be limited to relatively few lymphocytes, perhaps only 1 in 1,000,000. Consequently, in a primary immune response there are generally an insufficient number of specific lymphocytes to eliminate the invading pathogen. However, when an antigen receptor is engaged by its corresponding antigen, the lymphocyte usually becomes activated (primed), ceases temporarily to migrate, enlarges (blast transformation) and proliferates rapidly so that, within 3–5 days, there are numerous daughter cells – each specific for the antigen that initiated the primary immune response.

Such antigen-driven clonal expansion accounts for the characteristic delay of several days before adaptive immunity becomes effective in defending the body. In addition to the effector cells generated by clonal expansion and differentiation, so-called memory cells are generated; these may be very long-lived and are the basis of immunological memory characteristic of adaptive immunity [207]. Functionally, immunological memory enables a more rapid and effective secondary immune response upon re-exposure to the same antigen. In contrast to innate immunity, the antigen recognition profile of the adaptive immune system reflects the individual’s lifetime exposure to stimuli from infectious agents and other antigens, and will consequently differ among individuals.

### Immune response and immune reaction

The purpose of adaptive immunity is primarily to combat infections by preventing colonisation of pathogens and keep them out of the body (immune exclusion), and to seek out specifically and destroy invading microorganisms (immune elimination). In addition, specific immune responses are, through regulatory mechanisms, involved in avoidance of overreaction (hypersensitivity or allergy) against harmless antigens as well as discrimination between components of ‘self’ and ‘non-self’. Autoimmunity occurs when the latter control mechanism breaks down [199].

It follows from the previous section that an adaptive immune response includes every aspect of cellular acti-
Immune protection and hypersensitivity

Although immune reactions are principally directed specifically against the antigen(s) that stimulated the adaptive response(s), extensive non-specific events will usually ensue, often appearing clinically and histologically as variations on the theme of inflammation. When the result is judged to be damaging, the underlying immune reactions are referred to as hypersensitivity, and the tissue-destructive effect is called immunopathology (Fig. 3). For didactic reasons, hypersensitivity mechanisms are categorised into four main types in accordance with the Coombs and Gell classification: type I (immediate type, IgE-mediated allergy); type II (cytolytic); type III (immune complex-mediated); and type IV (cell-mediated, DTH). These types rarely occur completely on their own, but with one type usually dominating in a particular immunopathological lesion [199].

It should also be emphasised that hypersensitivity is principally an expression of protective immune reactions that, however, become tissue-damaging mainly when the immunologically driven elimination of antigens for some reason does not succeed within a reasonable time period (Fig. 3). This may be due to a continuous supply of antigens (e.g. chronic virus infection,
autoimmunity, allergens) or excessive exposure to commensal and normally non-pathogenic bacteria through a defective mucosal barrier. An adverse immune reaction may, in addition, reflect an inefficient, skewed or exaggerated immune response on a hereditary (usually polygenic) background (e.g. unbalanced cytokine profile or overproduction of IgE antibodies in subjects with atopy). The hypersensitivity mechanisms involved cannot always be clearly defined, and a high degree of complexity is emerging as research is performed with steadily more sophisticated methods.

Cell-mediated immunity or DTH depends on a so-called Th1 profile of cytokines, including particularly interferon (IFN)-γ and tumour necrosis factor (TNF)-α. These cytokines activate macrophages and induce killer mechanisms, including cytotoxic T cells (Fig. 4). A Th2 profile includes mainly IL-4, IL-5 and IL-13, which are necessary for an IgE-mediated allergic reaction, involving degranulation of IgE-sensitised, strongly proinflammatory mast cells and activation of potentially tissue-damaging eosinophils (Fig. 4). Thus, IL-4 and IL-13 drive B-cell differentiation to IgE production, while IL-5 stimulate and prime eosinophilic granulocytes [207, 213–215].

**Regulation and the Th1/Th2 balance**

Great efforts have been made to elucidate the mechanisms involved in the induction and regulation of a polarised cytokine profile characterising activated Th-cell subsets [206]. There is particularly great interest in the role of APCs in shaping the phenotypes of naive T cells during their initial priming, partly because the differential expression level of various costimulatory molecules on activated and matured DCs may exert a decisive impact (Fig. 4). Thus, interaction of the T-cell CD28 receptor with B7.1 (CD80) appears to favour Th1 differentiation, and with B7.2 (CD86) the Th2 phenotype [216]. Certain cytokines secreted by the developed Th1 and Th2 cells act in an autocrine and reciprocally inhibitory fashion (Fig. 2); IL-4 promotes Th2 cell expansion and limits proliferation of Th1 cells, whereas IFN-γ enhances growth of Th1 cells but decreases Th2-cell development. In fact, the cytokine microenvironment clearly represents a potent determinant of Th1/Th2 polarisation, with IL-4 and IL-12 as the initiating key factors – being derived principally from innate immune responses during the T-cell priming (Fig. 4). Activated macrophages and DCs are the main source of IL-12, whereas an early burst of IL-4 may come from NKT cells, mast cells, basophils or already matured bystander Th2 cells [206].

Altogether, exogenous stimuli such as pathogen-derived products, the maturational stage of APCs and cytokines will influence Th1/Th2 differentiation on a background of genetic factors. In addition, there is an impact from complex interactions between antigen dose, T-cell receptor (TCR) engagement and MHC antigen affinities. High antigen levels appear to favour Th1 development, while low levels favour the Th2 subset [217]. Influenza antigenic properties include the nature of the antigen, with bacteria and viruses promoting Th1-cell differentiation and helminths the Th2 subset. Th2 differentiation also appears to be promoted by small soluble proteins characteristic of allergens (Fig. 4). Some important allergens (e.g. from house dust mite) are proteases, and it is suggested that this favours Th2 development because helminths secrete proteases to aid tissue penetration [207].
As mentioned previously (Fig. 2), Th1- and Th2-cell responses are cross regulatory, and the Th1/Th2 cytokine balance is also influenced by regulatory T (T<sub>reg</sub>) cells [218], which may secrete the suppressive cytokines IL-10 and transforming growth factor (TGF)-β. In summary, therefore, the nature of the APC (usually a DC) that stimulates the naïve T cells in a primary immune response will to a large extent influence the development of Th1, Th2 and T<sub>reg</sub> cells via its costimulatory molecules and cytokine secretion. In this manner the ‘signature’ of the microbial environment imprinted through PRRs, is important for maintenance of homeostasis in the adaptive immune system. Interestingly, the T<sub>reg</sub> cells may also exert a dampening effect directly on innate pathology-inducing mechanisms [219], and may themselves be directly influenced by microbial products such as LPs through TLRs which they express [220]. The anti-inflammatory regulatory network may furthermore include IL-10 and TGF-β derived from activated innate cells such as macrophages and DCs [221]. Lack of appropriate microbial stimulation controlling the homeostatic immune balance is central to the hygiene hypothesis and its implications for prevention of allergy by functional food [222–224].

**Mucosal immunity**

Mucosal immunity can be viewed as a first line of protection that reduces the need for systemic immunity, which is principally proinflammatory and therefore a ‘two-edged sword’ as explained above (Fig. 3). Numerous genes are involved in the regulation of innate and adaptive immunity, with a variety of modifications introduced over millions of years. During such evolutionary modulation, the mucosal immune system has generated two non-inflammatory layers of defence: (a) immune exclusion performed by secretory antibodies to inhibit surface colonisation of microorganisms and dampen penetration of potentially dangerous soluble substances; and (b) immunosuppressive mechanisms to avoid local and peripheral hypersensitivity to antigens that are normally innocuous (Fig. 5). The latter mechanism is referred to as ‘oral tolerance’ when induced via the gut [225], and probably explains why overt and persistent allergy to food proteins is relatively rare [226]. A similar downregulatory tone of the immune system normally develops against antigenic components of the commensal microbial flora [227–229].

Mucosally induced tolerance is a robust adaptive im-

![Diagram](image-url)

**Fig. 5** Schematic depiction of two major adaptive immune mechanisms operating at mucosal surfaces. 1 Productive immunity providing immune exclusion limits epithelial colonisation of pathogens and inhibits penetration of harmful foreign material. This first line of defence is principally mediated by secretory antibodies of the IgA (and IgM) class in cooperation with various nonspecific innate protective factors (not shown). Secretory immunity is preferentially stimulated by pathogens and other particulate antigens taken up through thin M cells (M) located in the dome epithelium covering inductive mucosa-associated lymphoid tissue (see Fig. 6). 2 Innocuous soluble antigens (e.g. food proteins; magnitude of uptake indicated) as well as the indigenous microbial flora are less stimulatory for secretory immunity (graded arrows), but induce suppression of pro-inflammatory humoral immune responses (IgG and Th2 cytokine-dependent IgE antibodies) as well as Th1 cytokine-dependent delayed-type hypersensitivity (DTH). The homeostatic Th2/Th1 balance is regulated by a complex and poorly defined phenomenon called ‘oral tolerance’ when induced via the gut. Its downregulatory effects can be observed both locally and in the periphery (modified from [212]).
mune function as more than a ton of food and drink may pass through the gut of an adult every year. This results in a substantial uptake of intact antigens, usually without causing any harm. However, the neonatal period is particularly critical, both with regard to infections and priming for allergic disease, because the mucosal barrier function and the immunoregulatory network are poorly developed for a variable period after birth [215, 230]. Notably, the postnatal development of mucosal immune homeostasis depends on the establishment of a normal bacterial flora as well as on adequate timing and dose of dietary antigens when first introduced [225, 229, 231].

Antibody-mediated mucosal defence

- **Immune exclusion.** The intestinal mucosa contains at least 80% of the body's activated B cells, which are terminally differentiated to Ig-producing blasts and plasma cells. Most of these immunocytes produce dimeric IgA, which along with pentameric IgM, can be actively transported through secretory epithelia by the polymeric Ig receptor (pIgR), also known as membrane secretory component or SC (Fig. 6). The binding site for this receptor depends on a small peptide called 'joining' (J) chain incorporated selectively into dimeric IgA and pentameric IgM [211]. Immune exclusion is then mediated by the generated secretory IgA (SIgA) and secretory IgM (SIgM) antibodies in cooperation with innate non-

---

**Fig. 6** Schematic depiction of the human mucosal immune system. Inductive sites for mucosal T and B cells are constituted by regional mucosa-associated lymphoid tissue (MALT) with their B-cell follicles and M cell (M)-containing follicle-associated epithelium through which exogenous luminal antigens are actively transported to reach professional antigen-presenting cells (APCs), including dendritic cells (DCs), macrophages, B cells and follicular dendritic cells (FDCs). In addition, intra- or subepithelial DCs may capture antigens and migrate via draining lymph to regional lymph nodes where they become active APCs, which stimulate T cells for productive or down-regulatory (suppressive) immune responses. Naive B and T cells enter MALT (and lymph nodes) via high endothelial venules (HEVs). After being primed to become memory/effector B and T cells, they migrate from MALT and regional lymph nodes via lymph and peripheral blood for subsequent extravasation at mucosal effector sites. This process is directed by the profile of adhesion molecules and chemokines expressed on the microvasculature, the endothelial cells thus exerting a local gatekeeper function for mucosal immunity. The mucosal lamina propria is illustrated with its various immune cells, including B lymphocytes, J chain-expressing IgA and IgM plasma cells, IgG plasma cells with a variable J-chain level (J), and CD4⁺ T cells. Additional features are the generation of secretory IgA (SIgA) and secretory IgM (SIgM) via pIgR (SC)-mediated epithelial transport, as well as paracellular leakage of smaller amounts (broken arrow) of both locally produced and serum-derived IgG antibodies into the lumen. Note that IgG cannot interact with J chain to form a binding site for pIgR. The distribution of intraepithelial lymphocytes (mainly T-cell receptor α/β⁺ CD8⁻ and some γδ⁺ T cells) is schematically depicted. Insert (lower left corner) shows details of an M cell and its ‘pocket’ containing various cell types.
specific defence mechanisms. In addition, some serum-derived or locally produced IgG antibodies may be transferred passively to the lumen by paracellular (Fig. 6). Importantly however, because IgG is complement-activating, its contribution to surface defence is potentially proinflammatory, which could jeopardise the epithelial barrier function [232].

Neonatal mucosal immunity. In contrast to several animal species, the human foetus acquires maternal IgG via the placenta and not as a result of breast-feeding. Also notably, intestinal uptake of SlgA antibodies from breast milk is of no importance for systemic immunity in humans, except perhaps in the preterm neonate [215]. Although so-called 'gut closure' normally occurs in humans mainly before birth, an effective mucosal barrier may not be established until after 2 years of age; the different variables involved in this process are poorly defined. Interestingly, the postnatal colonisation of commensal gut bacteria is important both to establish and regulate an appropriate intestinal epithelial barrier function [197, 198, 233].

Only occasional traces of SlgA and SlgM occur in human intestinal juice during the first postnatal period, whereas some IgG is often present, reflecting bulk leakage from the lamina propria, which after 34 weeks of gestation contains readily detectable maternal IgG [234]. IgA-producing cells are usually undetectable in the intestinal mucosa before the infant is 10 days of age, but thereafter a rapid increase takes place, IgM normally remaining predominant up to 1 month. Little increase of intestinal IgA production usually takes place after 1 year. A much faster establishment of secretory immunity may be seen in developing countries because of a much greater mucosal exposure to microorganisms [234].

Critical role of breast-feeding. A vast majority of pathogens use the mucosae as portals of entry. Thus, mucosal infections are a major killer below the age of 5 years – being responsible for more than 14 million deaths of children annually in developing countries. In those parts of the world, infants are highly dependent on SlgA antibodies from breast milk to protect their mucosae; epidemiological data suggest that the risk of dying from diarrhoea is reduced 14–24 times in breast-fed infants [235].

Although the value of breast-feeding in westernised countries is clinically most apparent in preterm infants, population studies show that exclusively breast-fed infants are in general better protected against a variety of infections and probably also against allergy, asthma and coeliac disease [215]. This strongly suggests that the mucosal barrier function in newborns can be reinforced by breast-feeding. Experiments in neonatal rabbits have demonstrated convincingly that SlgA is a crucial antimicrobial component of breast milk [236], in addition to a variety of other factors that may enhance mucosal homeostasis [215]. The protective role of secretory antibodies is further supported by the fact that knock-out mice lacking both SlgA and SlgM show increased mucosal leakiness as well as decreased resistance against bacterial toxins and colonising pathogens [237–239].

Immune induction in mucosa-associated lymphoid tissue (MALT)

Inductive lymphoid tissue. Mucosal lymphoid cells are located in three compartments: (a) organised MALT structures; (b) the mucosal lamina propria or glandular stroma; and (c) the mucosal surface epithelium. Gut-associated lymphoid tissue (GALT) constitutes a major part of MALT and comprises Peyer’s patches, the appendix and numerous solitary or isolated lymphoid follicles [210, 211]. All these structures are believed to represent inductive sites contributing to intestinal immune responses, while the lamina propria and epithelial compartment principally constitute effector sites (Fig. 6). The domes of GALT are covered by a characteristic follicle-associated epithelium, which contains M cells. These thin and bell-shaped specialised epithelial cells effectively transport microorganisms as well as dead antigens (especially of particulate nature) from the gut lumen into the organised lymphoid tissue [240]. Many enteropathogenic infectious agents use the M cells as portals of entry, so they represent extremely vulnerable parts of the epithelial barrier.

MALT structures resemble lymph nodes with B-cell follicles, intervening T-cell zones and a variety of APCs such as macrophages and DCs, but there are no afferent lymphatics (see Table 14). Exogenous stimuli therefore come directly from the lumen mainly via the M cells, perhaps aided by DCs which may penetrate the surface epithelium with their processes [241]. Induction and regulation of mucosal immunity hence takes place primarily in MALT and to a lesser extent at the effector sites (Fig. 6).

Stimulation and homing of intestinal B cells. Antigens are presented to naïve T cells in GALT by APCs after intracellular processing (degradation) to immunogenic peptides [211]. In addition, luminal peptides may be taken up and presented by APCs and epithelial cells directly to various subsets of intra- and subepithelial T lymphocytes (Fig. 6). Not only do mucosal APCs, but also the small intestinal villous epithelium and the follicle-associated epithelium of GALT surrounding the M cells, express MHC class II molecules in humans, particularly HLA-DR, and in addition they express classical and nonclassical MHC class I molecules [210, 242]. As discussed previously (Fig. 2), such molecules are essential for an antigen-presenting function. Interestingly, MHC class II-positive naïve and memory B lymphocytes
abound juxtaposed to the M cells. Such B cells may present antigens efficiently to T cells in cognate downregulatory or immunostimulatory interactions [243].

T cells primed in GALT release cytokines such as TGF-β and IL-10, which may drive the differentiation of mucosal antigen-specific B cells to predominantly IgA-committed plasma blasts, although their regulation still remains unclear [244, 245]. Most B cells primed by ‘first signals’ in GALT structures migrate rapidly via draining lymphatics to mesenteric lymph nodes where they are further stimulated; they may then reach peripheral blood and become seeded by preferential homing mechanisms into distant mucosal effector sites, particularly the intestinal lamina propria where they finally develop to Ig-producing plasma cells (Fig. 6). This terminal differentiation requires ‘second signals’ that are modulated by available antigens, various cell types expressing MHC class II molecules, and Th cells [211, 243]. Most B cells included in this homing or trafficking to mucosal effector sites apparently belong to clones of an early maturation stage, as indicated by their high level of J-chain expression regardless of co-concomitant isotype, although the IgA class is normally predominant. J chain-containing dimeric IgA and pentameric IgM are finally translocated to the lumen as SlgA and SlgM by the pIgR as mentioned above (Fig. 6).

The homing of primed T and B cells to the gut lamina propria is facilitated by mucosal ‘homing receptors’ interacting with ligands on the local microvascular endothelium (‘addressins’), with an additional fine-tuned navigation conducted by chemokines [211, 212]. Under normal conditions, therefore, the local microvasculature exerts a ‘gatekeeper’ function to allow preferential extravasation of primed lymphoid cells belonging to the intestinal immune system (Fig. 6). Other MALT structures such as the nasopharynx-associated lymphoid tissue (adenoids and tonsils) of Waldeyer’s ring, appear to have an immune-inductive function similar to that of GALT [246]. However, immune cells primed in Waldeyer’s ring differ from GALT-derived cells with regard to homing properties; this disparity contributes to a certain regionalisation of the mucosal immune system (Table 14). Altogether, effector cells generally tend to home back to the body region where they were initially stimulated.

The mammary glands are, in an integrated manner, part of the mucosal immune system, and milk antibodies reflect antigenic B-cell stimulation both in GALT and Waldeyer’s ring. Thus, breast milk contains SlgA targeted against potentially infectious agents in the mother’s environment. Breast-feeding, therefore, represents an ingenious immunological integration of mother and child [215].

Perinatal development of GALT. Very few B-cell blasts with IgA-producing capacity normally circulate in peripheral blood of newborns (<8 per million mononuclear cells), although this number is substantially increased already after 1 month (~600 per million mononuclear cells), reflecting the progressive stimulation of GALT with ensuing homing of immune cells to mucosal effector sites [247]. An initial early elevation of Ig-producing cells (mainly of the IgM class) can be seen in preterm infants, especially in those with intrauterine infections [248]. Thus, mucosal immune cells are competent at least during the final trimester, but APCs need to be activated by exogenous ‘danger signals’ such as microbial PAMPs, which enables them to provide appropriate co-stimulatory signals to prime naïve T cells (Fig. 2). The indigenous microbial flora is very important in this context as shown by the fact that the intestinal IgA system of germ-free mice is normalised after 4 weeks of conventionalisation. Bacteroides and Escherichia coli are considered to be most immunostimulatory of the commensal flora [229].

It is possible that suboptimal development of the SlgA-dependent mucosal barrier function and inadequate tolerance mechanisms together may explain the increasing frequency of certain diseases in industrialised countries, particularly allergies and autoimmune inflammatory disorders [224]. This ‘hygiene’ hypothesis has been tested in several studies by evaluating the beneficial effect of probiotic bacterial preparations. Especially viable strains of the commensal intestinal microflora, such as lactobacilli and bifidobacteria, have been reported to enhance IgA responses – both in humans and experimental animals – apparently in a T cell-dependent manner [229, 249–251]. A recent double-blind study of infants with a family history of atopy, reported the prevalence of atopic eczema to be reduced by 50% at the age of 2 years in those receiving a probiotic Lactobacillus strain daily for 6 months [252]. It remains to be determined whether this beneficial clinical

![Table 14 Normal ranges for immune cell numbers in the circulation of adult humans](image-url)

<table>
<thead>
<tr>
<th>Number per litre of blood (x 10^9)*</th>
<th>Total leucocytes</th>
<th>Neutrophils</th>
<th>Eosinophils</th>
<th>Basophils</th>
<th>Monocytes</th>
<th>Lymphocytes</th>
<th>T lymphocytes (CD3+)</th>
<th>Helper T lymphocytes (CD3+CD4+)</th>
<th>Cytotoxic T lymphocytes (CD3+CD8+)</th>
<th>B lymphocytes (CD19+)</th>
<th>Natural killer cells (CD56+)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.0–11.0</td>
<td>2.0–7.5</td>
<td>0.0–0.4</td>
<td>0.0–0.1</td>
<td>0.2–0.8</td>
<td>1.0–3.4</td>
<td>0.6–2.5</td>
<td>0.35–1.5</td>
<td>0.23–1.1</td>
<td>0.04–0.7</td>
<td>0.2–0.7</td>
<td></td>
</tr>
</tbody>
</table>

* 5th and 95th percentile [358]
effect was mediated via SIgA enhancement or oral tolerance promotion, or both.

Immune elimination and oral tolerance

- **Mucosal immune elimination.** This term refers to mechanisms involved in removal of foreign material that has penetrated the epithelial barrier. It represents a ‘second line’ of mucosal defence that depends partly on systemic immune components, such as serum-derived in addition to locally produced antibodies, probably often operating in combination with cytotoxicity mediated by CD8+ T cells and NK cells. Notably, immune elimination is enhanced by non-specific biological amplification systems of innate immunity (Fig. 3). As discussed above, this scenario may evolve into overt immunopathology if satisfactory removal of antigen is not rapidly achieved, thereby giving rise to inflammatory disease. Such ‘frustrated’ immunological elimination mechanisms are apparently part of the pathogenesis in various gut disorders including food allergy, coeliac disease and inflammatory bowel disease [210, 211, 214].

- **Oral tolerance.** It is believed that the adverse development alluded to above, is normally counteracted by a variety of mucosally induced immunosuppressive mechanisms collectively called oral tolerance (Fig. 5), including activation of T<sub>reg</sub> cells and balancing of the Th1/Th2 cytokine profile (Figs. 2 and 4). Identifiable experimental variables of this complex phenomenon include: genetics; age, dose and timing of postnatal antigen feeding; antigenic structure and composition; epithelial barrier integrity; and the degree of concurrent local immune activation as reflected by microenvironmental cytokines and expression of co-stimulatory molecules on APCs [225, 231]. Rodent studies suggest that the commensal microflora is important both for induction of oral tolerance and for reconstitution of this function alluded to above, is normally counteracted by a variety of mucosally induced immunosuppressive mechanisms collectively called oral tolerance (Fig. 5). As discussed above, this scenario may evolve into overt immunopathology if satisfactory removal of antigen is not rapidly achieved, thereby giving rise to inflammatory disease. Such ‘frustrated’ immunological elimination mechanisms are apparently part of the pathogenesis in various gut disorders including food allergy, coeliac disease and inflammatory bowel disease [210, 211, 214].

It is justified to believe that oral tolerance operates also in humans. Thus, in the normal state the vulnerable gut mucosa exhibits virtually no local IgG response and contains very few hyperactivated T cells [210, 211]. Moreover, the systemic IgG response to dietary antigens tends to decrease with increasing age [253, 254], and a hypo-responsive state to bovine serum albumin has been demonstrated by intradermal testing in adults [255]. The fact that resident APCs from normal human gut mucosa are quite inert in terms of their immuno-stimulatory properties [256], supports the notion that they play a central role in oral tolerance. One possibility is that mucosal DCs carry penetrating dietary and innocuous microbial antigens away from the mucosa, thereby avoiding local hyperactivation of immune cells [215].

Alltogether, a complex scenario apparently exits for oral tolerance (Fig. 5), depending on apoptosis of T cells (clonal deletion) when intestinal antigen exposure is excessive, but on anergy due to lack of costimulatory APC molecules, antigen clearance from the mucosa, or induction of T<sub>reg</sub> cells and immune deviation (skewing of T-cell cytokine profiles) at lower antigen doses [215, 231]. This scenario is further complicated by the fact that suppressive cytokines, for instance IL-10 and TGF-β, are produced not only by T<sub>reg</sub> cells (Fig. 2), but also by APCs and epithelial cells [221]. In addition, it remains unclear whether the most important immunoregulatory events for oral tolerance take place in Peyer’s patches, intestinal mucosa, mesenteric lymph nodes or the liver.

Importance of homeostatic mucosal immune regulation

According to the hygiene hypothesis, the increasing incidence of allergy in westernised societies may to some extent be explained by a reduced microbial load early in infancy, resulting in too little Th1-cell activity with insufficient IFN-γ to cross-regulate optimally IgE-inducing Th2-cell responses [257–259]. In this context, an appropriate composition of the commensal flora and exposure to foodborne and orofaecal microbes most likely exert an important homeostatic impact, both by enhancing the SIgA-mediated barrier function and by promoting oral tolerance through a shift from a predominant Th2-cell activity in the newborn period to a more balanced cytokine profile later on [215, 260]. Thus, the intestinal microflora of young children in Sweden was found to contain a relatively large number of *Clostridium* spp., whereas high levels of *Lactobacillus* spp. and *Eubacterium* spp. were detected in an age-matched population from Estonia [261]. Perhaps this difference could explain the lower incidence of allergy in the Baltic countries compared with Scandinavia [262]. A recent Finnish study likewise reported that IgE-sensitised infants had more Clostridia and tended to have fewer bifidobacteria in their stools than normal controls with a similar family history of atopy [263]. Such observations make a good case for studying the potential benefits of functional foods.

The feeding and treatment regimen (e.g. antibiotics) to which the newborn is subjected, as well as its nutritional state, exert a significant impact on the composition of its indigenous microbiota and on its gut integrity. Thus, the balance of the infant’s developing mucosal immune system may be easily disturbed [264]. As mentioned above, the immunoregulatory role of commensal bacteria in relation to atopic dermatitis was highlighted in a recent clinical trial with postnatal colonisation of a probiotic lactobacillus strain [252]. In experimental animal models, bifidobacteria have been shown to promote intestinal production of anti-rotavirus IgA antibodies [229]. Intestinal colonisation of both lactobacilli and bifidobacteria is enhanced by feeding on breast milk because of its large amounts of oligosaccharides.
of the immune system, but there is no increase susceptibility to infections. Whether this is so or not is discussed further below. Thus, by interacting with PRRs on various innate cell types such as macrophages and DCs (Fig. 2), the indigenous microbiota may have an impact on mucosal homeostasis beyond that of enhancing the SIgA system and promoting a Th1-cytokine profile that counterbalances Th2-cell responsiveness.

Measurements of immune status and function

Functional food claims related to the immune system are suggestive of an enhancement of activity with the assumption that this would result in improved host defence against infectious agents. In addition, a number of claims implicitly suggest immune enhancement or improved host defence against infectious agents or could be interpreted in these ways (Table 3).

Definition of the concept

The immune system acts to protect the host from pathogenic organisms (see above). In the face of infection the host is required to mount a coordinated and highly regulated response involving cellular activation, cellular movement, and the production of protective proteins, e.g. antibodies, acute phase proteins and mediators, e.g. cytokines. It is known that individuals with impaired immune responses are more susceptible to infections and are more likely to suffer from infectious morbidity and mortality. It is known that improvements in the immune responses of such individuals can decrease susceptibility to infections and decrease infectious morbidity and mortality. Such observations have led to the idea that enhancing the immune response will, ipso facto, decrease susceptibility to infections. Whether this is so or not is discussed further below.

There is a wide range of methodologies available with which to assess the status and functional capacity of the immune system, but there is no single marker of either its status or functional capacity. The activity of many of the separate components of the immune system can be measured, most frequently by studying that component under controlled ex vivo conditions. It is also possible to study a coordinated immune response in vivo. Animal studies can be used to investigate the functional responses of immune cells isolated from the blood, thymus, spleen, lymph nodes, and peritoneal cavity and, in some cases, from the bone marrow, lungs, liver and gastrointestinal tract. Human studies are often limited by the ability to sample only blood and external secretions such as saliva, although in some experimental settings it is possible to take biopsies of the gut, which may include immune tissue, or to collect bronchial lavage fluid. In most human settings, circulating cell numbers, their activation state and their responses to challenge can be measured. However, it must be remembered that the majority of immune cells are not in the bloodstream; for example, only 2% of total lymphocytes are circulating at any given time [271]. Normal ranges have been established for circulating immune cell numbers and Ig concentrations, but there are no normal ranges for other immune cell functional responses. Thus, it is vitally important that when immune cell functions are measured, appropriate controls and standardised, validated experimental protocols are used.

Assessments of immune status most frequently involve the measurement of various leukocyte numbers in the bloodstream, and the size and cellularity of lymphoid organs, where accessible.

Assessments of the functional capacity of the immune system can be made by:

- Measuring specific cell functions ex vivo, i.e. of cells isolated and studied in short- or long-term culture;
- Measuring in vivo responses to challenge, e.g. by measuring the changes in the concentrations of antibodies in the bloodstream or saliva, or by measuring the clinical response to administration of antigen;
- Measuring the incidence and severity of infections. In animal studies, resistance to challenge with live pathogens can be used; the outcome is usually survival, and this can be coupled with some of the above in vivo and ex vivo measures and with measures of the numbers of pathogens which are found in various organs, e.g. spleen, lymph nodes, liver. For ethical reasons human studies are largely restricted to naturally occurring infectious episodes or those using attenuated pathogens.

General comments on study design, subjects etc

Many immune parameters are affected by hormones, e.g. cortisol, adrenaline. Therefore it is important that measures of immune status and function and blood drawing be performed at the same time of day and in highly standardised conditions. Unless there is a need to study the response to some other factor, e.g. exercise, smoking, acute exposure to a drug etc, blood should be sampled in the morning after an overnight fast (> 10 hours) and from fully rested subjects. Since the blood collected will be required for measurement of immune cell numbers, immune cell functions, or plasma markers or for preparation of immune cells, it should be collected into an anti-coagulant. For most immune measures heparin is a suitable anti-coagulant. EDTA should be
avoided since it will interfere with responses involving cellular calcium fluxes, particularly if these are to be measured in whole blood. Immune cell numbers and functional responses may be affected by gender, age, obesity, dietary habits, other lifestyle factors, e.g. smoking, acute and chronic exercise, and the presence of disease, consumption of alcohol and pharmaceuticals, and pregnancy (Fig. 7). Thus, it is important (a) that subjects in any study be selected carefully and properly matched and (b) that measures of these confounding factors be accurately determined. This may simply involve questionnaires or may include additional biochemical analyses. Obviously subjects should be free of infections whilst being studied, unless susceptibility or responses to infection are under investigation.

In any dietary intervention study measuring immune functions, subjects should be studied before and after the intervention. Almost all measures of immune function allow for this, although an exception is studies where subjects’ response to vaccination is followed. A control group for any intervention should always be used; the control group should be well matched with the intervention group.

Since there are many factors affecting immune cell numbers and functional responses, observations made in one group of individuals may always not be readily extrapolated to other groups. Indeed, different groups of subjects may respond differently to the same dietary change. An example of this was seen in the study by Meydani et al. [272] in which dietary supplementation with fish oil significantly influenced immune cell functions in healthy older women (aged 51 to 68 years) but not in healthy younger women (aged 22 to 33 years). Furthermore, the effect of one dietary change may be influenced by another component of the diet. As an example, the effect of dietary fish oil on the proliferation of lymphocytes from healthy, non-smoking men aged 24 to 57 years was totally abolished if the subjects also consumed 200 mg α-tocopherol/day [273].

Taken together these studies suggest that observations made in one group of subjects can only be applied to groups with similar characteristics: the same effects may not occur in men and women, in individuals of different age, in smokers and non-smokers, in sedentary and physically active individuals, in individuals with different habitual diets, in different ethnic groups, and so on.

Individuals also express inherent genetic differences that may influence aspects of immune function. For example, polymorphisms in genes or in the promoter regions of genes coding for a vast array of proteins involved in the immune response have been identified, and many of these polymorphisms have functional consequences, e.g., polymorphisms in the promoter regions of cytokine genes influence the level of production of the cytokine (see [274, 275] for references). Likewise, polymorphisms in human leukocyte antigen genes affect the capacity for antigen presentation (see [275]). Such genotypic differences provide a genetic basis for differential sensitivity to infectious agents, for many diseases involving immune dysfunction, and for some cancers [275]. Recent studies also indicate that some polymorphisms may influence individual sensitivity to dietary components. For example, the effect of dietary fish oil supplementation on TNF-α production by immune cells from healthy males was strongly influenced by genotypes, i.e., combinations of polymorphisms, in the promoter regions of the TNF-α and TNF-β genes [276]. This study suggests that, even when other factors, e.g., gender, age, ethnicity, body mass index, physical activity, smoking status, are controlled and standardised, there may be some factors, e.g., genotype, that result in variable immune responses and in variable responses to dietary change. Ultimately, this may
mean that extrapolation of findings from individuals with one genotype to another may not be possible. However, at this stage it is too soon to be certain about this.

### Assessment of immune status

#### Cell counts (absolute and %)

The total number of white cells and of the subclasses of white cells, e.g. neutrophils, monocytes, lymphocytes, T lymphocytes, B lymphocytes, CD4+ cells, CD8+ cells, NK cells, in the circulation can be determined using immunological staining procedures, e.g. flow cytometry. Because “white cell counts” are used clinically, normal ranges have been identified (Table 14). In addition to total cell numbers, the percentage contribution of each class to the total is sometimes used. By combining antibody “stains” it is possible to obtain great detail about the sub-types of cells present. The ratio of CD4+ to CD8+ cells is often reported as a measure of the relative numbers of T-helper and T-suppressor/cytotoxic cells, but the true meaning of this ratio is unclear. The ratio of memory to naïve cells (CD45RO:CD45RA) can be determined; this is an indicator of long-term activation of the immune system. The number of memory T cells increases over an individual’s lifetime representing cumulative antigen exposure.

#### Size of lymphoid organs

In animal studies the thymus, spleen and lymph nodes can be removed and weighed. In human studies thymus size can be estimated by imaging techniques. This approach has been used to identify differences in thymus size among infants on different feeding regimens [277] and to show an increase thymus size in malnourished children given oral zinc [278].

#### Assessment of the functional activity and capacity of the immune response

**In vivo measures (see Table 15)**

**Circulating concentrations of total Ig and of the Ig subclasses.** Measurements are made by ELISA or similar
method. Typical ranges for these are: IgA (1.4–4 mg/ml); IgD (0–4 mg/ml); IgE (17–450 ng/ml); IgG (8–16 mg/ml); IgM (0.5–2 mg/ml). In the absence of an “immune challenge” measurements of Igs are not very useful.

Circulating concentrations of Ig specific for antigens can be measured after an antigen challenge of some sort, e.g. inoculation with a vaccine such as those to hepatitis B, influenza or Pneumococcus. Because blood can be sampled serially these measurements can provide a dynamic picture of both primary and secondary antibody responses. These measurements are very useful since they represent a coordinated, integrated immune response to a relevant challenge.

Concentration of secretory IgA (sIgA) in saliva and tears. Total and antigen-specific sIgA can be measured. This can be a useful measure of mucosal immune responses.

Circulating concentrations of cytokines or of soluble cytokine receptors. Measurements are made with ELISA, often requiring highly sensitive assays. The source of the cytokines measured is not known. For example adipose tissue and skeletal muscle are important sources of circulating TNF-α and IL-6, respectively, in some situations. However, elevated concentrations of some cytokines have been associated with chronic disease (e.g. atherosclerosis, type-2 diabetes, obesity, arthritis, Alzheimers etc.) and are predictive of poor outcome in critically ill patients [279].

Delayed-type hypersensitivity response to intradermal application of an antigen to which the individual has already been exposed. This measures the cell-mediated immune response, and is often referred to as a “skin test”. The response is measured as the size of the reaction, termed induration, around the area of application at a period, usually 48 hours, after the application. This measurement is useful since it represents a coordinated, integrated cell-mediated immune response to a relevant challenge. However, there is significant variation in the response among individuals (see below), the test cannot be repeated on the same area skin, and recent vaccination may interfere with the outcome. Furthermore, most studies that have made this measurement have used commercially available applicator kits which are no longer available.

Incidence and severity of infectious diseases. This has been widely used in human studies to suggest interactions between nutrient status or physiological perturbation (e.g. heavy exercise) and immune function.

Ex vivo measures (Table 16)

Ex vivo measures allow the functional responses of specific immune cell types, e.g. neutrophils, monocytes, T lymphocytes, B lymphocytes, NK cells etc, to be determined. To obtain a detailed overall view of the effect of a dietary change, a battery of immune cell functions should be measured. These should represent the functions of several types of immune cells.

Phagocytosis by neutrophils and monocytes. Substrates for phagocytosis include bacteria, sheep red blood cells and yeast particles; these can be studied in the opsonised and unopsonised states. Some techniques, e.g. flow cytometry, allow identification of both the number of cells participating in phagocytosis and the phagocytic activity per cell. Measures of phagocytosis can be coupled to measures of oxidative burst; bacterial phagocytosis measurements can be coupled to measures of bacterial killing.

Oxidative (respiratory) burst (superoxide generation) by neutrophils and monocytes. Stimuli to induce respiratory burst include bacteria and protein kinase C activators such as phorbol esters, e.g. phorbol myristyl acetate. Experimental conditions should allow for both increased and decreased oxidative burst to be measured. Some techniques, e.g. flow cytometry, allow identification of both the number of cells participating in oxidative burst and the activity per cell. Oxidative burst measurements can be coupled with measures of bacterial killing. Production of other reactive species such as hydrogen peroxide can also be made.

Chemotactic response of neutrophils or monocytes. This is the movement of these cells towards particular stimuli; stimuli used include leukotriene B4, bacterial cell wall peptides such as formyl-methionyl-leucyl-phenylalanine, IL-8 and autologous serum.

Eicosanoid production by neutrophils and monocytes. Isolated cells can be stimulated with appropriate agents such as calcium ionophores, phorbol esters, or bacterial LPS and the concentrations of eicosanoids in supernatants can be measured by ELISA, radioimmunoassay, gas chromatography/mass spectrometry, or HPLC. The timing of eicosanoid generation depends upon cell type and the stimulus used.

Natural killer cell activity. This is measured as killing of tumour cells known to be specific targets for natural killer cells. The K562 cell line is often used as a target for human natural killer cells. The assay is normally conducted at several ratios of killer to target cell, e.g. 100:1, 50:1, 25:1, 12.5:1. Typically the assay time is quite short, e.g. 4h. There are a number of ways to measure target cell killing. Classically, target cells are pre-loaded with $^{51}$Cr and the release of $^{51}$Cr into the medium as a result of target cell death is determined by standard γ counting. One advantage of this assay is that background
## Table 16  Summary of *ex vivo* methods used to assess the functional capacity of components of the immune system

<table>
<thead>
<tr>
<th>Component of the immune system examined</th>
<th>Method</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Reliability (C. V.)</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutrophil or monocyte phagocytic activity</td>
<td>Measurement of uptake of particles (e. g. bacteria, yeast cell walls) often by flow cytometry (requires particles to be fluorescently labelled)</td>
<td>Can be performed in whole blood and so does not require purification of cells; Can determine the percentage of cells involved and the activity per cell; Can be coupled to measurement of oxidative burst</td>
<td></td>
<td>5 to 10 %</td>
<td>Moderate</td>
</tr>
<tr>
<td>Neutrophil or monocyte oxidative burst activity</td>
<td>Measurement of generation of superoxide or other reactive oxygen species by spectrophotometry or flow cytometry; Stimulated by bacteria (i.e. phagocytosis) or by protein kinase C activators such as phorbol esters</td>
<td>If using spectrophotometry: Can follow reaction with time to obtain rate of generation and total generated if using flow cytometry: Can be performed in whole blood and so does not require purification of cells; Can determine the percentage of cells involved and the activity per cell; Can be coupled to measurement of phagocytosis</td>
<td></td>
<td>5 to 10 %</td>
<td>Good</td>
</tr>
<tr>
<td>Neutrophil or monocyte chemotactic response</td>
<td>Measurement of movement towards chemoattractants such as peptides (e. g. fMLP, leukotriene B4, autologous serum)</td>
<td>Can measure number of cells migrating and distance migrated</td>
<td>Not known</td>
<td></td>
<td>Moderate</td>
</tr>
<tr>
<td>Cytokine production by monocytes</td>
<td>Measurement of production of cytokine protein by ELISA (also by flow cytometry) or cytokine mRNA following stimulation, often with LPS</td>
<td>Can be performed in whole blood culture so does not require purification of cells; Can measure intracellular cytokines by flow cytometry and therefore identify relative number and type of cytokine producing cells; Can measure both pro- and anti-inflammatory cytokines in the same sample</td>
<td>Different cytokines produced at different rates and with different sensitivities to the stimulant used</td>
<td>5 to 10 %</td>
<td>Moderate</td>
</tr>
<tr>
<td>Eicosanoid production by neutrophils or monocytes</td>
<td>Measurement of production of eicosanoid by ELISA, RIA, HPLC following stimulation, with LPS, calcium ionophore etc.</td>
<td>Can be performed in whole blood and so does not require purification of cells; Can be performed in whole blood and so does not require purification of cells; Can determine the percentage of cells expressing the marker and the level of expression per cell</td>
<td>Not known</td>
<td></td>
<td>Moderate</td>
</tr>
<tr>
<td>Surface expression of proteins involved in antigen presentation by monocytes</td>
<td>Measurement of cell surface expression of marker (e. g. HLA subtype) by flow cytometry following stimulation</td>
<td>Can be performed in whole blood and so does not require purification of cells; Can determine the percentage of cells expressing the marker and the level of expression per cell</td>
<td>Not known</td>
<td></td>
<td>Good</td>
</tr>
<tr>
<td>Natural killer cell activity</td>
<td>Measurement of killing of defined target cells (e. g. K562) which are often pre-labelled with 51Cr or fluorescently labelled (flow cytometry)</td>
<td>Use of 51Cr</td>
<td>5 to 10 %</td>
<td>Good</td>
<td></td>
</tr>
<tr>
<td>Cytotoxic T cell activity</td>
<td>Measurement of killing of defined target cells (e. g. P815) which are often pre-labelled with 51Cr or fluorescently labelled (flow cytometry)</td>
<td>Use of 51Cr</td>
<td>5 to 10 %</td>
<td>Good</td>
<td></td>
</tr>
</tbody>
</table>
counts can be low, giving a high level of sensitivity. However, the use of $^{51}$Cr requires suitable precautions. There are alternative methodologies for determining natural killer cell activity. It is possible to fluorescently label target cells and to determine target cell killing using flow cytometry. Alternatively, target cell death has been determined as the appearance of lactate dehydrogenase in the medium; this is released from dead target cells. If this approach is used a number of controls are required, because there may be spontaneous release of lactate dehydrogenase from both killer cells and target cells. Also this assay must be done in serum free medium, because serum contains lactate dehydrogenase. Whatever approach is used, the data can be expressed in various ways, such as % target cell killing at each killer to target cell ratio or "lytic ratio" which is the ratio required to kill a particular percentage, e.g. 25 or 50% of target cells.

- **Cytotoxic T lymphocyte activity.** This is measured as killing of virally infected cells known to be specific targets for cytotoxic T cells. The P815 cell line is often used as a target for human cytotoxic T cells. The assays are performed in the same way as described for natural killer cell activity.

- **Lymphocyte proliferation.** This is the increase in number of lymphocytes in response to a stimulus. Most often this is measured as the incorporation of radioactively labelled thymidine into the DNA of the dividing lymphocytes, although a number of other measures, not involving the use of radioactivity, are available. If thymidine incorporation is used then the thymidine is normally added towards the end of an extended cell culture period, e.g. for the final 18 hours of a 72 hour culture period. Agents used to stimulate lymphocyte proliferation include concanavalin A (Con-A), phyto-

### Table 16  
Continued

<table>
<thead>
<tr>
<th>Component of the immune system examined</th>
<th>Method</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Reliability (C. V.)</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lymphocyte proliferation</td>
<td>Measurement of an indicator of an increase in cell number following stimulation, often with PHA, Con A, anti-CD3, phorbol ester + ionophore; Most often involves measurement of the incorporation of $^{3}$H(thymidine into DNA although colourimetric and ELISA assays are available</td>
<td>Use of $^{3}$H</td>
<td>Sensitive but requires suitable precautions. One must consider the use of $^{51}$Cr.</td>
<td>5 to 10 %</td>
<td>Moderate</td>
</tr>
<tr>
<td>Cytokine production by T lymphocytes</td>
<td>Measurement of production of cytokine protein by ELISA (also by flow cytometry) or cytokine mRNA following stimulation, often with Con A or PHA or anti-CD3</td>
<td>Can be performed in whole blood culture so does not require purification of cells; Can measure intracellular cytokines by flow cytometry and therefore identify relative number and type of cytokine producing cells; Can measure a range of cytokines (e.g. Th1-type and Th2-type) in the same sample</td>
<td>Different cytokines produced at different rates and with different sensitivities to the stimulant used</td>
<td>5 to 10 %</td>
<td>Moderate</td>
</tr>
<tr>
<td>Antibody production</td>
<td>Measurement of total or antigen-specific immunoglobulins by ELISA following stimulation with antigens</td>
<td>Reflects B-cell activity</td>
<td>Use of radioimmunoassay. Serum free medium required since immunoglobulins are present in serum.</td>
<td>10 to 20 %</td>
<td>Moderate</td>
</tr>
<tr>
<td>Surface expression of activation markers on lymphocytes</td>
<td>Measurement of cell surface expression of marker (e.g. IL-2 receptor CD25, CD69) by flow cytometry following stimulation</td>
<td>Can be performed in whole blood and so does not require purification of cells; Can determine the percentage of cells expressing the marker and the level of expression per cell; By combining antibodies can get great detail of the responses of different cell types</td>
<td>Not known. Requires serum free medium. Serum contains chemotactic factors that can affect lymphocyte responses.</td>
<td>Not known</td>
<td>Good</td>
</tr>
</tbody>
</table>

C.V. coefficient of variation; SIgA secretory IgA; DTH delayed type hypersensitivity; fMLP f-methionine leucine phenylalanine; ELISA enzyme linked immunosorbant assay; LPS lipopolysaccharide; RIA radioimmunoassay; HPLC high performance liquid chromatography; PHA phytohaemagglutinin; Con-A concanavalin
haemagglutinin (PHA) and anti-CD3, which stimulate T lymphocytes; pokeweed mitogen that stimulates a mixture of T and B lymphocytes; and bacterial lipopolysaccharide, which stimulates B lymphocytes. These agents are all known as mitogens and the process as mitogen-stimulated lymphocyte proliferation. Most often, T cell mitogens are used. If the individual has been sensitised to an antigen (or allergen) then the antigen (or allergen) can be used to stimulate lymphocyte proliferation. If the process of lymphocyte proliferation is determined using thymidine incorporation then the results can be expressed as “cpm” or “dpm” of radioactivity incorporated per culture or this can be normalised to the number of lymphocytes initially cultured. If cells are cultured in both unstimulated and stimulated states then the results can be expressed as stimulation index, i.e. incorporation in the presence of stimulus divided by incorporation in the absence of stimulus. The proliferative response to mitogens or antibodies is much greater than that to an antigen or allergen. This is because mitogenic stimulation is non-specific and will target a large proportion, perhaps all, of the T or B cells in a cell preparation. In contrast, antigenic stimulation is highly specific and targets those few cells that will recognise the antigen.

**Production of cytokines by lymphocytes and monocytes.** This usually requires the cells to be stimulated. For lymphocytes, mitogens are used or antigens, if the individual has been sensitised, while for monocytes bacterial lipopolysaccharide is most often used. Cytokine protein concentrations in the cell culture medium are most frequently measured by ELISA. However, cellular mRNA levels can also be measured, e.g. by PCR technologies. Flow cytometry can be used to measure the intracellular concentration of cytokine protein. This technique also allows the relative number of cytokine producing cells to be identified and, if combined with other immunological stains, the type of cells producing the cytokine. A similar method is ELISpot; which allows the absolute number and type of cytokine producing cells to be identified. Whichever approach is used, cytokine production is a dynamic process and the concentration of cytokine mRNA or protein represents a balance between synthesis and degradation or utilisation. Thus, several time points should be studied, each one of these providing a “snap-shot” of the situation at that specific moment. The production of Th1- and Th2-type cytokines by isolated lymphocytes can be used to indicate the balance between the two types of response. IFN-γ is frequently used as a marker for the Th1-type response. IL-4 has sometimes been used as a marker for the Th2-type response, but IL-4 is often produced in low amounts and only after prolonged periods in culture. IL-5 is an alternative to IL-4.

**Production of Ig by lymphocytes.** This involves measurement of total or antigen-specific immunoglobulins by ELISA following stimulation with antigens and reflects B cell activity.

**Cell surface expression of molecules involved in antigen presentation, e.g. HLA subtypes, and in cellular activation, e.g. cytokine receptors, CD69, after stimulation.** Stimulants used can include mitogens or antigens. Cell surface expression is most frequently determined by flow cytometry after immunological staining. The percentage of cells expressing the molecule and the average level of expression per cell can both be determined. If combined with other immunological stains, the type of cell expressing the molecule can be identified. CD69 is expressed relatively early by lymphocytes stimulated with mitogens, e.g. within 6 hours, while cytokine receptors such as the CD25, the α-chain of the IL-2 receptor, appear later, e.g. after 12 to 24 hours. Thus, surface molecule expression is a dynamic process and represents a balance between appearance on the surface and internalisation. Therefore, several time points should be studied, each one of these providing a “snap-shot” of the situation at that specific moment.

**Specific measures of gut integrity and gut-associated immune responses (Table 17)**

There are a number of techniques that can be used to determine mucosal structural and functional integrity. Although these do not measure directly aspects of immune function, they give an indication of the intactness of mucosal barrier function and the likelihood of increased translocation of antigens and microorganisms across the gut and so of the potential for exposure of these to the immune system. A number of proteins can be measured in faeces and some of these are indicative of responses of the intestinal immune system. For example, faecal TNF-α and calprotectin (a granulocyte-derived protein) can indicate intestinal inflammation. However, direct effects of dietary interventions on the gut-associated immune system cannot be easily measured in humans.

**Specific comments on in vivo and ex vivo measures of functional activity and capacity of the immune response**

By definition, ex vivo measures require that cell functions be studied outside of the normal environment in which they normally occur, i.e. within the body. Ex vivo cell responses may not be the same as those observed in the more complex in vivo situation. This effect may be exaggerated by studying cells in increasingly purified states. Thus, measurements of cell function made in whole blood may be more similar to those seen in vivo.
than functions measured using purified cell preparations. Whole blood systems retain all blood components (including plasma) and they are kept at the same ratios at which they exist in vivo; by definition cell purification removes many blood components. Since measures of immune cell functions require a period of culture, which can be from minutes to several days, this raises a number of technical issues with regard to the appropriate additions to make to the cell culture medium. A major issue is that of serum/plasma source and concentration. Cultured cells typically require a source of serum/plasma, although there are serum-free supplements available for use. There are several options for the choice of serum/plasma: foetal bovine serum, autologous serum or plasma, i.e. from the same donor as the cells, or pooled human AB serum or plasma. The nature of the serum/plasma used can affect the absolute functional response observed, as can the concentration of serum/plasma used. One advantage of using purified cells for measuring some ex vivo functional responses, e.g. lymphocyte proliferation, cytokine production, antibody production, is that the number of cells cultured can be carefully controlled; this may not be the case where whole blood is cultured.

When making measures of immune function, either in vivo or ex vivo, it must be remembered that the responses being measured are dynamic in nature. Thus, the absolute response measured may be different at different time points; for example, the concentration of a given cytokine in the cell culture medium may be higher at 48 hours of culture than at 24 hours. Furthermore, different responses follow different time courses; for example the concentration of one cytokine may be highest after 24 hours of cell culture while the concentration of a second cytokine may be highest at 72 hours of cell culture. Thus, if there is a desire to more fully understand the effect of an intervention it is appropriate to study the functional responses at several time points. Another issue is that immune responses are related to the concentration of the stimulant used to trigger those responses in a dose-dependent fashion. Thus, once again the absolute response and the timing of that response will depend upon the concentration of the stimulus used (see Fig. 8 for an example), and it may be desirable to use several concentrations of stimulus in order to more fully understand the effect of an intervention.

The sections above highlight a number of factors that may influence any given immune functional outcome: whether whole blood or purified cells are used; the choice of type and concentration of plasma/serum; the timing of the response being studied; and the relationship of the response to the concentration of the stimulus used. Furthermore, the number of responder cells will influence the absolute response, the timing of that response and the sensitivity to stimulus concentration. Thus, it is absolutely imperative that for a given study or set of studies a highly standardised protocol be used. The effect of this is that results for the same assay between laboratories, or even within a laboratory, if some aspect of the experimental protocol is changed, may not be directly comparable.

Even when highly standardised experimental conditions are used there are wide variations in all in vivo and ex vivo measurements of immune responses. Some of this variation is likely due to factors mentioned earlier, e.g. age, smoking status, obesity, dietary habits, acute and chronic exercise, acute and chronic consumption of alcohol, pregnancy etc. Nevertheless, even when as many of these factors are standardised significant variation remains (Tables 18, 19). Genetic polymorphisms, early life events, hormone status and gut flora may be additional factors contributing to such variation [280].

Because ex vivo cell culture is susceptible to variation in many factors, in vivo measures of immune competence are ultimately of superior value in predicting host resistance to infections. Because these are conducted in the whole body setting they are the result of a coordinated, intact immune response and they are less susceptible...
Fig. 8  Increase in expression of CD69 on the surface of human lymphocytes following stimulation with the T cell mitogen Con A. Blood was collected from a healthy male after a 10 hour fast. Blood was diluted 1:1 with culture medium and was stimulated with different concentrations of Con A for various periods of time; final culture volume was 0.25 ml. At the end of each culture period, cells were stained with a monoclonal antibody to CD69 and analysed by flow cytometry. Median fluorescence intensity of CD69-positive cells is shown. Date are previously unpublished (G. C. Burdge and P. C. Calder)

Table 18  Phagocytic activity in response to E. coli and oxidative burst activity in response to E. coli or phorbol ester

<table>
<thead>
<tr>
<th>Cell Type</th>
<th>Function</th>
<th>Stimulus</th>
<th>Median</th>
<th>10th Percentile</th>
<th>90th Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutrophil</td>
<td>Phagocytosis</td>
<td>E. coli</td>
<td>Males</td>
<td>74.0</td>
<td>92.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>81.9</td>
<td>93.2</td>
</tr>
<tr>
<td></td>
<td>MFI</td>
<td>E. coli</td>
<td>Males</td>
<td>1548</td>
<td>4848</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>2308</td>
<td>4247</td>
</tr>
<tr>
<td>Monocyte</td>
<td>Phagocytosis</td>
<td>E. coli</td>
<td>Males</td>
<td>24.7</td>
<td>40.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>23.0</td>
<td>38.8</td>
</tr>
<tr>
<td></td>
<td>MFI</td>
<td>E. coli</td>
<td>Males</td>
<td>1138</td>
<td>3093</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>1533</td>
<td>3460</td>
</tr>
<tr>
<td>Neutrophil</td>
<td>Oxidative Burst</td>
<td>E. coli</td>
<td>Males</td>
<td>94.5</td>
<td>98.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>94.2</td>
<td>98.7</td>
</tr>
<tr>
<td></td>
<td>MFI</td>
<td>E. coli</td>
<td>Males</td>
<td>620</td>
<td>1321</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>585</td>
<td>1365</td>
</tr>
<tr>
<td>Neutrophil</td>
<td>Oxidative Burst</td>
<td>Phorbol ester</td>
<td>Males</td>
<td>94.8</td>
<td>98.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>95.5</td>
<td>99.6</td>
</tr>
<tr>
<td></td>
<td>MFI</td>
<td>Phorbol ester</td>
<td>Males</td>
<td>1445</td>
<td>2589</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>1596</td>
<td>2864</td>
</tr>
<tr>
<td>Monocyte</td>
<td>Oxidative Burst</td>
<td>E. coli</td>
<td>Males</td>
<td>54.0</td>
<td>84.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>60.9</td>
<td>80.6</td>
</tr>
<tr>
<td></td>
<td>MFI</td>
<td>E. coli</td>
<td>Males</td>
<td>173</td>
<td>873</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>121</td>
<td>846</td>
</tr>
<tr>
<td>Monocyte</td>
<td>Oxidative Burst</td>
<td>Phorbol ester</td>
<td>Males</td>
<td>64.7</td>
<td>97.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>76.5</td>
<td>96.4</td>
</tr>
<tr>
<td></td>
<td>MFI</td>
<td>Phorbol ester</td>
<td>Males</td>
<td>196</td>
<td>621</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>173</td>
<td>602</td>
</tr>
</tbody>
</table>

Blood was collected into heparin from males (n = 55) and females (n = 40) after an overnight fast. Subjects were aged 25 to 72 years and were healthy; they were not taking any prescribed medication; they did not have diagnosed cardiovascular disease, diabetes, liver or endocrine dysfunction or chronic inflammatory disease; they were not pregnant or lactating; they were not vegetarian; they did not consume fish oil, evening primrose oil or vitamin supplements; they smoked < than 15 cigarettes/day; they exercised strenuously < 3 times/wk; they had a body mass index between 18 and 34 kg/m²; they did not consume > 2 portions of oily fish/week. Habitual nutrient intakes of the subjects were in accordance with UK averages. Neutrophil and monocyte phagocytosis of E. coli and respiratory burst in response to E. coli or phorbol ester were determined in whole blood by flow cytometry (see [359] for details). Data shown are % of active cells and median fluorescence intensity (MFI a measure of the activity per cell). Data are reproduced with permission from the American Society for Clinical Nutrition from Kew et al. [359]
tible to the various confounding effects associated with cell culture. Nevertheless, in vivo approaches are not straightforward and they are still highly variable between individuals. Eighty-six of the individuals described in Tables 19 and 20 underwent the DTH “skin test” which involved standardised application of seven antigens to the forearm. Median response was to two antigens (10th and 90th percentile values 0 and 3, respectively) with a median cumulative area of response of 15.6 mm² (10th and 90th percentile values 0 and 50.8 mm², respectively) [281].

The large variation among individuals in all potential immune outcomes means that intervention studies must be adequately powered to identify significant effects.

Specific comments relating to the biological significance of any effects demonstrated on immune function

Decreases or increases in indicators of immune function (up to 10% at least) may not be relevant to host defence. There are two main reasons for this. First, there is significant redundancy in the immune system, such that a small change in the functional capacity of one component of the immune response may be compensated for by a change in the functional capacity of another component. Secondly, there may be “excess” capacity in some immune functional responses, particularly those that are measured ex vivo by challenging the cells with a high concentration of stimulant. To get a detailed overall view of the effect of a dietary change, a battery of immune cell functions should be measured. It is not clear whether the wide variation in in vivo and ex vivo functional responses among apparently healthy individuals results in variable susceptibility to infection. If it does not, then the notions of redundancy and excess capacity outlined above would explain this. Thus, it is not absolutely certain that an increase in one or more immune function parameters among healthy individuals will improve host resistance. Thus, healthy individuals may not benefit from altered immune function. In other words increasing the activity of one or more components of the immune system, as tested by the above methodologies, may not necessarily be beneficial to the individual, just as decreasing the activity of one or more components of the immune system, as tested by the above methodologies, may not necessarily be detrimental to the individual. However, large variations or changes in some immune functions have been related to improved host defence. For example, data from recent studies indicate that individuals with low natural killer cell activity have increased risk of infections and cancer compared with individuals with moderate or high activity.

<table>
<thead>
<tr>
<th>Cell Type</th>
<th>Function</th>
<th>Units</th>
<th>Median</th>
<th>10th Percentile</th>
<th>90th Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBMC</td>
<td>IL-2 production</td>
<td>kU/L</td>
<td>Males</td>
<td>6.3</td>
<td>3.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>6.2</td>
<td>2.7</td>
</tr>
<tr>
<td>PBMC</td>
<td>IFN-γ production</td>
<td>kU/L</td>
<td>Males</td>
<td>63.7</td>
<td>15.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>94.0</td>
<td>8.6</td>
</tr>
<tr>
<td>PBMC</td>
<td>IL-4 production</td>
<td>ng/L</td>
<td>Males</td>
<td>33.5</td>
<td>9.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>26.8</td>
<td>5.7</td>
</tr>
<tr>
<td>PBMC</td>
<td>IL-6 production</td>
<td>ng/L</td>
<td>Males</td>
<td>37288</td>
<td>6517</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>30089</td>
<td>7225</td>
</tr>
<tr>
<td>PBMC</td>
<td>IL-1β production</td>
<td>ng/L</td>
<td>Males</td>
<td>4205</td>
<td>1277</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>4729</td>
<td>1608</td>
</tr>
<tr>
<td>PBMC</td>
<td>TNF-α production</td>
<td>ng/L</td>
<td>Males</td>
<td>10352</td>
<td>2660</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>9371</td>
<td>3277</td>
</tr>
<tr>
<td>Lymphocyte</td>
<td>Proliferation</td>
<td>cpm/well</td>
<td>Males</td>
<td>29224</td>
<td>15251</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Females</td>
<td>31347</td>
<td>5875</td>
</tr>
</tbody>
</table>

Blood was collected into heparin from males (n = 88) and females (n = 62) after an overnight fast. Subjects were aged 25 to 72 years and were healthy; they were not taking any prescribed medication; they did not have diagnosed cardiovascular disease, diabetes, liver or endocrine dysfunction or chronic inflammatory disease; they were not pregnant or lactating; they did not consume fish oil, evening primrose oil or vitamin supplements; they smoked < than 15 cigarettes/day; they exercised strenuously < 3 times/wk; they had a body mass index between 18 and 34 kg/m²; they did not consume > 2 portions of oily fish/week. Habitual nutrient intakes of the subjects were in accordance with UK averages. Mononuclear cells were purified from the blood and cultured under standard conditions (see [359]). Lipopolysaccharide was used to stimulate production of TNF-α, IL-1β and IL-6. Concanavalin A was used to stimulate production of IL-2, IFN-γ and IL-4 and lymphocyte proliferation; the latter was determined as thymidine incorporation into DNA. Cytokines were measured by ELISA. Data are reproduced with permission from the American Society for Clinical Nutrition from [359].

Table 19 Production of cytokines by peripheral blood mononuclear cells (PBMC) and proliferation of lymphocytes
in vivo nutrients used can greatly exceed concentrations present in cells in isolation from other components with highly unphysiological in nature. For example, they use in vivo ing examined may differ from the form in which it exists. For these reasons, extrapolations from in vitro studies to the whole body context should only be made cautiously, and effects identified in in vitro studies must be confirmed in controlled human dietary studies.

Comment on the role of solely in vitro studies

In vitro studies refer to studies in which isolated cells are exposed directly, in culture, to agents, e.g. nutrients, under examination for immunomodulatory properties. The value of this approach is that the experimental conditions are highly controlled, that detailed dose response studies can be performed, that high-throughput screening is possible, and that mechanisms of action can be identified. However, in vitro systems frequently are highly unphysiological in nature. For example, they use cells in isolation from other components with which they would normally interact, the concentrations of nutrients used can greatly exceed concentrations present in vivo, and the form of the nutrient, or other factor, being examined may differ from the form in which it exists in vivo. For these reasons, in vitro studies must be confirmed in controlled human dietary studies.

Infections, atopy and inflammatory disorders

One of the most successful uses for functional foods affecting the gastrointestinal tract has been in the reduction of risk of diarrhoeal diseases. This derives from the concept that infection with pathogenic bacteria or viruses can be resisted more effectively if the natural barrier provided by the indigenous or commensal flora is strengthened. The principal conditions where functional foods, mainly probiotics, have been used are in the prevention of antibiotic-associated diarrhoea (AAD), traveller’s diarrhoea and acute diarrhoea of children. In these conditions functional claims for strengthening the barrier function or natural defences of the gut seem legitimate, whilst reduction of disease risk has also been clearly demonstrated. Probiotics have also been used in the management of inflammatory bowel diseases although this clearly becomes an area for medical rather than functional claims. Similarly, their use in the treatment of acute infectious diarrhoea of children is medicinal but was the starting point for probiotic use early in the last century.

The mucosae of the gastrointestinal tract form an important organ of the host. In addition to its principal physiological function, digestion and absorption of nutrients, the intestinal mucosa provides a protective interface between the internal environment and the constant challenge from antigens of the external environment, also carrying defence mechanisms against infectious and inflammatory diseases. Protection against potentially harmful agents encountered by the enteric route is provided by a number of non-immunological factors and immunological mechanisms.

While it appears evident that balanced normal microbiota may become aberrant secondary to gut-related disease such as infectious diarrhoea, and thus constitute a target for probiotic intervention, it is not known whether aberrancies in the early composition of the microbiota per se can be a primary cause of disease. Nevertheless, the predisposition to infectious and inflammatory disease is associated with immaturity and dysfunction of the gut barrier, and the establishment of the gut microbiota provides the critical maturational signals. Their interaction is particularly apparent during the postnatal development, when major maturational events occur in the gut-associated lymphoid tissue.

Infections including acute diarrhoea

The most fully documented probiotic intervention is the nutritional management of subjects with acute infectious diarrhoea. Rotavirus is recognised as the leading cause of these infections in children. Rotaviruses invade the highly differentiated absorptive columnar cells of the small intestinal epithelium, where they replicate causing defective sodium and chloride transport. The invasion results in partial disruption of the intestinal mucosa with loss of microvilli and decrease in the villus/crypt ratio, and diarrhoea is mainly due to a failure of the epithelium to differentiate during rapid migration to repair the disruption. In addition, increased intestinal permeability and aberrant absorption of intraluminal antigens ensue during gastrointestinal infection also the gut microbiota balance is disturbed. Rotavirus diarrhoea is associated with an increased concentration of faecal urease, an inflammatory mediator which predisposes the gut mucosa to further ammonia-induced damaging effects and to the overgrowth of urease-producing bacteria.

The current accepted guidelines for treatment of acute diarrhoea are based on correcting the dehydration by oral rehydration solutions. In addition, immediately after the completion of oral rehydration full feedings of previously tolerated diet can be reintroduced. Well-controlled clinical studies have shown that specific strains of probiotics can modulate gut barrier functions in rotavirus diarrhoea and thus constitute a safe adjunct nutritional management. A multicentre study by the European Society for Paediatric Gastroenterology, Hepatology and Nutrition working group tested the clinical efficacy and safety of a probiotic administered in an oral...
rehydration solution [290]. In rotavirus diarrhoea, a significant decrease of episodes was observed, while in non-specific or bacterial diarrhoea no clear effect was found. A recent meta-analysis confirmed the efficacy of a variety of probiotics in acute diarrhoea in children in reducing duration of symptoms among children less than 5 years with acute, nonbacterial diarrhoea [291]. The estimated benefit was approximately one day.

On this basis, specific probiotic strains have preventive potential, in reducing the risk of acquisition of viral gastroenteritis. Saavedra et al. [292] conducted a double-blind, placebo-controlled trial in hospitalised infants randomised to receive a standard infant formula or the same formula supplemented with *Bifidobacterium"* and *Streptococcus thermophilus*. Altogether, 31% of the patients given the standard infant formula, but only 7% of those receiving the probiotic-supplemented formula developed diarrhoea during a 17-month follow-up, and the prevalence of rotavirus shedding was significantly lower in those receiving probiotic-supplemented formula. Probiotic supplementation resulted in a significant decrease in the incidence of diarrhoea in undernourished non-breast-fed Peruvian children followed-up for 15 months [293]. Szajewska and colleagues [294] evaluated the efficacy of orally administered *Lactobacillus rhamnosus* GG in the prevention of nosocomial diarrhoea in young children. Eighty-one children aged 1 to 36 months, who were hospitalised for reasons other than diarrhoea, were enrolled in a randomised double-blind trial to receive probiotics or placebo for the duration of their hospital stay. *Lactobacillus* GG reduced the risk of nosocomial diarrhoea in comparison with placebo; 6.7% versus 33.3%; relative risk 0.2 (95% CI 0.06–0.6). The prevalence of rotavirus infection was similar in probiotic and placebo groups, while the risk of developing rotavirus gastroenteritis was reduced by probiotic preparation. Recently, Mastretta and colleagues [295] assessed the effect of probiotics (*Lactobacillus GG*) and breast-feeding on nosocomial rotavirus infections in 220 infants hospitalised during one rotavirus epidemic season. The incidence of nosocomial rotavirus infections was 27.7%. The probiotic preparation was ineffective, whereas breast-feeding was effective in reducing the risk of nosocomial rotavirus infection.

The beneficial effect in diarrhoea by probiotics has been explained by reduction in the duration of rotavirus shedding and in increased gut permeability caused by rotavirus infection, together with a significant increase in IgA-secreting cells to rotavirus [296]. A lactobacillus strain isolated from human microbiota was recently shown to counter cellular damage associated with a diarrhoeagenic pathogen, thus corroborating reports of a normalising of intestinal permeability by selected probiotics [297]. The principal effect of probiotics, however, is characterised by stabilisation of the gut microflora [143].

The value of probiotic preparations for reducing the risk of traveller’s diarrhoea has been studied, but the results have been conflicting, due to differences in probiotic species and vehicles used, in dosage schedule, as well as in travel destinations in which the studies have been conducted. Recent double-blind, placebo-controlled studies indicate, however, that there is evidence that some probiotic species may provide protection against traveller’s diarrhoea [298].

### Antibiotic associated diarrhoea (AAD)

Preventing AAD by giving probiotic microorganisms also exploits the idea that probiotic species can strengthen colonisation resistance to pathogen growth in the gut. AAD occurs in over 20% of adult patients who receive antibiotics and is particularly common in the elderly, during use of multiple antibiotic regimes and in enterally fed patients. It is variably associated with the presence of toxin producing *Clostridium difficile*.

There are over 40 published reports on this subject [299] dating back to 1975 [300], and which are summarised in several recent reviews [301–304]. Of these reports, nine are randomised double-blind placebo controlled trials and are the subject of two recent meta-analyses [299, 305]. These reports show that probiotic organisms are a safe, relatively cheap and effective preventive strategy against AAD. Both meta-analyses show a similar, significant, overall reduction in relative risk in favour of probiotics of 0.39 (0.25–0.62 95% confidence interval, CI) [305] and 0.40 (0.27–0.57 CI) [299].

Antimicrobial treatment disturbs the colonisation resistance of the gut microbiota, which may induce symptoms, most frequently diarrhoea. The preventive potential of probiotics on AAD in children has also been shown [306]. To avoid confusion caused by recent antimicrobial treatments, the incidence of diarrhoea after a single antimicrobial treatment and the effect of probiotics was evaluated in children with no history of antimicrobial use during the previous 3 months. The incidence of diarrhoea was 5% in the group given *Lactobacillus rhamnosus* GG and 16% in the placebo group, substantiating the efficacy of the probiotics approach. In a further study, Vanderhoof and colleagues [307] studied 188 children aged 6 months to 10 years who were given oral antibiotics in an outpatient setting. *Lactobacillus rhamnosus* GG as compared to placebo significantly reduced stool frequency and increased stool consistency during antibiotic therapy. In contrast, the same preparation did not reduce the rate of occurrence of diarrhoea compared to placebo in adult patients taking antibiotics initially administered in a hospital setting [308].

Whilst this is one of the best-demonstrated benefits of probiotics, many questions are unanswered. Several
different probiotics have been used, principally from the genera Bifidobacteria and Lactobacillus and the yeast Saccharomyces, but the choice of species is probably determined by their known safety and history of use rather than the result of a systematic search for beneficial properties amongst the many possible candidate probiotic species that normally inhabit the gut. The design of trials of probiotics for AAD also raises questions that are as yet unanswered about the ideal dose, its timing and duration and whether combinations of probiotics are better than single species. There are no reports of safety issues from controlled trials. Case reports of septicemia have been reported [309], however, these mostly occurred in patients who were immunosuppressed [305]. Clinically AAD is a very diverse disease without a clear pathogenesis, except in the minority of cases with C. difficile endotoxin, and can vary from mild diarrhoea to life threatening pseudomembranous colitis. It can be treated with metronidazole or vancomycin and in some trials the probiotic has been combined with an antibiotic [310–312].

Despite the lack of understanding of how probiotics work in AAD, and the pleomorphic nature of the condition, clear benefits are evident. Two groups have reported the use of prebiotics for AAD [313, 314].

### Inflammatory disorders

Inflammation is frequently accompanied by imbalance in the intestinal microbiota. A strong inflammatory response may then be mounted to microbiota bacteria, leading to perpetuation of the inflammation and gut barrier dysfunction [315]. Duchmann and associates [316] have demonstrated that healthy individuals are tolerant to their own microbiota, and that such tolerance is abrogated in patients with inflammatory bowel disease. An altered gut microbiota is reported in patients with rheumatoid arthritis [317] and allergic disease [262, 318] implying that the host-microbe interaction in the gut responds to inflammation in the gut and elsewhere in the human body.

Probiotic bacteria may counteract the inflammatory process by stabilising the gut microbial environment and the intestine’s permeability barrier, and by enhancing the degradation of enteral antigens and altering their immunogenicity. Another explanation for the gut-stabilising effect could be improvement of the intestine’s immunological barrier, particularly intestinal IgA responses. Probiotic effects in reducing the risk of disease may also be mediated via control of the balance between pro- and anti-inflammatory cytokines [315].

An increasing number of clinical and experimental studies demonstrate the importance of constituents within the intestinal lumen, in particular the resident microbiota, in driving the inflammatory responses in Crohn’s disease and ulcerative colitis. Specific infectious agents or antigens initiating or perpetuating the inflammation have not been identified. Recent reports have revealed a possible link between intestinal microbiota and inflammatory bowel disease. Consequently, modifying the intestinal immunological milieu by specific strains of the gut microbiota may be seen as one target of prophylactic intervention. Most evidence on a role of aberrant gut microbiota in inflammatory bowel disease, however, derives from experimental animal models. Transgenic mice with targeted deletion of the T cell receptor spontaneously develop colitis in response to the gut microbiota [319]. If organised GALT is removed from the mice by appendectomy at neonatal age, but not later, tolerance to gut microbiota with no colitis develops, indicating that the initial colonisation pattern may determine subsequent immunological processes.

Preliminary reports have shown benefit in reversing some of the immunological disturbances characteristic of Crohn’s disease [320]. In addition, reduction in disease activity and increased intestinal permeability has been achieved in paediatric patients with Crohn’s disease by probiotic intervention [321]. In adults operated for the condition, however, a probiotic preparation failed to reduce the risk of endoscopic recurrence during one-year follow-up [322]. An interesting recent study demonstrates that intervention with a non-pathogenic E. coli aids in maintaining remission in ulcerative colitis [323].

### Allergic disease

The rise in prevalence of atopic diseases has reached epidemic proportions in the industrialised societies. The phenomenon cannot be explained by genetic factors only nor by novel emerging allergens sensitising the host in addition to the traditional dietary allergens and aeroallergens. Two candidate explanatory factors relating to the modern Western lifestyle stem from altered hygiene and nutrition [315].

The hygiene hypothesis of allergy conceives the rapid increase in atopy to be related to reduced exposure to microbes at an early age. The earliest and most massive source of such exposure is associated with the establishment of the gut microbiota. Indeed, the initial compositional development of the gut microbiota is considered a key determinant in the development of normal gut barrier functions and healthy host-microbe interactions.

Initial signals to counter IL-4 and thereby IgE and atopy, and IL-5-generated eosinophilic inflammation may stem from components of the innate immunity (Figs. 2 and 4). The Th1-promoting potential of the healthy gastrointestinal microbiota is possibly associated with two structural components of bacteria, the
LPS portion of gram-negative bacteria (endotoxin) and a specified CpG motif in bacterial DNA [324, 325]. These structures (PAMPs) activate immunomodulatory genes via PRRs such as TLRs present, e.g. on macrophages, DCs and intestinal epithelial cells [326, 327]. LPS binds to TLR4 on the cell surface, whereas the CpG motif must be taken up into the cell by endocytosis before binding to TLR9 [328, 329]. As reported in an important paper by Neish and colleagues [233], non-pathogenic microbes elicit an immunosuppressive effect on intestinal epithelial cells by inhibition of the transcription factor NF-κB pathway thus demonstrating a direct anti-inflammatory effect. In addition, specific strains of the gut microbiota have been shown to contribute to a Treg cell population (Fig. 2) amenable to oral tolerance induction [330], and to counter allergy by generation of anti-inflammatory IL-10 and TGF-β [331, 332].

Results obtained from a murine model indicate that animals kept in germ-free conditions exhibit impaired development of the intestinal immune system resulting in Th2-skewed immune responsiveness with abrogation of oral tolerance [330]. Interestingly, reconstitution of the intestinal microbiota with bifidobacteria at the neonatal stage, but not at a later stage, was shown to restore the susceptibility to oral tolerance. In one prospective clinical study, intestinal microbiota from 76 infants at high risk of atopic diseases were analysed at 3 weeks and 3 months of age by conventional bacterial cultivation and two culture-independent methods, gas-liquid chromatography of bacterial cellular fatty acids and quantitative fluorescence in situ hybridisation of bacterial cells [252]. Positive skin prick reaction at 12 months was observed in 22/76 (29%) children. At 3 weeks of age the bacterial cellular fatty acid profile in faecal samples differed significantly between infants developing and not developing atopy. Upon fluorescence in situ hybridisation atopics were found to have more clostridia and tended to have fewer bifidobacteria in their stools than non-atopics. Differences in the neonatal gut microbiota were thus shown to precede the development of atopy, suggesting a crucial role of the balance of indigenous intestinal bacteria for the maturation of human immunity to a nonatopic mode. Similarly, a recent study utilising traditional plate culture methods showed that children who developed allergy were less often colonised with bifidobacteria during their first year of life as compared to those who did not develop allergic disease [333].

Preliminary studies have revealed that cow's milk casein, a common allergen in cow's milk allergy, hydrolysed with Lactobacillus rhamnosus GG -derived enzymes suppresses lymphocyte proliferation and, more specifically, production of allergenic IL-4 in vitro [334, 335]. In experiments conducted by von der Weid and colleagues [336], Lactobacillus paracasei inhibited proliferation of T cells and reduced secretion of both Th1 and Th2 cytokines whilst inducing the development of a population of T<sub>reg</sub> cells producing TGF-β and IL-10 reminiscent of tolerogenic Th3 cells (Fig. 2). Furthermore, a strain of Lactobacillus casei has been demonstrated to suppress IgE responses and systemic anaphylaxis in a murine model of food allergy [337]. According to a recent paper by [338], lactobacilli inhibited allergen-specific Th2 cytokine production by peripheral blood mononuclear cells from allergic individuals. However, different strains of lactobacilli appear to induce distinct and even opposing responses in murine DCs and thus specific strains of the gut microbiota and probiotics may play a crucial role in determining the Th1/Th2/Th3-driving capacity of intestinal DC. In parallel, recent observations indicate that the cytokine production patterns induced by intestinal bifidobacteria are strain-specific [318]. The results of clinical studies evaluating the effects of probiotics in allergic disease tend to substantiate this suggestion.

There are data on record concerning the beneficial effects of probiotics in subjects with established allergic disorders. In a double-blind placebo-controlled trial of infants with atopic eczema, infant formula supplemented with the probiotics, either Lactobacillus GG or Bifidobacterium lactis Bb-12 resulted in significant improvement of the skin condition along with a decrease in markers of systemic allergic inflammation as compared to infants receiving unsupplemented formula [339]. Furthermore, in infants with cow's milk allergy and atopic eczema, extensively hydrolysed formula fortified with probiotics was shown to alleviate intestinal inflammation associated with disorder [340]. In adults, a milk challenge in conjunction with a probiotic strain reduced the immunoinflammatory response characteristic of the challenge without probiotics [341].

In the first double-blind, placebo-controlled clinical trial, Lactobacillus GG administered prenatally and during the first months of life resulted in a significant reduction in the prevalence of atopic eczema in at-risk infants [263]. However, no risk reduction was observed for atopy, as characterised by antigen-specific IgE generation. Furthermore, the prevalence of food allergy was comparable in infants receiving probiotics and placebo. In a subgroup of this study population in whom the probiotics were administered to the lactating mother, probiotic supplementation increased the concentration of TGF-β in breast milk [332], furnishing one mechanism by which the risk of infant atopic eczema could be reduced.

Taking the accruing data on probiotics and allergy together, it may be concluded that whilst Lactobacillus GG appears to be effective in reducing the risk of atopic eczema, it offers no protection from food allergy even though it alleviates inflammation of already established food allergy. These findings demonstrate the heterogeneous nature of allergic disorders on one hand and the strain-specificity of the probiotic effects on the other.
Suggested markers for use in clinical studies of acute diarrhoeal disorders, inflammatory bowel disease and allergy are listed in Table 20.

**Conclusion**

The gut and the immune system have many functions. When working optimally, they provide efficient digestion and absorption of nutrients, ensure the safe detoxification and excretion of both environmental and endogenously created toxins and mutagens, and provide a competent, lifelong protection and defence against allergens and invading organisms. The very large number of commensal bacteria in the gut play an important part in digestion and help to shape, monitor and maintain the body’s defences. Optimal function of these integrated systems is not defined and wide variation is seen during physiological testing in healthy subjects. To substantiate a claim that a food or nutrient maintains healthy digestion or strengthens the body’s natural resistance is, therefore, difficult. It is still more challenging to substantiate claims for enhanced function or an improved sense of well being.

As a first step, in what is an important process of optimising health, we have defined normal bowel habit and transit time, identified their role as risk factors for disease and described how they may be measured. Similarly, we have tried to define what is a healthy gut flora in terms of the dominant genera and their metabolism and listed the many, varied and novel methods for determining these parameters.

It has proved less easy to put boundaries on what constitutes optimal or improved gastric emptying, gut motility, nutrient and water absorption and the function of organs such as the liver, gallbladder and pancreas. The many tests of these functions are described (and also in [27]).

Similarly, we have discussed gastrointestinal well being. Sensations arising from the gut can be both pleasant and unpleasant. However, the characteristics of well being are ill defined and merge imperceptibly from acceptable to unacceptable, a state that is also subjective. Nevertheless, we feel this is an important area for future work and method development.

The immune system is even more difficult to make quantitative judgements about. When it is defective, then clinical problems ensue, but this is an uncommon state. The two arms of the immune system, innate and adaptive, work synergistically together and comprise many cellular and humoral factors. The adaptive system is extremely sophisticated and between the two arms of immunity there is great redundancy, which provides robust defences. New aspects of normal immune function are discovered regularly. It is not clear whether normal immune function can be “improved”.

Measuring aspects of immune function is possible. We have given information about at least 50 possible

<table>
<thead>
<tr>
<th>Effect</th>
<th>Method of assessment</th>
<th>Identification of markers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nutritional management of</td>
<td>Randomised double-blind clinical study</td>
<td>Duration of symptoms</td>
</tr>
<tr>
<td>diarrhoea</td>
<td>(Determination of the sample size appropriate for the condition)</td>
<td>Eradication of the infectious agent</td>
</tr>
<tr>
<td>allergic/inflammatory disease</td>
<td></td>
<td>Need of symptomatic therapy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Symptom score</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Well being: negative challenge/provocation tests,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>macroscopic, microscopic evaluation, indirect indices</td>
</tr>
<tr>
<td></td>
<td></td>
<td>specific for the condition</td>
</tr>
<tr>
<td>Alleviation of disease activity/reactions/</td>
<td>Clinical follow-up studies</td>
<td>Activity indices specific for the condition</td>
</tr>
<tr>
<td>relapses/inflammation</td>
<td>Cross-over challenge studies (double-blind placebo-</td>
<td>Symptom score (e.g. IBD, arthritis, atopic eczema)</td>
</tr>
<tr>
<td></td>
<td>controlled)</td>
<td>Proinflammatory cytokines specific for the condition and site</td>
</tr>
<tr>
<td></td>
<td>(Determination of the sample size appropriate for the</td>
<td>(TNF-α, IFN-γ, IL-4, IL-5, IL-13)</td>
</tr>
<tr>
<td></td>
<td>condition)</td>
<td>Endoscopy, histology, local inflammation</td>
</tr>
<tr>
<td>Enhanced host defence</td>
<td>Intestinal permeability</td>
<td>Dual sugar permeability tests, PEGs, perfusion,</td>
</tr>
<tr>
<td></td>
<td>Immunomodulation in vitro/in vivo</td>
<td>macromolecular transport, IgA antibodies,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Antigen-specific antibody, markers of innate immunity</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Anti-inflammatory cytokines (IL-10, TGF-β) and</td>
</tr>
<tr>
<td></td>
<td></td>
<td>mediators</td>
</tr>
<tr>
<td>Reduction in risk of disease</td>
<td>Randomised double-blind placebo-controlled study</td>
<td>Prevalence of the condition after appropriate follow-up</td>
</tr>
<tr>
<td>diarrhoea</td>
<td>(Determination of the sample size appropriate for the</td>
<td>Early markers of sensitisation, immune deviation</td>
</tr>
<tr>
<td>allergic/inflammatory disease</td>
<td>condition and population: general vs. at-risk)</td>
<td>In vitro assessment of host-microbe interaction,</td>
</tr>
<tr>
<td>Gut microbiota stabilisation</td>
<td>Gut microbiota aberrancy assessment</td>
<td>Characterisation of the healthy vs. aberrant</td>
</tr>
<tr>
<td></td>
<td>Modern techniques of evaluation of the gut microecology</td>
<td>microbiota appropriate for the age</td>
</tr>
<tr>
<td></td>
<td>(Site of sampling appropriate for the condition)</td>
<td></td>
</tr>
</tbody>
</table>

Table 20  Methodology in studies of infection and atopy
tests. However, there is no one test that will define either the status or functional capacity of the immune system. Human studies are often limited by the ability to sample only blood or secretions such as saliva but it should be remembered that only 2% of lymphocytes circulate at any given time, which limits interpretation of data.

There are normal ranges for immune cell numbers and Ig concentrations but none for other immune cell functional responses. In the light of this, we recommend assessing the functional capacity of the immune system by

a) measuring specific cell functions \textit{ex vivo},

b) measuring \textit{in vivo} responses to challenge, e.g. change in antibody levels in peripheral blood or response to antigens,

c) determining the incidence and severity of infection in target populations during naturally occurring episodes or in response to attenuated pathogens.

References


130. Petrakis NL, King EB (1981) Cytological abnormalities in nipple aspirates of breast fluid from women with severe constipation. Lancet ii:1203–1205
233. Neish AS, Gewirtz AT, Zeng H, Youn JH Cummings et al II/169


